ELSEVIER

HANDBOOK OF GLOBAL ANALYSIS

Edited by

DEMETER KRUPKA
DAVID SAUNDERS

Handbook of Global Analysis

This page intentionally left blank

Handbook of Global Analysis

Edited by

Demeter Krupka and David Saunders
Palacky University, Olomouc, Czech Republic

ELSEVIER

Amsterdam • Boston • Heidelberg • London • New York • Oxford
Paris • San Diego • San Francisco • Singapore • Sydney • Tokyo

Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
Linacre House, Jordan Hill, Oxford OX2 8DP, UK

First edition 2008

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-0-444-52833-9

Printed and bound in The Netherlands

08 09 10 11 12 10 9 8 7 6 5 4 3 2 1

Preface

The group of topics known broadly as ‘Global Analysis’ has developed considerably over the past twenty years, to such an extent that workers in one area may sometimes be unaware of relevant results from an adjacent area. The many variations in notation and terminology add to the difficulty of comparing one branch of the subject with another.

Our purpose in preparing this Handbook has been to try to overcome these difficulties by presenting a collection of articles which, together, give an overall survey of the subject. We have been guided in this task by the MSC2000 classification, and so the scope of the Handbook may be described by saying that it covers the 58-XX part of the classification: ranging from the structure of manifolds, through the vast area of partial differential equations, to particular topics with their own distinctive flavour such as holomorphic bundles, harmonic maps, variational calculus and non-commutative geometry. The coverage is not complete, but we hope that it is sufficiently broad to provide a useful reference for researchers throughout global analysis, and that it will also be of benefit to mathematical physicists and to PhD and post-doctoral students in both areas.

The main work involved in the preparation of the Handbook has, of course, been that of the authors of the articles, who have carried out their task with skill and professionalism. Our debt to them is immediate and obvious. Some other potential authors have, for personal reasons, been unable to offer contributions to the Handbook, but we hope that those omissions will not detract too much from its value. The editors also wish to acknowledge the assistance of Petr Volný in the formatting of the LaTeX manuscripts, and of Andy Deelen, Kristi Green and Simon Pepping at Elsevier for their help and advice during the preparation of the book. In addition, we should like to record our particular thanks to Arjen Sevenster from Elsevier, who commissioned the project and gave us support and encouragement during its development.

We should finally like to acknowledge the support of the Czech Science Foundation (grant 201/06/0922) and the Czech Ministry of Education, Youth, and Sports (grant MSM 6198959214) for our work on this project.

Demeter Krupka
David Saunders

Palacký University, Olomouc.

This page intentionally left blank

Contents

Demeter Krupka and David Saunders

Global aspects of Finsler geometry[1]

Tadashi Aikou and László Kozma

Contents

1 Finsler metrics and connections

1.1 Finsler metrics

Let $\pi : TM \to M$ be the tangent bundle of a connected smooth manifold M of $\dim M = n$. We denote by $v = (x, y)$ the points in TM if $y \in \pi^{-1}(x) = T_xM$. We denote by $z(M)$ the zero section of TM, and by $TM^\times$ the slit tangent bundle $TM \backslash z(M)$. We introduce a coordinate system on TM as follows. Let $U \subset M$ be an open set with local coordinate $(x^1, \cdots, x^n)$. By setting $v = \sum y^i \left(\partial/\partial x^i\right)_x$ for every $v \in \pi^{-1}(U)$, we introduce a local coordinate $(x, y) = (x^1, \cdots, x^n, y^1, \cdots, y^n)$ on $\pi^{-1}(U)$.

Definition 1.1 A function $F : TM \longrightarrow \mathbb{R}$ is called a *Finsler metric* on M if

(1) $F(x, y) \geq 0$, and $F(x, y) = 0$ if and only if $y = 0$,

(2) $F(x, \lambda y) = \lambda F(x, y)$ for $\forall \lambda \in \mathbb{R}^+ = \{\lambda \in \mathbb{R} :\ \lambda > 0\}$,

(3) $F(x, y)$ is smooth on $TM^\times$, the out-side of the zero section,

[1] Tadashi Aikou: work supported in part by Grant-in-Aid for Scientific Research No. 17540086(2006), The Ministry of Education, Science Sports and Culture. László Kozma: partially supported by the Hungarian Scientific Research Fund OTKA T048878.

(4) $G = F^2/2$ is *strictly convex* on each tangent space T_xM, that is, the Hessian (G_{ij}) defined by

$$G_{ij}(x,y) = \frac{\partial^2 G}{\partial y^i \partial y^j} \tag{1.1}$$

is positive-definite,

are satisfied. The pair (M, F) is called a *Finsler manifold*.

We note that the last condition in this definition is equivalent to the convexity of the unit ball $B_x = \{y \in T_xM \mid F(x,y) \leq 1\}$.

If a Finsler metric F is defined, then the norm $\|y\|$ of each $y \in T_xM$ is defined by $\|y\| = F(x,y)$, and the length $s(t)$ of a smooth curve $c(t) = (x^1(t), \cdots, x^n(t))$ is defined by $s(t) = \int_0^1 \|\dot{c}(t)\| dt = \int_0^1 F(x(t), \dot{x}(t))\, dt$.

Example 1.2 *(Funk metric)* Let g be a Riemannian metric on M. We define $\alpha : TM \to \mathbb{R}$ by $\alpha(v) = \sqrt{g(v,v)}$. Since α is convex, there exists a 1-from β such that $\beta(v) \leq \alpha(v)$. The function $F = \alpha + \beta$ defines a convex Finsler metric on M so-called *Randers metric*. We shall review a typical example of Randers metric (see [37] or [14]). Let $\mathbb{R}^n$ be an n-dimensional Euclidean space with the standard coordinate $(x^1, \cdots, x^n)$, and $\mathbb{B}$ the unit ball centered the origin: $\mathbb{B} = \{x \in \mathbb{R}^n \mid \phi(x) = 1 - \|x\|^2 > 0\}$. The Riemannian metric g_H defined by

$$g_H = \frac{(1 - \|x\|^2)\left(\sum dx^i\right)^2 + \left(\sum x^i dx^i\right)^2}{(1 - \|x\|^2)^2}$$

is called the *Hilbert metric* on $\mathbb{B}$. We define a 1-form β by

$$\beta = \frac{\sum x^i dx^i}{1 - \|x\|^2} = -\frac{1}{2} d \log \phi.$$

The norm $\|\beta\|_H$ of β with respect to g_H is given by $\|\beta(x)\|_H = \|x\| < 1$, and thus the function F on $\mathbb{B}$ defined by $F(v) = \sqrt{g_H(v,v)} + \beta(v)$ is a Finsler metric called the *Funk metric* on $\mathbb{B}$. We note that the relation between g_H and F is given by

$$\|v\|_H = \frac{1}{2}\left[F(v) + F(-v)\right]$$

for all $v \in TM$.

For the differential π_* of the submersion $\pi : TM^\times \longrightarrow M$, the *vertical subbundle* V of $T(TM^\times)$ is defined by $V = \ker \pi_*$, and V is locally spanned by $\{\partial/\partial y^1, \cdots, \partial/\partial y^n\}$ on each $\pi^{-1}(U)$. Then it induces the exact sequence

$$0 \longrightarrow V \xrightarrow{\ i\ } T(TM^\times) \xrightarrow{\ \pi_*\ } \widetilde{TM} \longrightarrow 0, \tag{1.2}$$

where $\widetilde{TM} = \{(y,v) \in TM^\times \times TM \,\big|\, v \in T_{\pi(y)}M\}$ is the pull-back bundle π^*TM.

$$\begin{array}{ccc} \widetilde{TM} & \longrightarrow & TM \\ \big\downarrow & & \big\downarrow \pi \\ TM^\times & \underset{\pi}{\longrightarrow} & M \end{array}$$

Since the natural local frame field $\{\partial/\partial x^i\}_{i=1,\cdots,n}$ on each U is identified with the one of $\widetilde{TM}$ on $\pi^{-1}(U)$, any section X of $\widetilde{TM}$ is written in the form $X = \sum(\partial/\partial x^i) \otimes X^i$ for smooth functions X^i on each $\pi^{-1}(U)$. Furthermore, since $\ker \pi_* = V$, the differential π_* is given by $\pi_* = \sum(\partial/\partial x^i) \otimes dx^i$.

We define a metric G on the bundle $\widetilde{TM}$ by

$$G(X,Y) = \sum G_{ij} X^i Y^j \tag{1.3}$$

for every section $X = \sum(\partial/\partial x^i) \otimes X^i$ and $Y = \sum(\partial/\partial x^j) \otimes Y^j$. We also set

$$C_{ijk} = \frac{1}{2}\frac{\partial g_{ij}}{\partial y^k} = \frac{1}{4}\frac{\partial^3 L^2}{\partial y^i\,\partial y^j\,\partial y^k}.$$

Then we define a symmetric tensor field $C : \otimes^3 \widetilde{TM} \to \mathbb{R}$ by

$$C(X,Y,Z) = \sum C_{ijk} X^i Y^j Z^k \tag{1.4}$$

for all sections X, Y, Z of $\widetilde{TM}$. It is trivial that C vanishes identically if and only if G is a Riemannian metric on M. This tensor field C is called the *Cartan tensor field*.

In the sequel, we use the notation $A^0(\widetilde{TM})$ for the space of smooth sections of $\widetilde{TM}$. Since $\widetilde{TM}$ is naturally identified with $V \cong \ker \pi_*$, any section X of $\widetilde{TM}$ is considered as a section of V. We denote by X^V the section of V corresponding to $X \in A^0(\widetilde{TM})$:

$$A^0(\widetilde{TM}) \ni X = \sum \frac{\partial}{\partial x^i} \otimes X^i \quad \Longleftrightarrow \quad \sum \frac{\partial}{\partial y^i} \otimes X^i := X^V \in A^0(V).$$

The following is trivial since (1.2) is exact.

$$\pi_*(X^V) = 0 \tag{1.5}$$

for every $X \in A^0(\widetilde{TM})$.

The multiplier group $\mathbb{R}^+ \cong \{cI \in GL(n,\mathbb{R});\ c \in \mathbb{R}^+\} \subset GL(n,\mathbb{R})$ acts on the total space by multiplication

$$m_\lambda : TM^\times \ni v = (x,y) \to \lambda v = (x, \lambda y) \in TM^\times$$

for every $\lambda \in \mathbb{R}^+$. This action induces a canonical section $\mathcal{E}$ of V defined by $\mathcal{E}(v) = (v,v)$ for all $v \in TM^\times$. By the homogeneity of F, we have

$$\mathcal{E}(F) = \frac{d}{dt}\Big|_{t=0} F(x, y + t\mathcal{E}) = F.$$

We shall consider $\mathcal{E}$ as a section of $\widetilde{TM}$, and we denote it by the same notation $\mathcal{E}$, that is, $\mathcal{E}(x,y) = \sum(\partial/\partial x^i) \otimes y^i$. This section $\mathcal{E}$ is called the *tautological section* of $\widetilde{TM}$. Then it is easily shown that $F = \sqrt{G(\mathcal{E},\mathcal{E})}$ and

$$C(\mathcal{E}, \bullet, \bullet) \equiv 0. \tag{1.6}$$

1.2 Ehresmann connection

For the submersion $\pi : TM^\times \to M$, the vertical subbundle is defined by $V = \ker \pi_*$, while the *horizontal subbundle* H is defined by a subbundle $H \subset T(TM^\times)$ which is complementary to V. These subbundles give a smooth splitting

$$T(TM^\times) = H \oplus V. \tag{1.7}$$

Although the vertical subbundle V is uniquely determined, the horizontal subbundle is not canonically determined. An *Ehresmann connection* of the submersion $\pi : TM^\times \to M$ is a selection of horizontal subbundles. In this report, we shall define this as follows.

Definition 1.3 An *Ehresmann connection* of the submersion $\pi : TM^\times \to M$ is a bundle morphism $\theta : T(TM^\times) \to \widetilde{TM}$ satisfying

$$\theta(X^V) = X \tag{1.8}$$

for every $X \in A^0(\widetilde{TM})$.

If an Ehresmann connection θ is given, then a horizontal subbundle H is defined by $H = \ker \theta$. In this report, we shall assume that the subbundle H defined by θ is invariant by the action $m_\bullet$, that is, $(m_\lambda)_* H = H \circ m_\lambda$ for all $\lambda \in \mathbb{R}^+$. This assumption is equivalent to

$$\mathcal{L}_\mathcal{E} H \subset H. \tag{1.9}$$

Remark 1.4 A *linear connection* of the tangent bundle TM is a selection of horizontal subbundles in $GL(n, \mathbb{R})$-invariant way. Thus, an Ehresmann connection θ in our sense is sometimes called a *non-linear connection* of TM.

In the sequel, we denote by A^k and $A^k(\widetilde{TM})$ the space of smooth k-forms and $\widetilde{TM}$-valued k-form on $TM^\times$ respectively. We suppose that an Ehresmann connection θ is given. Then, the exterior differential $d : A^k \to A^{k+1}$ is decomposed into the form $d = d^H \oplus d^V$ according to the decomposition (1.7), where d^H is the differential along H and d^V is the one along V. If a covariant derivation $D : A^0(\widetilde{TM}) \to A^1(\widetilde{TM})$ of the bundle $\widetilde{TM}$ is also decomposed into the form $D = D^H \oplus D^V$.

Proposition 1.5 *If an Ehresmann connection θ is given, then there exists a covariant exterior derivation D of $\widetilde{TM}$ satisfying*

$$\theta = D\mathcal{E}, \tag{1.10}$$

or equivalently

$$D^H \mathcal{E} = 0. \tag{1.11}$$

Proof We define a covariant derivation D by $D^V = d^V$ and $D^H_X Y = \theta[X^H, Y^V]$ for all $X, Y \in A^0(\widetilde{TM})$. It is easily shown that $D = D^H \oplus d^V$ is a covariant derivation on $\widetilde{TM}$. Then we have

$$D^V_X \mathcal{E} = X^V(\mathcal{E}) = X^V = \theta(X)$$

and, from (1.9) we obtain

$$D^H_X \mathcal{E} = \theta[X^H, \mathcal{E}] = -\theta(\mathcal{L}_\mathcal{E} X^H) = 0.$$

Therefore we obtain (1.11). □

Since we are concerned with the tangent bundle, $\widetilde{TM}$ is also naturally identified with the horizontal subbundle H, and any section X of $\widetilde{TM}$ is considered as a section of H. We denote by X^H the section of H corresponding to $X \in A^0(\widetilde{TM})$:

$$A^0(\widetilde{TM}) \ni X = \sum \frac{\partial}{\partial x^i} \otimes X^i \quad \Longleftrightarrow \quad \sum \frac{\delta}{\delta x^i} \otimes X^i := X^H \in A^0(H),$$

where

$$\left\{ \frac{\delta}{\delta x^1} = \left(\frac{\partial}{\partial x^1} \right)^H, \cdots, \frac{\delta}{\delta x^n} = \left(\frac{\partial}{\partial x^n} \right)^H \right\}$$

denotes the horizontal lift of natural local frame field $\left\{\partial/\partial x^1, \cdots, \partial/\partial x^n\right\}$ with respect to the given Ehresmann connection θ. The set $\{dx^1, \cdots, dx^n\}$ is the dual basis of H^*. For the two bundle morphism π_* and θ from $T(TM^\times)$ onto $\widetilde{TM}$, we have

Proposition 1.6 *The bundle morphisms π_* and θ satisfy*

$$\pi_*(X^H) = X, \quad \pi_*(X^V) = 0 \tag{1.12}$$

and

$$\theta(X^H) = 0, \quad \theta(X^V) = X \tag{1.13}$$

for every $X \in A^0(\widetilde{TM})$.

1.3 Chern connection

If a Finsler metric F is given on TM, then there exists a natural metric G on $\widetilde{TM}$ defined by (1.3). Then we shall introduce a covariant derivation ∇ which satisfies some natural axioms.

For a given covariant derivation ∇ on $\widetilde{TM}$, we always define an Ehresmann connection $\theta : T(TM^\times) \to \widetilde{TM}$ by

$$\theta = \nabla \mathcal{E}. \tag{1.14}$$

With respect to the splitting (1.7), ∇ is also decomposed into the form $\nabla = \nabla^H \oplus \nabla^V$.

Definition 1.7 ([10]) The *Chern connection* on (M, F) is a covariant exterior differentiation $\nabla : A^k(\widetilde{TM}) \to A^{k+1}(\widetilde{TM})$ uniquely determined from the following conditions.

(1) ∇ is *symmetric*:

$$\nabla \pi_* = 0, \tag{1.15}$$

where we considered $\pi_* = \sum(\partial/\partial x^i) \otimes dx^i$ as a section of $A^1(\widetilde{TM})$.

(2) ∇ is *almost G-compatible*:

$$\nabla^H G = 0. \tag{1.16}$$

Remark 1.8 In the case of $C = 0$, the metric F is the norm function of a Riemannian metric g, and the Chern connection ∇ is given by $\nabla = \pi^*\nabla^M$ for the Levi-Civita connection ∇^M of (M, g). The Chern connection is also called the *Rund connection* of (M, F) (cf. [3], [12] [32]).

We can easily show that θ defined by (1.14) is invariant by the natural action $m_\bullet$ of $\mathbb{R}^+$. In local coordinate, θ is given by

$$\theta = \nabla\left(\sum \frac{\partial}{\partial x^i} \otimes y^i\right) = \sum \frac{\partial}{\partial x^i} \otimes \left(dy^i + \sum \omega^i_j y^j\right),$$

where ω^i_j is the connection forms of ∇ with respect to $\left\{\partial/\partial x^1, \cdots, \partial/\partial x^n\right\}$. The set $\{\theta^i, \cdots, \theta^n\}$ of 1-forms defined by $\theta^i = dy^i + \sum \omega^i_j y^j$ $(i = 1, \cdots, n)$ is the dual basis of V^* defined by θ.

Then the covariant derivative ∇ is also decomposed into the form $\nabla = \nabla^H \oplus \nabla^V$, where $\nabla^H : A^0(\widetilde{TM}) \longrightarrow A^0(\widetilde{TM} \otimes H^*)$ is defined by $\nabla^H_X Y = \nabla_{X^H} Y$, and $\nabla^V : A^0(\widetilde{TM}) \longrightarrow A^0(\widetilde{TM} \otimes V^*)$ is defined by $\nabla^V_X Y = \nabla_{X^V} Y$ for all $X, Y \in A^0(\widetilde{TM})$ respectively. The covariant derivative ∇G of the metric G is decomposed into the form $\nabla G = \nabla^H G + \nabla^V G$, and thus the assumption (1.16) is equivalent to $\nabla^H_X G = 0$:

$$X^H G(Y, Z) = G(\nabla^H_X Y, Z) + G(Y, \nabla^H_X Z) \tag{1.17}$$

for all $X, Y, Z \in A^0(\widetilde{TM})$. By the definition (1.4) of Cartan tensor field C, we have

$$(\nabla^V_X G)(Y, Z) = 2C(X, Y, Z) \tag{1.18}$$

for all $X, Y, Z \in A^0(\widetilde{TM})$.

On the other hand, (1.17) implies $\nabla^H_X \mathcal{E} = \theta(X^H) = 0$ and

$$X^H F^2 = X^H G(\mathcal{E}, \mathcal{E}) = G(\nabla^H_X \mathcal{E}, \mathcal{E}) + G(\mathcal{E}, \nabla^H_X \mathcal{E}) = 0.$$

Therefore we obtain

Proposition 1.9 *Let* $\theta \in A^1(\widetilde{TM})$ *be the Ehresmann connection of* $\pi : TM^\times \to M$ *defined by* (1.14) *for the Chern connection* ∇ *of* (M, F). *Then we have*

$$d^H F \equiv 0. \tag{1.19}$$

Since the condition (1.15) is equivalent to $\sum \omega^i_j \wedge dx^j = 0$, the connection form ω is given in the form $\omega^i_j = \sum \Gamma^i_{jk}(x, y)dx^k$ with the coefficients Γ^i_{jk} satisfying the symmetric property $\Gamma^i_{jk} = \Gamma^i_{kj}$. Then the condition (1.17) is written as $d^H G - {}^t\omega G - G\omega = 0$, and thus the coefficients Γ^i_{jk} are given by

$$\Gamma^i_{jk}(x, y) = \frac{1}{2} \sum G^{il} \left(\frac{\delta G_{lk}}{\delta x^j} + \frac{\delta G_{jl}}{\delta x^k} - \frac{\delta G_{jk}}{\delta x^l}\right), \tag{1.20}$$

where (G^{ij}) denotes the inverse of (G_{ij}).

1.4 Parallel translation

Let (M, F) be a Finsler manifold with the Chern connection ∇. For a non-vanishing vector field $v = \sum v^i(x)(\partial/\partial x^i)$ on M, we define its covariant derivative ∇v with respect to ∇. Let $\tilde{v} : M \to \widetilde{TM}$ be the natural lift of v defined by $\tilde{v}(x) = \mathcal{E}(v(x)) = (v^*\mathcal{E})(x)$. The covariant derivative ∇v with respect to ∇ is given by $\nabla v = \tilde{v}^*\nabla\mathcal{E} = \tilde{v}^*\theta$:

$$\nabla v = \tilde{v}^*\nabla\mathcal{E} = \sum \frac{\partial}{\partial x^i} \otimes (dv^i + \sum v^j \Gamma^i_{jk}(x, v)dx^k). \tag{1.21}$$

If v satisfies $\nabla v = 0$, then v is said to be *parallel* with respect to ∇.

Let $c = (x(t)) : I = [0, 1] \to M$ be a smooth curve, and $v(t)$ be a non-vanishing vector field along c. Then we define a lift $\tilde{c}_v : I \to TM^\times$ of c by $\tilde{c}_v = (x(t), v(t))$. A lift $\tilde{c}_v$ is said to be *horizontal* if it satisfies $\tilde{c}_v^*\theta = 0$:

$$\tilde{c}_v^*\nabla\mathcal{E} = \sum \frac{\partial}{\partial x^i} \otimes \left[\frac{dv^i}{dt} + \sum v^j(t)\Gamma^i_{jk}\left(\tilde{c}_v(t)\right)\frac{dx^k}{dt}\right] = 0. \tag{1.22}$$

If $v(t)$ satisfies this equation, $v(t)$ is said to be *parallel* along c.

The system (1.22) has a unique solution $v_\zeta(t)$ which depends on the initial condition $\zeta = v_\zeta(0)$ smoothly. From smooth dependence of solutions on ζ, the mapping $P_{c(t)} : T_{c(0)}M \to T_{c(t)}M$ defined by $P_{c(t)}(\zeta) = (c(t), v_\zeta(t))$ is a diffeomorphism for every $t \in I$. Because of homogeneity of θ, if $v_\zeta(t)$ is a solution of (1.22), then $\lambda v_\zeta(t)$ is also a solution satisfying $\lambda v_\zeta(0) = \lambda\zeta$, and thus the uniqueness of solutions implies that $v_{\lambda\zeta}(t) = \lambda v_\zeta(t)$. Hence the horizontal lift $\tilde{c}_v(t)$ of a curve c starting at $\zeta \in T_{c(0)}M$ satisfies the homogeneity $\tilde{c}_{\lambda v}(t) = (c(t), \lambda v_\zeta(t)) = \lambda\tilde{c}_v(t)$, and $P_{c(t)}$ also satisfies the homogeneity $P_{c(t)}(\lambda\zeta) = \lambda P_{c(t)}(\zeta)$ for all $\lambda > 0$ and $\zeta \in T_{c(0)}M$. The family $P_c = \{P_{c(t)} : t \in I\}$ is called the *parallel translation* along c with respect to ∇.

The tangent space T_xM at every point $x \in M$ becomes a normed linear space with a norm $\| \bullet \|_x = F(x, \bullet\,)$. If we put $P_{c(t)}(\zeta) = (c(t), v_\zeta(t))$ for any point ζ in $T_{c(0)}M$, the norm $\|v_\zeta(t)\|$ of the vector field $v_\zeta(t)$ along $c(t)$ is given by $F\left(c(t), v_\zeta(t)\right)$. Then, because of Proposition 1.9 we have

$$dF(c(t), v_\zeta(t)) = d(\tilde{c}_v^*F) = \tilde{c}_v^*(d^VF + d^HF) = \tilde{c}_v^*(d^HF) = 0.$$

Hence the parallel translation P_c is norm-preserving:$\|P_{c(t)}(\zeta)\|_{c(t)} = \|\zeta\|_{c(0)}$.

Proposition 1.10 *The parallel translation P_c along any curve $c = c(t)$ on M is a norm-preserving map between the tangential normed-spaces.*

The parallel translation P_c is said to be *isometry* if it satisfies

$$\left\|P_{c(t)}(\zeta) - P_{c(t)}(\eta)\right\|_q = \|\zeta - \eta\|_p \tag{1.23}$$

for all $\zeta, \eta \in T_pM$. The parallel translation P_c along a curve $c = c(t)$ is norm-preserving, but not isometry in general. It is trivial that, if P_c is a linear mapping, then P_c is an isometry. In a later section, we shall consider the case where every tangent spaces are isometric mutually as normed linear spaces.

We denote by C_p the set of all (piecewise) smooth curves c with starting point $p = c(0)$ and ending point $p = c(1)$. Then there exists a natural product "$\circ$" in C_p. We also set

$H_p = \{P_{c(1)} : T_pM \longrightarrow T_pM \mid c \in C_p\}$. Defining $P_{c_1(1)} \circ P_{c_2(1)} = P_{(c_1 \cdot c_2)(1)}$, we can easily show that H_p is a group with the multiplication "$\circ$". This group H_p is called the *holonomy group* with reference point $p \in M$. In general, since the parallel translation P_c is not linear between the fibres, H_p is not a Lie group.

1.5 Torsion

The canonical bundle morphism $\pi_* : T(TM^\times) \to \widetilde{TM}$ in the sequence (1.2) is identified as a section of $A^1(\widetilde{TM})$, and the Chern connection ∇ makes π_* parallel by the assumption (1.15). We shall calculate the covariant exterior derivative $\nabla \pi_*$:

$$\nabla \pi_* = \sum \frac{\partial}{\partial x^i} \otimes \Big(\sum \omega^i_j \wedge dx^j\Big).$$

Since V is integrable and π_* satisfies (1.12), we have $\nabla \pi_*(X^V, Y^V) = 0$. Furthermore, for all $X, Y \in A^0(\widetilde{TM})$, we have

$$\begin{aligned}(\nabla \pi_*)(X^H, Y^H) &= \nabla_{X^H} \pi_*(Y^H) - \nabla_{Y^H} \pi_*(X^H) - \pi_*[X^H, Y^H] \\ &= \nabla^H_X Y - \nabla^H_Y X - \pi_*[X^H, Y^H]\end{aligned}$$

and

$$\begin{aligned}(\nabla \pi_*)(X^V, Y^H) &= \nabla_{X^V} \pi_*(Y^H) - \nabla_{Y^H} \pi_*(X^V) - \pi_*[X^V, Y^H] \\ &= \nabla^V_X Y - \pi_*[X^V, Y^H].\end{aligned}$$

Therefore we have

Proposition 1.11 *The assumption* (1.15) *of the Chern connection* ∇ *is equivalent to*

$$\nabla^H_X Y - \nabla^H_Y X = \pi_*[X^H, Y^H] \tag{1.24}$$

and

$$\nabla^V_X Y = \pi_*[X^V, Y^H] \tag{1.25}$$

for all $X, Y \in A^0(\widetilde{TM})$.

On the other hand, the Chern connection ∇ defines another bundle morphism $\theta : T(TM^\times) \to \widetilde{TM}$ by (1.10). In Finsler geometry, the covariant exterior differential $\nabla\theta$ plays an important role.

Definition 1.12 The *torsion* T of ∇ is defined by

$$T = \nabla\theta. \tag{1.26}$$

We remark that the vertical part of T vanishes identically. In fact, we have

$$\begin{aligned}T(X^V, Y^V) &= \nabla_{X^V} \theta(Y^V) - \nabla_{Y^V} \theta(X^V) - \theta[X^V, Y^V] \\ &= \nabla^V_X Y - \nabla^V_Y X - \theta[X^V, Y^V] = 0\end{aligned}$$

for all $X, Y \in A^0(\widetilde{TM})$, since the vertical subbundle V is integrable. Similar computations lead us to

Proposition 1.13 *The horizontal part T^{HH} and the mixed part T^{HV} are given by*

$$T^{HH}(X,Y) = -\theta[X^H, Y^H] = d\theta(X^H, Y^H) \tag{1.27}$$

and

$$T^{HV}(X,Y) = \nabla^H_X Y - \theta[X^H, Y^V] \tag{1.28}$$

for all $X, Y \in A^0(\widetilde{TM})$ respectively. In particular, the mixed part T^{HV} satisfies

$$T^{HV}(X, \mathcal{E}) = 0, \tag{1.29}$$

where $T^{HH}(X,Y) = T(X^H, Y^H)$ and $T^{HV}(X,Y) = T(X^H, Y^V)$.

Remark 1.14 Using local coordinate $(x^1, \cdots, x^n)$, the torsion form T is given by

$$T = \sum \frac{\partial}{\partial x^i} \otimes \left(d\theta^i + \sum \omega^i_j \wedge \theta^j \right). \tag{1.30}$$

Case of $T^{HH} \equiv 0$

The horizontal subbundle H is integrable if and only if $[H, H] \subset H$. Then, from (1.13), H is integrable if and only if $\theta([X^H, Y^H]) = 0$ for all $X, Y \in A^0(\widetilde{TM})$.

Definition 1.15 The *integrability tensor* $\Theta \in A^2(\widetilde{TM})$ of θ is defined by

$$\Theta(X,Y) = d\theta(X^H, Y^H) = -\theta[X^H, Y^H]. \tag{1.31}$$

From (1.27), we have $T^{HH} = \Theta$, and thus $T^{HH} = 0$ means that the horizontal subbundle H is integrable. Therefore H defines a foliation $\mathcal{F}$ on the total space $TM^\times$ which is transversal to the fibres if $T^{HH} \equiv 0$.

Definition 1.16 A non-vanishing smooth section $v : M \to TM^\times$ is said to be *horizontal* with respect to ∇ if it satisfies $\nabla v = v^*\nabla\mathcal{E} = v^*\theta = 0$.

For a horizontal section v, we have $v^* d^V \equiv 0$, and thus $v^* \circ d^H = d \circ v^*$.

From this, the integrability condition $d(v^*\theta) = 0$ for v to be horizontal is given by $v^*\Theta = 0$. Hence, if we assume $\Theta \equiv 0$, this assumption guarantees the existence of a horizontal section $v(x) = (x, y(x))$ satisfying $v(x_0) = \zeta (\neq 0)$ for an arbitrary initial point $\zeta \in T_{x_0}M$. The n-dimensional submanifold of TM defined by $y = y(x)$ is the maximal integrable manifold, which is the leaf $\mathcal{F}_\zeta$ of the foliation $\mathcal{F}$ through the point (x_0, ζ).

For the metric G on $\widetilde{TM}$, we define a Riemannian metric g_v on M by $g_v = v^*G$ for a horizontal section v. For the connection form ω of ∇, the form $\omega_v := v^*\omega$ defines a connection ∇^M on TM. Then we have

Proposition 1.17 *Assume that $T^{HH} \equiv 0$. Let $v : M \to TM^\times$ be a horizontal section, and g_v the induced metric on M. The induced connection ∇^M on M is the Levi-Civita connection of (M, g), and v is parallel with respect to ∇^M.*

Case of $T^{HV} \equiv 0$: Landsberg spaces

The metric G on $\widetilde{TM}$ makes each fibre T_xM a Riemannian space with the metric G_x. Since the parallel translation P_c along any curve c is given by the horizontal lift $\tilde{c}_v$ and its tangent vector field lies in horizontal space:

$$\frac{d\tilde{c}_v(t)}{dt} = \sum \frac{\delta}{\delta x^i} \otimes \frac{dx^j}{dt} \in H_{\tilde{c}_v(t)},$$

P_c is an isometry between the fibres if and only if

$$\mathcal{L}_{X^H} G = 0 \tag{1.32}$$

for every $X \in A^0(\widetilde{TM})$ (cf. [23]). We have then

Proposition 1.18 *The Lie derivative $\mathcal{L}_{X^H} G$ is given by*

$$[\mathcal{L}_{X^H} G]\,(Y, Z) = G(T^{HV}(X, Y), Z) + G(Y, T^{HV}(X, Z)) \tag{1.33}$$

for all $X, Y, Z \in A^0(\widetilde{TM})$.

Definition 1.19 A Finsler manifold (M, F) is said to be a *Landsberg* if $T^{HV} \equiv 0$.

From the relation (1.33), we can prove that (1.32) is equivalent to $T^{HV} = 0$.

Theorem 1.20 ([22]) *A Finsler manifold (M, F) is Landsberg if and only if the parallel translation P_c along any curve c is an isometry between the tangential Riemannian spaces.*

From this theorem, we see that any parallel translation $P_c \in H_p$ along $c \in C_p$ is a isometry in the tangential Riemannian space T_pM if (M, F) is a Landsberg space. Then it is shown that H_p is a Lie group, since the isometry group G of any Riemannian manifold is a Lie group.

Proposition 1.21 ([26]) *Suppose that (M, F) is a Landsberg space. Then the holonomy group H_p with reference point $p \in M$ is a Lie group.*

Remark 1.22 Each fibre T_xM is a Riemannian manifold with the metric induced form G. Then, the condition $T^{HV} = 0$ means that each fibre T_xM is totally geodesic in TM with some Sasakian-type metric (cf. [1]).

On the other hand, the volume form on each fibre T_xM is induced from the n-form $\Pi = \sqrt{\det G}\, \theta^1 \wedge \cdots \wedge \theta^n$. We shall consider the case where each fibre T_xM is minimal in TM. From (1.25), we have

$$(\mathcal{L}_{X^H} \Pi)(Y_1, \cdots, Y_n) = \mathcal{L}_{X^H}(\Pi(Y_1, \cdots, Y_n)) + \sum_{k=1}^{n} \Pi\left(Y_1, \cdots, T^{HV}(X, Y_k), \cdots, Y_n\right)$$

for all $X,\ Y_1, \cdots,\ Y_n \in A^0(\widetilde{TM})$. If we put $(\operatorname{tr} T^{HV})(X) := \operatorname{trace}\{Y \to T^{HV}(X, Y)\}$, then, by direct computations, we can show the following

$$\mathcal{L}_{X^H} \Pi = (\operatorname{tr} T^{HV})(X) \Pi.$$

Thus $\mathcal{L}_{X^H} \Pi = 0$ if and only if $(\operatorname{tr} T^{HV})(X) = 0$.

For any $X \in A^0(\widetilde{TM})$, we denote by φ_t the one parameter group of local transformation in TM induced from X^H. For a compact subset $K_0 \subset T_{x_0}M,\ x_0 \in M$, we set

$K_t = \varphi_t(K_0) \subset T_{\varphi_t(x_0)}M$. Then the volume $V(K_t)$ is defined by $V(K_t) = \int_{K_t} \Pi$. Suppose that (M, F) is Landsberg space. Then by Theorem 1.20, each φ_t is isometry between the fibres, and we have

$$\frac{d}{dt}V(K_t) = \frac{d}{dt}\int_{K_t} \Pi = \frac{d}{dt}\int_{K_0} \varphi_t^* \Pi = \int_{K_0} \frac{d}{dt}(\varphi_t^* \Pi) = \int_{K_0} \mathcal{L}_{X^H} \Pi = 0.$$

Hence, if (M, F) is Landsberg, the volume $V(K_t)$ is constant(cf. [7]).

A Finsler manifold (M, F) satisfying $\operatorname{tr} T^{HV} = 0$ is called a *weak Landsberg space* (cf. [44]). The condition $\operatorname{tr} T^{HV} = 0$ means that each fibre T_xM is is minimal submanifold in TM with some Sasakian-type metric (cf. [1]).

1.6 Curvature

An important quantity in geometry is the *curvature* which measures the flatness of the space.

Definition 1.23 The *curvature* R of ∇ is defined by

$$R = \nabla^2. \tag{1.34}$$

Similarly to the case of torsion T, we first remark that the vertical part of R vanishes identically. In fact, we obtain

$$\begin{aligned} R(X^V, Y^V)Z &= \nabla^V_X \nabla^V_Y Z - \nabla^V_Y \nabla^V_X Z - \nabla_{[X^V, Y^V]} Z \\ &= \sum \frac{\partial}{\partial x^i} \otimes \left[X^V Y^V(Z^i) - Y^V X^V(Z^i) - [X^V, Y^V](Z^i) \right] = 0 \end{aligned}$$

for all sections $X, Y, Z \in A^0(\widetilde{TM})$, since the vertical subbundle V satisfies $[V, V] \subset V$. Hence the surviving part of R are the horizontal part R^{HH} and the mixed part R^{HV}:

$$\begin{aligned} R^{HH}(X, Y) = R(X^H, Y^H)Z &= \nabla_{X^H} \nabla_{Y^H} Z - \nabla_{Y^H} \nabla_{X^H} Z - \nabla_{[X^H, Y^H]} Z \\ &= \nabla^H_X \nabla^H_Y Z - \nabla^H_Y \nabla^H_X Z - \nabla_{[X^H, Y^H]} Z \end{aligned}$$

and

$$\begin{aligned} R^{HV}(X, Y) = R(X^H, Y^V)Z &= \nabla_{X^H} \nabla_{Y^V} Z - \nabla_{Y^V} \nabla_{X^H} Z - \nabla_{[X^H, Y^V]} Z \\ &= \nabla^H_X \nabla^V_Y Z - \nabla^V_Y \nabla^H_X Z - \nabla_{[X^H, Y^V]} Z. \end{aligned}$$

The following is trivial from the definition.

$$R^{HH}(X, Y)Z + R^{HH}(Y, X)Z = 0. \tag{1.35}$$

From the assumption (1.15), the Ricci identity $\nabla^2 \pi_* = R \wedge \pi_*$ gives $R \wedge \pi_* = 0$. Then

$$\begin{aligned} (R \wedge \pi_*)(X^H, Y^H, Z^H) &= R(X^H, Y^H)\pi_*(Z^H) + R(Y^H, Z^H)\pi_*(X^H) \\ &\quad + R(Z^H, X^H)\pi_*(Y^H) = R^{HH}(X, Y)Z + R^{HH}(Y, Z)X + R^{HH}(Z, X)Y \end{aligned}$$

and

$$(R \wedge \pi_*)(X^H, Y^V, Z^H) = R(X^H, Y^V)\pi_*(Z^H) + R(Y^V, Z^H)\pi_*(X^H) + R(Z^H, X^H)\pi_*(Y^V) = R^{HV}(X,Y)Z - R^{HV}(Z,Y)X$$

induce the following.

Proposition 1.24 *For all $X, Y, Z \in A^0(\widetilde{TM})$, the curvatures R^{HH} and R^{HV} satisfy the following identities:*

$$R^{HH}(X,Y)Z + R^{HH}(Y,Z)X + R^{HH}(Z,X)Y \equiv 0 \tag{1.36}$$

and

$$R^{HV}(X,Y)Z - R^{HV}(Z,Y)X \equiv 0. \tag{1.37}$$

By the definition (1.26), the Ricci identity $\nabla^2\mathcal{E} = R\mathcal{E}$ implies the relation $T = R\mathcal{E}$.

Proposition 1.25 *The curvature R and the torsion T of the Chern connection ∇ satisfy the relations*

$$R^{HH}(X,Y)\mathcal{E} = T^{HH}(X,Y) \tag{1.38}$$

and

$$R^{HV}(X,Y)\mathcal{E} = T^{HV}(X,Y) \tag{1.39}$$

for all $X, Y \in A^0(\widetilde{TM})$.

The symmetry assumption (1.15) derives Proposition 1.24. The almost G-compatibility assumption (1.16) derives the followings. Firstly, concerning with R^{HH}, we have

Proposition 1.26 *The horizontal curvature R^{HH} satisfies*

$$G(R^{HH}(X,Y)Z,W) + G(R^{HH}(X,Y)W,Z) + 2C(T^{HH}(X,Y),Z,W) = 0 \tag{1.40}$$

for all X, Y, Z and $W \in A^0(\widetilde{TM})$.

Secondary, concerning the mixed part R^{HV}, we have

Proposition 1.27 *The mixed part R^{HV} satisfies the following identities.*

$$G(R^{HV}(X,Y)Z,W) + G(Z, R^{HV}(X,Y)W) + 2(\nabla^H_X C)(Y,Z,W) + 2C(T^{HV}(X,Y),Z,W) = 0 \tag{1.41}$$

for all sections X, Y, Z and W of $\widetilde{TM}$.

As an application of Proposition 1.27, we obtain the following

Proposition 1.28 *The mixed part R^{HV} of the curvature R satisfies the following identity*

$$R^{HV}(X,\mathcal{E}) = 0 \tag{1.42}$$

for every $X \in A^0(\widetilde{TM})$.

Remark 1.29 The *curvature form* Ω^i_j of ∇ is the 2-form on $TM^\times$ defined by setting $R\dfrac{\partial}{\partial x^j} = \sum \dfrac{\partial}{\partial x^i} \otimes \Omega^i_j$. By definition of R, we obtain the 1-st Bianchi identity $\Omega = d\omega + \omega \wedge \omega$:

$$\Omega^i_j = d\omega^i_j + \sum \omega^i_l \wedge \omega^l_j. \tag{1.43}$$

Differentiating this identity, we obtain the 2-nd Bianchi identity $\nabla\Omega = 0$:

$$d\Omega^i_j + \sum \omega^i_l \wedge \Omega^l_j - \sum \Omega^i_l \wedge \omega^l_j = 0. \tag{1.44}$$

In the case of Finsler geometry, this identity induces some complicated identities, since ∇ has non-zero torsions T^{HH} and T^{HV}. For example, if we calculate the horizontal 3-form of the left-hand-side of (1.44), we obtain

$$\begin{aligned}(\nabla^H_X R^{HH})(Y,Z) &+ (\nabla^H_Y R^{HH})(Z,X) + (\nabla^H_Z R^{HH})(X,Y)\\ &= R^{HV}(Y, T^{HH}(X,Z)) + R^{HV}(Z, T^{HH}(Y,X)) + R^{HV}(X, T^{HH}(Z,Y)).\end{aligned}$$

The other identities including the terms $\nabla^H R^{HV}$, $\nabla^V R^{HH}$ and $\nabla^V R^{HV}$, see the book [10] or [32].

Case of $R^{HH} \equiv 0$

We suppose $R^{HH} \equiv 0$. Then (1.38) gives $\Theta = T^{HH} \equiv 0$, and thus $TM^\times$ admits a horizontal section $v : M \to TM^\times$.

Proposition 1.30 *If $R^{HH} \equiv 0$, the induced metric $g_v = v^*G$ is a flat Riemannian metric on TM, and so M is locally Euclidean.*

Case of $R^{HV} \equiv 0$: Berwald spaces

In this subsection, we shall consider the case of $R^{HV} \equiv 0$.

Definition 1.31 A Finsler manifold (M, F) is said to be *Berwald* if $R^{HV} \equiv 0$.

Because of (1.39), a Berwald space is a special class of Landsberg spaces. If (M, F) is Berwald, the Chern connection ∇ is linear, that is, there exists a symmetric linear connection ∇^M on TM such that $\nabla = \pi^*\nabla^M$.

If (M, F) is a Berwald space, then Szabó's theorem [49] showed that we can find a Riemannian metric g on M which is compatible with ∇^M. We show the outline of the proof of this fact. For this, we define a isometric group $\mathcal{G}$ of F. For an arbitrary point $x \in M$, we set $\mathcal{G} = \{g \in GL(n, \mathbb{R}) \mid \|gy\| = \|y\|,\ \forall y \in T_xM\}$. By the continuity of the norm $\|\cdot\|$ and the homogeneity of F, we can prove that $\mathcal{G}$ is a compact Lie group [21]. Then we define an inner product $\langle\cdot,\cdot\rangle_x$ on T_xM by

$$\langle y_1, y_2\rangle_x = \int_{\mathcal{G}} (gy_1, gy_2)dg$$

for an arbitrary inner product $(\cdot,\cdot)$ on T_xM and a bi-invariant Haar measure dg on $\mathcal{G}$. By definition, it is trivial that this inner product $\langle\cdot,\cdot\rangle_x$ is $\mathcal{G}$-invariant.

On the other hand, the holonomy group H_x of ∇^M with the reference point x is a subgroup of $\mathcal{G}$, since ∇^M preserves F invariant. Hence $\langle\cdot,\cdot\rangle_x$ is also H_x-invariant. Thus we can extend the inner product $\langle\cdot,\cdot\rangle_x$ to a Riemannian metric g on TM by the help of the parallel displacement with respect to ∇^M. It is trivial that ∇^M is compatible with respect to this metric g. Hence we have

Theorem 1.32 ([49]) *Suppose that a Finsler manifold (M, F) is a Berwald space. Then there exists a Riemannian metric g on M such that the Chern connection ∇ of (M, F) is given by $\nabla = \pi^*\nabla^M$ for the Levi-Civita connection ∇^M of (M, g).*

The curvature is related with parallel translation. The following theorem due to [21] characterizes Berwald spaces in terms of parallel translations.

Theorem 1.33 ([21]) *A Finsler manifold (M, F) is Berwald if and only if the parallel translation P_c along any curve c is an isometry between the tangential normed spaces.*

Proof We suppose that the parallel translation P_c along any curve $c = c(t)$ in M is an isometry between the tangential normed spaces. Then, by the Mazur-Ulam's theorem (cf. [33] and [50]), the mapping P_c is linear and there exists a $GL(n, \mathbb{R})$-valued function $A_c = (A^i_j(t))$ satisfying $P_c(\zeta) = (c(t), A_c(t)\zeta) = (x^i(t), \sum A^i_j(t)\zeta^j)$. From (1.22) we get

$$\sum \frac{dA^i_j(t)}{dt}\zeta^j = -\sum N^i_j\left(c(t), A_c(t)\zeta\right)\frac{dx^j}{dt}$$

for all curves $c(t)$ and initial point ζ. Therefore the coefficients $N^i_j(x, y) = \sum \varGamma^i_{lj}y^l$ are linear with respect to the fibre coordinate $(y^1, \cdots, y^n)$, namely, $\varGamma^i_{lj} = \varGamma^i_{lj}(x)$, which shows $R^{HV} = 0$.

Conversely we suppose that (M, F) is Berwald. Then, since the functions $N^i_j(x, y)$ are linear in the fibre coordinate $(y^1, \cdots, y^n)$, the solutions $y_\zeta(t)$ of (1.22) are linear in ζ. Hence the parallel translation P_c is linear, and from (1.23), we have

$$\|P_c(\zeta) - P_c(\eta)\|_{c(t)} = \|P_c(\zeta - \eta)\|_{c(t)} = \|\zeta - \eta\|_{c(0)}$$

for all $\zeta, \eta \in T_{c(0)}M$. Hence P_c is an isometry. □

Example 1.34 Let $F = \alpha + \beta$ be a Randers metric on M. Then, by the well-known theorem due to [25], (M, F) is Berwald if and only if β is parallel 1-form on the base Riemannian manifold (M, g).

Remark 1.35 From (1.39), if R^{HV} vanishes identically, then T^{HV} also vanishes. Hence the class of Landsberg spaces contains the class of Berwald spaces. There exists a lot of example of Berwald spaces. However it is still an open problem to find an example of non-Berwald Landsberg space.

Case of $R^{HH} = R^{HV} \equiv 0$: Locally Minkowski spaces

In this section, we shall be concerned with flat Finsler manifolds.

Definition 1.36 A Finsler manifold (M, F) is said to be *locally Minkowski* if there exists a local coordinate system on M with respect to which the function F is independent of the base point $x \in M$.

We have the following well-known theorem (cf. [32]).

Theorem 1.37 *A Finsler manifold* (M, F) *is locally Minkowski if and only if the Chern connection* ∇ *is flat.*

Proof We suppose that ∇ is flat, that is, $R^{HH} = R^{HV} \equiv 0$. Then, the Chern connection ∇ is induced from a flat connection ∇^M on TM. Hence there exists an open cover $\mathcal{U}$ of M and local frame fields $(e_1, \cdots, e_n)$ on $U \in \mathcal{U}$ such that $\nabla e_j = 0$. This condition is equivalent to the existence of the change of frames $e_j = \sum \frac{\partial}{\partial x^i} A^i_j(x)$, $A = (A^i_j(x)) : U \to GL(n, \mathbb{R})$, on each U satisfying $dA + \omega A = 0$. With respect to such a local frame field e, the connection form $\tilde{\omega} = A^{-1}dA + A^{-1}\omega A$ of ∇ vanishes on U. Then, $\tilde{\omega} = 0$ and (1.19) imply the independence of F on the base point $x \in U$. If we denote by $B = (B^i_j(x))$ the inverse of the matrix A, the condition above is equivalent to

$$\frac{\partial B^i_j(x)}{\partial x^k} = \sum B^i_m \Gamma^m_{jk}.$$

Since ∇ is symmetric, there exist some functions $w^i(x)$ such that $B^i_j = \partial w^i / \partial x^j$, and the local frame e_j is given by

$$e_j = \sum \frac{\partial}{\partial x^i} \frac{\partial x^i}{\partial w^j} = \frac{\partial}{\partial w^j}.$$

Hence there exists a local coordinate system $\{U, (w^i)\}$ on M with respect to which the function F is independent on the base point $x \in U$.

Conversely we assume that F is independent of the base point x. By definition, the metric tensor G_{ij} is also independent of the base point x. Then, we get $N^i_j = \sum \Gamma^i_{lj} y^l = 0$ and (1.20) imply $\Gamma^i_{jk} = 0$, and thus the Chern connection ∇ is flat. □

Example 1.38 Let $F = \alpha + \beta$ be a Randers metric on M. Then (M, F) is locally Minkowski if and only if (M, F) is Berwald and the base Riemannian manifold (M, g) is flat ([25]).

This example is true for any locally Minkowski space.

Proposition 1.39 *A Finsler manifold* (M, F) *is locally Minkowski if and only if* (M, F) *is Berwald and its associated Riemannian metric is flat.*

1.7 Flag curvature

Let X be a tangent vector at $x \in M$. We may consider X a section of $\widetilde{TM}$. Then the 2-plane $\mathcal{F}(X)$ spanned by X and $\mathcal{E}$ is called the *flag* with the *flagpole* $\mathcal{E}$. For the curvature tensor R of ∇, the sectional curvature

$$\frac{G(R(X, \mathcal{E})\mathcal{E}, X)}{\|X\|^2 \|\mathcal{E}\|^2 - G(X, \mathcal{E})^2}$$

is called the *flag curvature* of the flag $\mathcal{F}(X)$, and denoted by K. From (1.42), we have $G(R^{HV}(X, \mathcal{E})\mathcal{E}, X) = 0$ for every section $X \in \Gamma(\widetilde{TM})$, and so the flag curvature K is given by

$$K(X) = \frac{G(R^{HH}(X, \mathcal{E})\mathcal{E}, X)}{\|X\|^2 \|\mathcal{E}\|^2 - G(X, \mathcal{E})^2}. \tag{1.45}$$

The flag curvature K depends on X and the point $(x, y) \in TM^\times : K = K(x, y, X)$. If the flag curvature K is independent of X at every point $(x, y) \in TM^\times$, the space is said to be *of scalar flag curvature* $K(x, y)$. A Finsler manifold (M, F) is said to be of *constant flag curvature* if K is constant. For Finsler manifolds of constant flag curvature, see Chapter 12 in [10]. The proof of the following theorem is found in [32].

Theorem 1.40 (Schur's lemma) *Let (M, F) be a Finsler manifold of scalar flag curvature $K = K(x, y)$. If K is a function of position $x \in M$ alone, then (M, F) is of constant flag curvature provided* $\dim M \geq 3$.

Example 1.41 Let F be the Funk metric on the unit ball $\mathbb{B}$ defined in Example 1.1. It is well-known that F has negative constant flag curvature $K = -1$ (see [37] or [14]).

2 Geodesics in Finsler manifolds

Let $\gamma : I = [0, 1] \to M$ be a smooth curve. Since the symmetry condition $F(x, y) = F(x, -y)$ is not assumed, the orientation of curves is essential, that is, if a curve γ is given, then we always assume that γ is *oriented* by the parameter t. A smooth curve $\gamma = \gamma(t)$ is said to be *regular* if $\dot{\gamma}(t) := d\gamma/dt \neq 0$ for every $t \in I$.

Let $\mathit{\Gamma}(p, q)$ be the set of all regular oriented curves with the initial point $p = \gamma(0)$ and the terminal point $q = \gamma(1)$. Then we define a functional $L_F : \mathit{\Gamma}(p, q) \to \mathbb{R}$ by

$$L_F(\gamma) = \int_0^1 \|\dot{\gamma}(t)\| \, dt = \int_0^1 F\left(\tilde{\gamma}(t)\right) dt.$$

Since F satisfies the homogeneity condition, this definition is well-defined. For an ordered pair $(p, q) \in M \times M$, the distance function $d_F(p, q)$ is defined by $d_F(p, q) = \inf_c L_F(\gamma)$, where infinimum is taken over of all oriented (piecewise) smooth curves from p to q. In general, since the symmetry condition is not assumed, the distance function d_F does not satisfy the symmetric property $d_F(p, q) = d_F(q, p)$. However, the distance d_F satisfies the following conditions:

(1) $d_F(p, q) \geq 0$,

(2) $d_F(p, q) = 0$ if and only if $p = q$,

(3) $d_F(p, q) \leq d_F(p, r) + d_F(r, q)$.

The metric topology of M is defined by the sets $B(p, \delta) = \{q \in M \mid d_F(p, q) < \delta\}$, and the metric topology of a connected Finsler manifold (M, F) coincides with the manifold topology of M.

2.1 Geodesics in Finsler manifolds

The *canonical lift* of a regular oriented curve γ is the curve $\tilde{\gamma} : I \to TM^\times$ defined by $\tilde{\gamma}(t) = (\gamma(t), \dot{\gamma}(t))$. For a vector field $X(t)$ along γ, we consider $X(t)$ as a section of $\widetilde{TM}$ along $\tilde{\gamma}$. Then we use the notation $\nabla_t X$ instead of $\tilde{\gamma}^* \nabla X$:

$$\nabla_t X = \sum \frac{\partial}{\partial x^i} \otimes \left[\frac{dX^i}{dt} + \sum X^j(t) \mathit{\Gamma}^i_{jk}\left(\tilde{\gamma}(t)\right) \frac{dx^k}{dt} \right]. \tag{2.1}$$

In particular, if $X(t) = \dot{\gamma}(t)$, then we have $\nabla_t \dot{\gamma}(t) = \tilde{\gamma}^* \theta$. We note that this equation is written as $\nabla_t X = \tilde{\gamma}^* \nabla^H X$, since X is a vector field along the curve $\gamma(t)$ in the base manifold M. The definition (2.1) and the metrical condition (1.14) imply

$$\frac{d}{dt} G(X, Y) = G(\nabla_t X, Y) + G(X, \nabla_t Y) \tag{2.2}$$

for all vector fields $X(t)$ and $Y(t)$ along γ. Hence we have

Proposition 2.1 *If* $\nabla_t X = \nabla_t Y = 0$*, then the inner product* $G(X, Y)$ *is constant along* γ.

Let (M, F) be a Finsler manifold with the Chern connection ∇.

Definition 2.2 A regular oriented curve $\gamma : I \to M$ is said to be a *path* if its canonical lift $\tilde{\gamma}$ is horizontal, that is, $\nabla_t \dot{\gamma} = \tilde{\gamma}^* \theta = 0$.

The length s of a regular curve γ is defined by $s(t) = \displaystyle\int_0^t \|\dot{\gamma}(t)\| \, dt$, and the function $s(t)$ is an increasing function of the parameter t. If the parameter t is positively proportional to s, then t is said to be *normal*.

Definition 2.3 Let (M, F) be a Finsler manifold. A path in M with a normal parameter is called a is a *geodesic* in (M, F).

From (2.1), a regular oriented curve $\gamma(t) = (x^i(t))$ with normal parameter t is a geodesic if and only if $\nabla_t \dot{\gamma}(t) = 0$:

$$\frac{d^2 x^i}{dt^2} + \sum \Gamma^i_{jk}(\tilde{\gamma}(t)) \frac{dx^j}{dt} \frac{dx^k}{dt} = 0 \tag{2.3}$$

is satisfied. From the metrical condition (2.2), if $\gamma = \gamma(t)$ with normal parameter t is a geodesic, the tangent vector $\dot{\gamma}(t)$ has a constant norm and γ has constant speed. In the sequel, we always assume that the parameter of a geodesic to be normal otherwise stated.

Let $\gamma_X : I \to M$ be a geodesic with initial point $x = \gamma_X(0)$ and the initial direction $X = \dot{\gamma}_X(0)$, where the parameter t is, of course, is normal. We shall define the *exponential map* $\exp$ by $\exp(x, X) = \gamma_X(1)$ if $X \neq 0$ and $\exp(x, 0) = x$. The restriction of $\exp$ to $\mathcal{D} \cap T_x M$ is denoted by $\exp_x$. The restricted exponential map $\exp_p$ maps the rays through the origin $0 \in T_x M$ to the unique geodesics through the point x in sufficiently small $B_x(r) = \{X \in T_x M \mid \|X\| < r\}$.

The exponential map $\exp$ is defined on an open neighborhood $\mathcal{D}$ of the zero section $z(M)$ of TM, and $\exp$ is C^∞-class away from $z(M)$. Furthermore $\exp$ is C^1-class at $z(M)$, and its derivative at $z(M)$ is the identity map. By a result due to Akbar-Zadeh, the map $\exp$ is C^2-class at $z(M)$ if and only if (M, F) is a Berwald space (see [10]).

For each $X \in T_x M$, the *radial geodesic* γ_X is given by $\gamma_X(t) = \exp_x(tX)$ for all $t \in I$ such that either side is defined. This geodesic segment γ_X has tangent vector field $\dot{\gamma}_X$ with $\dot{\gamma}_X(0) = X$. Since $\nabla_t \dot{\gamma}_X = 0$, (2.2) implies that $\|\dot{\gamma}_X\|^2 = G(\dot{\gamma}_X, \dot{\gamma}_X)$ is constant along γ_X, and thus $\|\dot{\gamma}_X(t)\| = \|\dot{\gamma}_X(0)\| = \|X\|$. Consequently we have $\displaystyle\int_0^1 \|\dot{\gamma}_X(t)\| \, dt = \|X\|$.

2.2 The first variation of arc length and geodesics

In this section, we shall show the first variation formula in Finsler manifolds. For this end, we introduce some definitions.

Let $\gamma = \gamma(t) \in \Gamma(p,q)$ be a regular oriented curve with unit speed, that is, $\|\dot{\gamma}(t)\| = 1$. Then a *variation* of γ is a family $\{\gamma_s\}$ of oriented curves $\gamma_s(t)$ parameterized by $s \in (-\varepsilon, \varepsilon)$ such that $\gamma_0(t) = \gamma(t)$ for all $t \in I$. A variation Γ_γ is said to be *proper* if it fixes the end points, that is, $\gamma_s(0) = p$ and $\gamma_s(1) = q$. We suppose that the map $\Gamma_\gamma : (-\varepsilon, \varepsilon) \times I \to M$ defined by $\Gamma_\gamma(s,t) = \gamma_s(t)$ is smooth. (For the variational problem of arc length, it is enough to assume that Γ_γ is piecewise differentiable with respect to the parameter t (cf. [32], Chapter VII), however, we shall assume the smoothness of Γ_γ for the simplicity of discussions.)

By the assumption, the map Γ satisfies $\Gamma_\gamma(0,t) = \gamma(t)$, $p = \Gamma_\gamma(s,0)$ and $q = \Gamma_\gamma(s,1)$. Setting $s =$ constant for each $s \in (-\varepsilon, \varepsilon)$, the parameterized curve $\gamma_s : I \to M$ defined by $\gamma_s(t) = \Gamma_\gamma(s,t)$ is called a *s-curve*, while the parameterized curve $\gamma_t(s) = \Gamma_\gamma(s,t)$ is a *t-curve* which is transversal curve to γ. In local coordinates, we set $\Gamma_\gamma(s,t) = \big(x^1(s,t), \cdots, x^n(s,t)\big)$. We denote by $\mathcal{S} = \partial\gamma_t/\partial s$ and $\mathcal{T} = \partial\gamma_s/\partial t$ the tangent vector fields of t-and s-curves respectively:

$$\mathcal{S} = \sum \frac{\partial}{\partial x^i} \otimes \frac{\partial x^i}{\partial s}, \quad \mathcal{T} = \sum \frac{\partial}{\partial x^i} \otimes \frac{\partial x^i}{\partial t}.$$

In particular, the vector field $\mathcal{V}(t)$ along γ defined by

$$\mathcal{V}(t) = \left(\frac{\partial \gamma_t}{\partial s}\right)_{(0,t)} = \mathcal{S}(0,t)$$

is called the *variational field* induced from Γ_γ. If Γ_γ is proper, that is, Γ_γ satisfies $\gamma_s(0) = \gamma(0) = p$ and $\gamma_s(1) = \gamma(1) = q$ for all $s \in (-\varepsilon, \varepsilon)$, then the variational field $\mathcal{V}$ is *proper*, that is, $\mathcal{V}$ satisfies $\mathcal{V}(0) = \mathcal{V}(1) = 0$.

We are always concerned with the variation Γ_γ whose variational field $\mathcal{V}$ is independent of the tangent vector $\dot{\gamma}$ at least one point on γ. Let $\mathcal{V} = \mathcal{V}(t)$ be any vector filed along a regular oriented curve $\gamma = \gamma(t)$. Then there exists a variation Γ_γ which induces $\mathcal{V}$ as its variational field. In fact, if we take $\Gamma_\gamma(s,t) = \exp(s\mathcal{V}(t))$, then $\Gamma_\gamma : (-\varepsilon, \varepsilon) \times I \to M$ is a variation of γ with variational field $\mathcal{V}$.

Lemma 2.4 *Let $\mathcal{V}$ be any vector field along γ. Then $\mathcal{V}$ is a variational field of some variation Γ_γ of γ. If $\mathcal{V}$ is proper, then $\mathcal{V}$ is the variational field induced from a some proper variation Γ_γ.*

The vector fields $\mathcal{S}$ and $\mathcal{T}$ are naturally identified with sections of $\widetilde{TM}$ along the canonical lift $\tilde{\gamma}_s$ of s-curve γ_s:

$$\mathcal{S}\left(\tilde{\gamma}_s(t)\right) = \sum \frac{\partial}{\partial x^i} \otimes \frac{\partial x^i}{\partial s}, \quad \mathcal{T}\left(\tilde{\gamma}_s(t)\right) = \sum \frac{\partial}{\partial x^i} \otimes \frac{\partial x^i}{\partial t}.$$

Lemma 2.5 *Let $\Gamma_\gamma : (-\varepsilon, \varepsilon) \times I \to M$ be a variation. Then we have*

$$\nabla^H_{\mathcal{S}} \mathcal{T} = \nabla^H_{\mathcal{T}} \mathcal{S} \tag{2.4}$$

along $\tilde{\gamma}_s = (\gamma_s(t), \dot{\gamma}_s(t))$.

Let Γ_γ be a proper variation of a regular oriented curve $\gamma \in \Gamma(p, q)$. We compute the first variation of length functional $L_F(\gamma_s)$. Since $G(\mathcal{T}, \mathcal{T}) = F(\gamma_s(t), \dot{\gamma}_s(t))^2 = L_F(\gamma_s)^2$, we have

$$\frac{d}{ds} L_F(\gamma_s) = \frac{1}{2} \int_0^1 \frac{1}{\|\mathcal{T}\|} \frac{\partial G(\mathcal{T}, \mathcal{T})}{\partial s} dt.$$

Furthermore, (1.17) and (2.4) imply

$$\frac{1}{\|\mathcal{T}\|} \frac{\partial G(\mathcal{T}, \mathcal{T})}{\partial s} = \frac{2}{\|\mathcal{T}\|} G\left(\nabla^H_{\mathcal{S}} \mathcal{T}, \mathcal{T}\right) = \frac{2}{\|\mathcal{T}\|} G\left(\nabla^H_{\mathcal{T}} \mathcal{S}, \mathcal{T}\right) \tag{2.5}$$

along $\tilde{\gamma}_s$. Consequently we have

$$\frac{1}{\|\mathcal{T}\|} \frac{\partial G(\mathcal{T}, \mathcal{T})}{\partial s} = \frac{2}{\|\mathcal{T}\|} \left[\frac{d}{dt} G(\mathcal{S}, \mathcal{T}) - G\left(\mathcal{S}, \nabla^H_{\mathcal{T}} \mathcal{T}\right) \right],$$

which gives

$$\frac{d}{ds} L_F(\gamma_s) = \int_0^1 \frac{1}{\|\mathcal{T}\|} \left[\frac{d}{dt} G(\mathcal{S}, \mathcal{T}) - G\left(\mathcal{S}, \nabla^H_{\mathcal{T}} \mathcal{T}\right) \right] dt.$$

Evaluating $s = 0$, then $\|\mathcal{T}\|_{s=0} = \|\dot{\gamma}(t)\| = 1$ derives the following:

Proposition 2.6 (First Variation Formula) *Let $\gamma : I \to M$ be a regular oriented curve, and Γ_γ a proper variation of γ. Then*

$$\left.\frac{d}{ds}\right|_{s=0} L_F(\gamma_s) = -\int_0^1 G(\mathcal{V}, \nabla_t \dot{\gamma})\, dt \tag{2.6}$$

where $\mathcal{V}$ is the variational field of Γ_γ.

A regular oriented curve γ is said to be a *stationary point* of the functional L_F if $(dL_F(\gamma_s)/ds)_{s=0} = 0$ for any proper variation Γ_γ. If a regular oriented curve $\gamma : I \to M$ is a geodesic, then γ satisfies (2.3), and thus γ is a stationary point of L_F from (2.6).

Conversely we suppose that γ is a stationary point of the functional L_F. Since the condition $(dL_F(\gamma_s)/ds)_{s=0} = 0$ is satisfied for any variational field $\mathcal{V}$ along γ, we take $\mathcal{V}(t) = \varphi(t)\nabla_t\dot{\gamma}$ for a smooth function φ satisfying $\varphi(0) = \varphi(1) = 0$ and $\varphi > 0$ elsewhere. Then, since $\mathcal{V}$ is proper and from (2.6), we have

$$-\int_0^1 \varphi(t) \left\|\nabla_t \dot{\gamma}\right\|^2 dt = 0,$$

which implies $\nabla_t \dot{\gamma} = 0$ on I.

Proposition 2.7 *A regular oriented curve in a Finsler manifold (M, F) is a stationary point of the functional L_F if and only if γ is a geodesic from p to q.*

A curve γ from $p = \gamma(0)$ to $q = \gamma(1)$ is said to be *minimizing* if $d_F(p, q) = L_F(\gamma)$. Since minimizing curve is a stational curve, we have

Theorem 2.8 *Every minimizing curve in (M, F) is a geodesic if γ is regular.*

The converse of this theorem is also true.

Theorem 2.9 *Every geodesic in a Finsler manifold (M, F) is locally minimizing.*

This theorem is proved by using the Gauss lemma. We define the *geodesic ball* $\mathcal{B}_x(r)$ centered at $x \in M$ of radius r by $\mathcal{B}_x(r) = \exp(B_x(r))$ for the tangential ball $B_x(r) = \{\zeta \in T_xM \mid \|\zeta\| < r\}$. Let $S_x(r) = \{X \in T_xM \mid \|X\| = r\}$ be the tangent sphere. Then the set $\mathcal{S}_x(r) = \exp(S_x(r))$ is called the *geodesic sphere* at x of radius r. Then the Gauss lemma is stated as follows.

Lemma 2.10 (The Gauss Lemma) *The radial geodesic γ_X is orthogonal to the geodesic sphere $\mathcal{S}_x(r)$ at $x \in M$.*

For the proof of Theorem 2.9, we need more technical preliminaries, and thus we omit it here. For the complete proof, see [10] or [13].

2.3 Euler-Lagrange equation

From Proposition 2.8, a geodesic in (M, F) is characterized as the critical points of length functional L_F. In general, the equation which characterizes the critical points of a functional on $\Gamma(p, q)$ is called the *Euler-Lagrange equation* of L_F. For details, refer to the book [4] or [32].

We consider an arbitrary proper variation $\Gamma_\gamma : \gamma_s(t) = \gamma(t) + sX(t)$ of a smooth curve $\gamma(t)$ with fixed end points, that is, $X(0) = X(1) = 0$. Then, by definition of norms, we have $\|\dot\gamma_s(t)\| = F(\gamma_s(t), \dot\gamma_s(t))$. The Taylor extension gives

$$\|\dot\gamma_s(t)\| - \|\dot\gamma(t)\| = s\left(\sum \frac{\partial F}{\partial x^i} X^i + \sum \frac{\partial F}{\partial y^i} \dot X^i\right) + \frac{s^2}{2!}(\cdots) + \cdots,$$

and this extension implies

$$\left.\frac{d}{ds}\right|_{s=0} L_F(\gamma_s) = \int_0^1 \left(\sum \frac{\partial F}{\partial x^i} X^i + \sum \frac{\partial F}{\partial y^i} \dot X^i\right) dt.$$

Because of

$$\frac{d}{dt}\left(\sum \frac{\partial F}{\partial y^i} X^i\right) = \sum \frac{\partial F}{\partial y^i} \dot X^i + \frac{d}{dt}\frac{\partial F}{\partial y^i} X^i,$$

we have

$$\begin{aligned}
\left.\frac{d}{ds}\right|_{s=0} L_F(\gamma_s) &= \int_0^1 \left[\sum \frac{\partial F}{\partial x^i} X^i + \frac{d}{dt}\left(\sum \frac{\partial F}{\partial y^i} X^i\right) - \frac{d}{dt}\left(\sum \frac{\partial F}{\partial y^i}\right) X^i\right] dt \\
&= \left[\sum \frac{\partial F}{\partial y^i} X^i\right]_0^1 + \int_0^1 \sum \left[\frac{\partial F}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial F}{\partial y^i}\right)\right] X^i dt \\
&= \int_0^1 \sum \left[\frac{\partial F}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial F}{\partial y^i}\right)\right] X^i dt.
\end{aligned}$$

Hence, $\gamma(t)$ is a critical point of L_F if and only if

$$E_i(\gamma) := \frac{\partial F}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial F}{\partial y^i}\right) = 0. \tag{2.7}$$

This equation is the *Euler-Lagrange equation* of the functional L_F. The quantity E_i is written as

$$E_i(\gamma) = \frac{1}{F}\left(\sum G_{ij}\frac{dy^i}{dt} + 2G_i - \frac{\partial F}{\partial y^i}\frac{dF}{dt}\right),$$

where we put

$$2G_i = \sum \frac{\partial^2}{\partial y^i \partial x^i}\left(\frac{F^2}{2}\right) y^j - \frac{\partial}{\partial x^i}\left(\frac{F^2}{2}\right).$$

Putting $G^i = \sum G^{im} G_m$ and $y^i = dx^i/dt$, the Euler-Lagrange equation $E_i(\gamma) = 0$ implies the equation of the geodesic as follows:

$$\frac{d^2x^i}{dt^2} + 2G^i\left(x, \frac{dx}{dt}\right) = p\,\frac{dx^i}{dt}$$

for the function $p = d\log F/dt$, where we note that t is an arbitrary parameter of γ. In particular, if we take Finslerian arc length as the parameter t of γ, that is, $dt = F(x, dx)$, we obtain the equation of geodesic as follows:

$$\frac{d^2x^i}{dt^2} + 2G^i\left(x, \frac{dx}{dt}\right) = 0. \tag{2.8}$$

Example 2.11 *(Geodesics of Randers metrics)* Let $F = \alpha + \beta$ be a Randers metric, where $\alpha^2 = \sum g_{ij}(x)y^i y^j$ and $\beta = \sum b_i(x)y^i$. We denote by γ^i_{jk} the Christoffel symbol of the base Riemannian manifold (M, g). Let t be the Finslerian arc length.

We define the quantities $\nabla^g_j b_i$ and b^i_j by

$$\nabla^g_j b_i = \frac{\partial b_i}{\partial x^j} - \sum b_m \gamma^m_{ij}, \quad b^i_j = \sum g^{im}\left(\frac{\partial b_m}{\partial x^j} - \frac{\partial b_j}{\partial x^m}\right).$$

Then, by direct computations, we see that the functions G^i in (2.8) are given by

$$2G^i = \sum \gamma^i_{jk} y^j y^k + \frac{y^i}{F}\left(\sum \nabla^g_k b_j y^j y^k - \alpha \sum b_m b^m_j y^j\right) + \alpha \sum b^i_j y^j.$$

Hence the equation (2.8) with the Finslerian arc length parameter t is given by the following complicated form:

$$\frac{d^2x^i}{dt^2} + \sum \gamma^i_{jk}\frac{dx^j}{dt}\frac{dx^k}{dt} + \left(\sum \nabla^g_k b_j \frac{dx^j}{dt}\frac{dx^k}{dt} - \alpha \sum b_m b^m_l \frac{dx^l}{dt}\right)\frac{dx^i}{dt} + \alpha \sum b^i_j \frac{dx^j}{dt} = 0.$$

If we take the Riemannian arc length u as the parameter, that is, $du = \alpha(x, dx)$, we have $dt = du + \sum b_i(x)dx^i$, and so

$$\frac{dt}{du} = 1 + \sum b_i(x)\frac{dx^i}{du}.$$

Then the equation above is reduced to the following form:

$$\frac{d^2x^i}{du^2} + \sum \gamma^i_{jk}\frac{dx^j}{du}\frac{dx^k}{du} + \sum b^i_j \frac{dx^j}{du} = 0.$$

In particular, if the 1-form $\beta = \sum b_i(x)dx^i$ is closed, then $b^i_j = 0$ implies

$$\frac{d^2x^i}{du^2} + \sum \gamma^i_{jk}\frac{dx^j}{du}\frac{dx^k}{du} = 0,$$

and so any geodesic in (M, F) coincides with one in the base Riemannian manifold (M, g). Consequently it is shown that, if the 1-form β is closed, then (M, F) is *projectively equivalent* to the base Riemannian manifold (M, g).

Example 2.12 *(Geodesics of Funk metric)* Let $F = \alpha + \beta$ be the Funk metric on the unit ball $\mathbb{B} \subset \mathbb{R}^n$ stated in Example 1.1. In this case, the 1-form β is exact form, and so any geodesic in $(\mathbb{B}, F)$ is given by the one in the Hilbert's space $(\mathbb{B}, g_H)$. We shall show the equation of geodesic parameterized by its Finslerian arc length t. From Example 2.1, it is given by

$$\frac{d^2x^i}{dt^2} + \sum \gamma^i_{jk} \frac{dx^j}{dt}\frac{dx^k}{dt} + \left(\sum \nabla^g_k b_j \frac{dx^j}{dt}\frac{dx^k}{dt}\right)\frac{dx^i}{dt} = 0.$$

Since the Hilbert's metric $g_H = \sum g_{ij}(x) dx^i \otimes dx^j$ is given by

$$g_{ij} = \frac{1}{1 - \|x\|^2}\left(\delta_{ij} + \frac{x^i x^j}{1 - \|x\|^2}\right),$$

its Christoffel symbol γ^i_{jk} is given by

$$\gamma^i_{jk} = \frac{1}{1 - \|x\|^2}\left(x^j \delta^i_k + x^k \delta^i_j\right).$$

Then, because of

$$b_j = -\frac{1}{2}\frac{\partial}{\partial x_j} \log\left(1 - \|x\|^2\right) = \frac{x^j}{1 - \|x\|^2},$$

we obtain

$$\nabla^g_j b_i = \frac{\partial b_i}{\partial x^j} - \sum b_m \gamma^m_{ij} = \frac{\delta_{ij}}{1 - \|x\|^2} = g_{ij} - b_i b_j.$$

Hence the equation (2.8) is given by

$$\frac{d^2x^i}{dt^2} + \frac{1}{1 - \|x\|^2}\left(2\sum x^k \frac{dx^k}{dt} + \left\|\frac{dx}{dt}\right\|^2\right)\frac{dx^k}{dt} = 0.$$

Furthermore, from $dt = \sqrt{\sum g_{ij}(x) dx^i dx^j} + \sum b_i(x) dx^i$, we obtain

$$\frac{1}{1 - \|x\|^2}\left(2\sum x^k \frac{dx^k}{dt} + \left\|\frac{dx}{dt}\right\|^2\right) = \sum (g_{ij} - b_i b_j)\frac{dx^i}{dt}\frac{dx^j}{dt} + 2\sum b_i(x)\frac{dx^i}{dt} = 1.$$

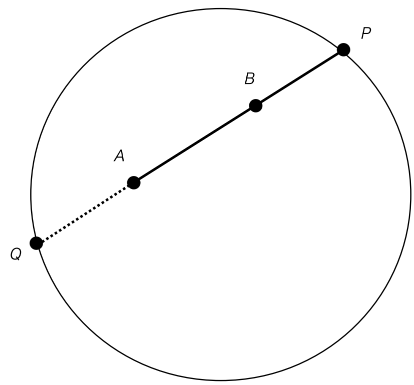
Consequently, the equation of geodesics in $(\mathbb{B}, F)$ is given by the following simple form:

$$\frac{d^2x^i}{dt^2} + \frac{dx^i}{dt} = 0.$$

The solution of this differential equation with initial conditions $A = (a^1, \cdots, a^n) = x(0)$ and $\lambda^i = \dfrac{dx^i}{dt}(0)(\neq 0)$ is given by $x^i = \lambda^i(1 - e^{-t}) + a^i$. Therefore any geodesic in $(\mathbb{B}, F)$ is given by a line in $\mathbb{B}$.

Let P the point $(\lambda^1 + a^1, \cdots, \lambda^n + a^n) = x(\infty)$ on the boundary $\partial\mathbb{B}$, and B a point on the line. We denote by AP (resp. BP) the Euclidean distance between the points A and P (resp. B and P). Because of

$$1 - e^{-t} = \frac{x^i - a^i}{\lambda^i} = \frac{\sqrt{\sum (x^i - a^i)^2}}{\sqrt{\sum (\lambda^i)^2}} = \frac{AB}{AP},$$



we obtain $e^{-t} = 1 - \dfrac{AB}{AP} = \dfrac{BP}{AP}$ which implies that the Funk's distance d_F is given by

$$d_F(A,B) = \log \frac{AP}{BP}.$$

Since the Hilbert's distance d_H is given by $d_H(A,B) = \left[d_F(A,B) + d_F(B,A)\right]/2$, the distance d_H is given by

$$d_H(A,B) = \frac{1}{2}\left(\log \frac{AP}{BP} + \log \frac{BQ}{AQ}\right).$$

2.4 The Jacobi fields and conjugate points

A variation $\Gamma_\gamma = \Gamma_\gamma(s,t)$ of a geodesic γ is said to be a *geodesic variation* if each s-curve γ_s is also a geodesic. Since each s-curve γ_s is a geodesic, we have $\nabla^H_{\mathcal{T}}\mathcal{T} = 0$.

Let X be a vector field along γ_s. Then, since $[\mathcal{S},\mathcal{T}] = 0$, we have

$$\nabla^H_{\mathcal{S}}\nabla^H_{\mathcal{T}}X - \nabla^H_{\mathcal{T}}\nabla^H_{\mathcal{S}}X = R^{HH}(\mathcal{S},\mathcal{T})X \tag{2.9}$$

along γ_s. From this equation, we get the so-called *the Jacobi equation*.

Proposition 2.13 *(The Jacobi Equation) Let γ be a geodesic, and $\mathcal{V}$ the variational field of a geodesic variation Γ_γ of γ in a Finsler manifold (M,F). Then $\mathcal{V}$ satisfies*

$$\nabla_t\nabla_t\mathcal{V} + R^{HH}(\mathcal{V},\dot\gamma)\dot\gamma = 0. \tag{2.10}$$

Definition 2.14 Let (M,F) be a Finsler manifold. The differential equation (2.10) is called the *Jacobi equation*. A vector field J along a geodesic satisfying (2.10):

$$\nabla_t\nabla_t J + R^{HH}(J,\dot\gamma)\dot\gamma = 0$$

is called a *Jacobi field* in (M, F).

By definition, the variational field $\mathcal{V}$ of a geodesic variation of a geodesic γ is a Jacobi field. Conversely, every Jacobi field along a geodesic γ is the variational field of some geodesic variation of γ. The differential equation (2.10) is linear and of second order, we have $2n$ linearly independent solution. Therefore, along any geodesic γ, the set of Jacobi field is a $2n$-dimensional vector space.

Let $\gamma \in \Gamma(p, q)$ be a geodesic segment in M. Then q is said to be *conjugate along* γ if there exists a Jacobi field $J(\neq 0)$ along γ such that J vanishes at p and q.

For $X \in T_pM$, we set $q = \exp_p X$. For an arbitrary $Y \in T_X(T_pM)$, we shall compute the differential $(\exp_p)_*Y$ at X:

$$(\exp_p)_*Y = \left.\frac{d}{ds}\right|_{s=0} \exp_p(X + sY).$$

To compute $(\exp_p)_*$, we define a geodesic variation Γ_γ of γ_X by $\Gamma_\gamma(s,t) = \exp_p t(X + sY)$. The variational field $J = \partial\Gamma_\gamma/\partial s$ is a Jacobi field along γ_X, and we have $J(1) = (\exp_p)_*Y$. The conjugate points are the image of the singularities by the exponential mapping.

Proposition 2.15 *Let* $\gamma_X(t) = \exp_p(tX)$ $(t \in I)$ *be the radial geodesic for* $X \in T_xM$. *Then* $\exp_p$ *is a local diffeomorphism if and only if* $q = \exp_p X$ *is not conjugate to* p *along* γ_X.

2.5 The second variational formula and index form

Let $\gamma : I \to M$ be a geodesic with unit speed. We shall compute the second variation of the length functional L_F. We shall compute

$$\left.\frac{d^2}{ds^2}\right|_{s=0} L_F(\gamma_s) = \int_0^1 \left[\frac{\partial}{\partial s}\frac{G\left(\nabla^H_{\mathcal{T}}\mathcal{S}, \mathcal{T}\right)}{\|\mathcal{T}\|}\right]_{s=0} dt$$

Differentiating with respect to s, we have

$$\frac{\partial}{\partial s}\frac{G\left(\nabla^H_{\mathcal{T}}\mathcal{S}, \mathcal{T}\right)}{\|\mathcal{T}\|} = -\frac{1}{\|\mathcal{T}\|^2}\frac{\partial\|\mathcal{T}\|}{\partial s}G\left(\nabla^H_{\mathcal{T}}\mathcal{S}, \mathcal{T}\right) + \frac{1}{\|\mathcal{T}\|}\frac{\partial}{\partial s}G\left(\nabla^H_{\mathcal{T}}\mathcal{S}, \mathcal{T}\right).$$

From (2.2) and (2.4), we get

$$\frac{\partial\|\mathcal{T}\|}{\partial s} = \frac{1}{\|\mathcal{T}\|}G\left(\nabla^H_{\mathcal{T}}\mathcal{S}, \mathcal{T}\right).$$

Furthermore

$$\begin{aligned}\frac{\partial}{\partial s}G\left(\nabla^H_{\mathcal{T}}\mathcal{S}, \mathcal{T}\right) &= G\left(\nabla^H_{\mathcal{S}}\nabla^H_{\mathcal{T}}\mathcal{S}, \mathcal{T}\right) + G\left(\nabla^H_{\mathcal{T}}\mathcal{S}, \nabla^H_{\mathcal{S}}\mathcal{T}\right)\\ &= G\left(\nabla^H_{\mathcal{T}}\nabla^H_{\mathcal{S}}\mathcal{S} + R^{HH}(\mathcal{S},\mathcal{T})\mathcal{S}, \mathcal{T}\right) + G\left(\nabla^H_{\mathcal{T}}\mathcal{S}, \nabla^H_{\mathcal{T}}\mathcal{S}\right).\end{aligned}$$

Consequently we have

$$
\begin{aligned}
&\frac{d^2 L_F(\gamma_s)}{ds^2} \\
&= \int_0^1 \frac{1}{\|\mathcal{T}\|}\left[G\left(\nabla^H_{\mathcal{T}}\nabla^H_{\mathcal{S}}\mathcal{S} + R^{HH}(\mathcal{S},\mathcal{T})\mathcal{S},\mathcal{T}\right) + \|\nabla^H_{\mathcal{T}}\mathcal{S}\|^2 - \frac{G\left(\nabla^H_{\mathcal{T}}\mathcal{S},\mathcal{T}\right)^2}{\|\mathcal{T}\|^2}\right] dt
\end{aligned}
$$

along γ_s. Since $\nabla^H_{\mathcal{T}}\mathcal{T} = 0$ and $\mathcal{V}(0) = \mathcal{V}(1) = 0$ imply

$$
\begin{aligned}
\int_0^1 \left[G\left(\nabla^H_{\mathcal{T}}\nabla^H_{\mathcal{S}}\mathcal{S},\mathcal{T}\right)\right]_{s=0} dt &= \int_0^1 \left[\frac{\partial}{\partial t} G\left(\nabla^H_{\mathcal{S}}\mathcal{S},\mathcal{T}\right)\right]_{s=0} dt \\
&= G\left(\nabla^H_{\mathcal{V}}\mathcal{V},\dot{\gamma}\right)_{t=1} - G\left(\nabla^H_{\mathcal{V}}\mathcal{V},\dot{\gamma}\right)_{t=0} = 0,
\end{aligned}
$$

we have

$$
\left.\frac{d^2}{ds^2}\right|_{s=0} L_F(\gamma_s) = \int_0^1 \left[G\left(R^{HH}(\mathcal{V},\dot{\gamma})\mathcal{V},\dot{\gamma}\right) + \|\nabla_t\mathcal{V}\|^2 - G\left(\nabla_t\mathcal{V},\dot{\gamma}\right)^2\right] dt. \tag{2.11}
$$

Let $\mathcal{V}^\top = G(\mathcal{V},\dot{\gamma})\dot{\gamma}$ be the tangential part of $\mathcal{V}$. We also denote by $\mathcal{V}^\perp$ the normal part of $\mathcal{V}$, that is, $\mathcal{V}^\perp = \mathcal{V} - \mathcal{V}^\top$. Then, $\nabla_t\dot{\gamma} = 0$ implies $\nabla_t\mathcal{V}^\top = \nabla_t(G(\mathcal{V},\dot{\gamma})\dot{\gamma}) = (\nabla_t\mathcal{V})^\top$ and $\nabla_t\mathcal{V}^\perp = \nabla_t\mathcal{V} - \nabla_t\mathcal{V}^\top = (\nabla_t\mathcal{V})^\perp$. Hence we have

$$
\|\nabla_t\mathcal{V}\|^2 = \|\nabla_t\mathcal{V}^\top\|^2 + \|\nabla_t\mathcal{V}^\perp\|^2 = G(\nabla_t\mathcal{V},\dot{\gamma})^2 + \|\nabla_t\mathcal{V}^\perp\|^2.
$$

Then, since $G(R^{HH}(\dot{\gamma},\dot{\gamma})\bullet,\bullet) = 0$ from (1.35) and $C(\dot{\gamma},\bullet,\bullet) = 0$ along γ from (1.6), we have $G(R^{HH}(\bullet,\bullet)\dot{\gamma},\dot{\gamma}) = 0$ from (1.36). Hence we get

$$
G(R^{HH}(\mathcal{V},\dot{\gamma})\mathcal{V},\dot{\gamma}) = G(R^{HH}(\mathcal{V}^\perp,\dot{\gamma})\mathcal{V}^\perp,\dot{\gamma}).
$$

Consequently, we obtain the *second variation formula* of L_F.

Proposition 2.16 (Second Variation Formula) *Let $\gamma : I \to M$ be any geodesic with unit speed, and Γ_γ a proper variation of γ, and $\mathcal{V}$ its variation field. Then*

$$
\left.\frac{d^2}{ds^2}\right|_{s=0} L_F(\gamma_s) = \int_0^1 \left[G(R^{HH}(\mathcal{V}^\perp,\dot{\gamma})\mathcal{V}^\perp,\dot{\gamma}) + \|\nabla_t\mathcal{V}^\perp\|^2\right] dt, \tag{2.12}
$$

where $\mathcal{V}^\perp$ is the normal part of $\mathcal{V}$.

Since (1.36) implies

$$
G(R^{HH}(\mathcal{V}^\perp,\dot{\gamma})\mathcal{V}^\perp,\dot{\gamma}) = -G(R^{HH}(\mathcal{V}^\perp,\dot{\gamma})\dot{\gamma},\mathcal{V}^\perp)
$$

along γ, and since $\mathcal{V}^\perp$ is normal to $\tilde{\gamma}$, we obtain

$$
G(R^{HH}(\mathcal{V}^\perp,\dot{\gamma})\mathcal{V}^\perp,\dot{\gamma}) = -\|\mathcal{V}^\perp\|^2 K(\mathcal{V}^\perp)
$$

for the flag curvature K. Hence the second variation formula (2.12) has the form

$$
\left.\frac{d^2}{ds^2}\right|_{s=0} L_F(\gamma_s) = \int_0^1 \left[\|\nabla_t\mathcal{V}^\perp\|^2 - \|\mathcal{V}^\perp\|^2 K(\mathcal{V}^\perp)\right] dt. \tag{2.13}
$$

Therefore we have

Proposition 2.17 *Let (M,F) be a Finsler manifold with non-positive flag curvature K. Then, the second variation of any geodesic satisfies*

$$\frac{d^2}{ds^2}\bigg|_{s=0} L_F(\gamma_s) > 0.$$

We define the *index form* on a Finsler manifold (M,F). Let γ be a unit speed geodesic in (M,F). We set

$$I(X,Y) = \int_0^1 \left[G(R^{HH}(X,\dot\gamma)Y,\dot\gamma) + G(\nabla_t X, \nabla_t Y)\right] dt \tag{2.14}$$

for normal proper vector fields X,Y along γ. The index form I is a symmetric bi-linear form on the space of normal proper vector fields. In fact, the Bianchi identity (1.36) implies

$$G(R^{HH}(X,\dot\gamma)Y,\dot\gamma) + G(R^{HH}(\dot\gamma,Y)X,\dot\gamma) + G(R^{HH}(Y,X)\dot\gamma,\dot\gamma) = 0.$$

Since the last term on the left hand side vanishes from (1.6) and (1.42), we have

$$G(R^{HH}(X,\dot\gamma)Y,\dot\gamma) = -G(R^{HH}(\dot\gamma,Y)X,\dot\gamma) = G(R^{HH}(Y,\dot\gamma)X,\dot\gamma)$$

along γ. Thus I is a symmetric bi-linear form: $I(X,Y) = I(Y,X)$.

Since (1.35) induces $G(R^{HH}(X,\dot\gamma)Y,\dot\gamma) = -G(R^{HH}(X,\dot\gamma)\dot\gamma,Y)$ along γ, if X and Y are proper, we have

$$\int_0^1 G\left(\nabla_t X, \nabla_t Y\right) = -\int_0^1 G\left(\nabla_t \nabla_t X, Y\right)$$

which implies

$$I(X,Y) = -\int_0^1 \left[G(\nabla_t\nabla_t X - R^{HH}(X,\dot\gamma)\dot\gamma, Y)\right] dt. \tag{2.15}$$

By the definition of I and (2.11), the second variation of L_F of unit speed geodesic is given by $I(X,X)$, and it can be thought as the Hessian of the length functional L_F. Thus, if γ is minimizing, then $I(X,X) \geq 0$ for any proper normal vector field X along γ. The following is a generalization of the well-known theorem in Riemannian geometry which shows that no geodesics is minimizing past its first conjugate point (e.g., Theorem 10.15 in [30]).

Theorem 2.18 *If $\gamma \in \Gamma(p,q)$ is a geodesic segment in a Finsler manifold (M,F) such that γ has an interior conjugate point to p, then there exists a proper normal vector field X along γ such that $I(X,X) < 0$. In particular, γ is not minimizing.*

We also consider the completeness of Finsler manifolds. For details of this contents, see Chapter VII in [32], Chapter 6 in [10]or Chapter 8 in [13].

Definition 2.19 A Finsler manifold (M,F) is said to be *geodesically complete* if the exponential mapping $\exp_x$ is defined on the whole of T_xM for every point $x \in M$.

We denote by $E(p,\delta)$ the subset of the closure $\overline{B(p,\delta)}$ consisting by the points joining by minimal geodesic with p. Then, if (M,F) is geodesically complete, the following three conditions are mutually equivalent.

(1) $E(p, \delta)$ is compact,

(2) $E(p, \delta) = \overline{B(p, \delta)}$ for all $\delta > 0$

(3) any ordered two points in M are joined by a minimal geodesic.

We shall introduce another completeness of Finsler manifolds.

Definition 2.20 A point sequence $\{p_m\}$ in (M, F) is called a *Cauchy sequence*, if for any $\varepsilon > 0$ there exists an integer N such that $d_F(p_i, p_j) < \varepsilon \ \ (i, j > N)$. Then (M, F) is said to be *metrically complete* if any Cauch sequence in M converges.

The following theorem is a natural generalization of the one in Riemannian geometry. The proof of it is omitted here. For the complete proof, see Chapter 8 in [13].

Theorem 2.21 (Hopf-Rinow Theorem) *Let (M, F) be a connected Finsler manifold. Then the following three conditions are mutually equivalent.*

(1) *(M, F) is geodesically complete.*

(2) *(M, F) is metrically complete with respect to the distance d_F.*

(3) *Any bounded closed subset of M is compact.*

3 Comparison theorems: Cartan–Hadamard theorem, Bonnet–Myers theorem, Laplacian and volume comparison

3.1 Cartan–Hadamard theorem

Before stating the theorems we need two notions:

(1) For a vector $y \in S_xM$, we define $c_y > 0$ to be the first number $r > 0$ such that there exists Jacobi field $J(t)$ along $c(t) = \exp(ty), \ 0 \leq t \leq r$, satisfying $J(0) = J(r) = 0$. c_y is called the conjugate value of y. Then we set

$$c_x = \inf_{y \in S_xM} c_y \qquad c_M = \inf_{x \in M} c_x,$$

called the *conjugate radius* at x and of M, resp.

(2) For a vector $y \in S_xM$, we define i_y to be the supremum of $r > 0$ such that $\exp_x(ty)$ is minimizing on $[0, r]$, and then the *injectivity radius* i_x at x is defined by $i_x = \inf\limits_{y \in S_xM} i_y$, and $i_M = \inf\limits_{x \in M} i_x$.

One can prove ([46]) that at each point $x \in M, \ i_y \leq c_y$ for all $y \in S_xM$, and hence $i_x \leq c_x$.

The generalization of Cartan-Hadamard theorem for Finsler spaces was first proved by Auslander [5], cf. [46].

Theorem 3.1 *Let (M, F) be a positively complete Finsler manifold. Suppose that the flag curvature satisfies $K \leq \lambda$. Then the conjugate radius satisfies $c_y \geq \pi/\sqrt{\lambda}$. In particular, if $K \leq 0$, then $c_y = \infty$ for any $y \in SM$. Hence $\exp_x \colon T_xM \to M$ is non-singular for any $x \in M$.*

A Finsler spacer is called a *Hadamard space* if it is positively complete, simply connected with $K \leq 0$. So, for a Hadamard space the exponential map $\exp_x : T_x M \to M$ is non-singular for all $x \in M$.

We mention that in [36] Neeb generalized the classical theorem of Cartan-Hadamard for Banach-Finsler manifolds endowed with a spray which have semi-negative curvature in the sense that the corresponding exponential function has a surjective expansive differential in every point.

A Finsler space is called *uniform* with uniformity constant C (cf. Egloff's work [17]) if for any $x \in M$ and $u, v \in T_x M$

$$C^{-1} g_u \leq g_v \leq C g_u.$$

We call a Finsler metric F *reversible*, if for all tangent vectors X we have: $F(-X) = F(X)$. Otherwise we call the metric *non-reversible* (or *irreversible*).

Theorem 3.2 ([46]) *Let (M, F) be a complete reversible uniform Finsler manifold. Assume that M is simply connected, and F satisfies $K \leq 0$. Then* $\exp_x : (T_x M, F_x) \to M$ *is distance quasi-nondecreasing:*

$$F_x(y_2 - y_1) \leq \sqrt{C} d(\exp_x(y_1) - \exp_x(y_2)), \qquad \forall\, y_1, y_2 \in T_x M.$$

This theorem expresses the dispersing of the geodesics emanating for the same point in the case of negative flag curvature. See a relating analysis about this in [10], p. 137.

In general, the conjugate radius is always less than or equal to the injectivity radius. However, in the case of positive flag curvature of an even-dimensional oriented Finsler manifolds they are equal.

Theorem 3.3 *Let (M, F) be an even-dimensional oriented closed Finsler manifold with $K > 0$. Then $i_M = c_M$.*

It can be proved that in this case M is simply connected (Synge theorem), cf. [5] and [27] for a different proof.

3.2 Bonnet–Myers theorem

The trace of the Riemann curvature $R_y : T_x M \to T_x M$ is called the *Ricci curvature (Ricci scalar)* of the Finsler manifold:

$$\mathrm{Ric}_y = \sum_{i=1}^{n} R^i{}_i(y).$$

If we use an orthonormal basis with respect to g_y such that $b_n = y/F(y)$, then the Ricci curvature can be expressed with the flag curvature. Namely, taking the flags $P_i = \mathrm{span}\,\{b_i, y\},\ i = 1, \ldots, n-1$, and the fact $R_y(y) = 0$, one can see easily that

$$\mathrm{Ric}_y = F^2(y) \sum_{i=1}^{n-1} K(P_i, y).$$

Theorem 3.4 *Let (M, F) be an n-dimensional positively complete Finsler manifold with* $\mathrm{Ric}_y \geq (n-1)\lambda$ *for all $y \in SM$, $\lambda > 0$. Then for any unit vector $y \in SM$, the conjugate value satisfies $c_y \leq \pi/\sqrt{\lambda}$.*

From this theorem of Bonnet and Myers it follows that the diameter of M is at most $\pi/\sqrt{\lambda}$, M is compact, and the fundamental group $\pi(M,x)$ is finite. See also [10].

3.3 Laplacian comparison

The Laplacian in the Finslerian case was given by Shen [46]. To obtain some Laplacian comparison theorems for the distance function, we need first the following function, which has an important role on the Rauch comparison theorem:

$$\mathrm{ct}_c(t) = \begin{cases} \sqrt{c}\cdot\cot(\sqrt{c}t), & c>0 \\ \frac{1}{t}, & c=0 \\ \sqrt{-c}\cdot\coth(\sqrt{-c}t), & c<0. \end{cases}$$

The S-curvature S_x is a real Finslerian quantity: For $y \in T_xM \setminus 0$, define the *distortion* of (M,F) as

$$\tau(y) := \log\frac{\sqrt{\det(g_{ij}(x,y))}}{\sigma},$$

where σ is the Busemann-Hausdorff volume form.

The *S-curvature* measures the rate of the distortion along geodesics:

$$S(y) := \frac{d}{dt}[\tau(\dot\gamma(t))]_{t=0},$$

where $\gamma(t)$ is the geodesic with $\dot\gamma(0)=y$. In local coordinates it can be expressed (cf. [46]) by

$$S(y) = N_i^i - \frac{y^i}{\sigma(x)}\frac{\partial\sigma}{\partial x^i}(x).$$

The Laplacian of a function f on M is given as follows: $\Delta f = \operatorname{div}(\nabla f) = \operatorname{div}(\ell^{-1}(df))$, where $\ell : TM \to T^*M$ is the Legendre transformation.

In the first theorem the assumption is given with the flag curvature, while in the second one with the Ricci curvature.

Theorem 3.5 ([51]) *Let $(M,F,d\mu)$ be a Finsler manifold, $r = d_F(p,\cdot)$ the distance function from a fixed point p. Suppose that the flag curvature of M satisfies $K(V,W) \le c$ for any $V,W \in TM$. Then for any vector X on M the following inequality holds whenever r is smooth:*

$$\Delta r \ge (n-1)\,\mathrm{ct}_c(r) - \|S\|,$$

where $\|S\|$ is the positive norm function of S-curvature which defined by

$$\|S\|_x = \sup_{X\in T_xM\setminus 0}\frac{S(X)}{F(X)}.$$

Theorem 3.6 ([51]) *Let $(M,F,d\mu)$ be a Finsler manifold with non-positive flag curvature. If the Ricci curvature of M satisfies* $\mathrm{Ric}_M \le c < 0$*, then the following inequality holds whenever r is smooth:*

$$\Delta r \ge \mathrm{ct}_c(r) - \|S\|$$

For the case where the Ricci curvature is bounded from below, the following comparison theorem is valid (cf. [51]):

Theorem 3.7 *Let $(M, F, d\mu)$ be a Finsler manifold with non-positive flag curvature. If the Ricci curvature of M satisfies $\mathrm{Ric}_M \geq (n-1)c$. Then the following inequality holds whenever r is smooth:*

$$\Delta r \leq \mathrm{ct}_c(r) + \|S\|.$$

3.4 Volume comparison

For real numbers c, Λ and positive integer n, let

$$V_{c,\lambda,n} := \mathrm{Vol}\,(S^{n-1}(1)) \int_0^r e^{\Lambda t} s_c(t)^{n-1}\, dt,$$

where

$$s_c(t) := \begin{cases} \sin(\sqrt{c}t), & c > 0 \\ t, & c = 0 \\ \mathrm{sh}\,(\sqrt{-c}t) & c < 0. \end{cases}$$

First the flag curvature and the norm of the S-curvature are bounded from above:

Theorem 3.8 ([51]) *Let $(M, F, d\mu)$ be a complete Finsler manifold which satisfies $K(V, W) \leq c$ and $\|S\| \leq \Lambda$. Then the function*

$$\frac{\mathrm{Vol}\,(B_p(r))}{V_{c,-\Lambda,n}(r)}$$

is monotone increasing for $0 < r < i_p$, where i_p is the injectivity radius of p. In particular, for $d\mu = dV_F$, the Busemann-Haussdorff volume form, one has

$$\mathrm{Vol}\,(B_p(r)) \geq V_{c,-\Lambda,n}(r), \qquad r \leq i_p.$$

Theorem 3.9 ([51]) *Let $(M, F, d\mu)$ be a complete Finsler manifold with non-positive flag curvature. If the Ricci curvature of M satisfies $\mathrm{Ric}_M \leq c < 0$ and $\|S\| \leq \Lambda$, then the function*

$$\frac{\mathrm{Vol}\,(B_p(r))}{V_{c,-\Lambda,2}(r)}$$

is monotone increasing for $0 < r < i_p$, where i_p is the injectivity radius of p. In particular, for $d\mu = dV_F$,

$$\mathrm{Vol}\,(B_p(r)) \geq \frac{\mathrm{Vol}\,(B^n(1))}{\mathrm{Vol}\,(B^2(1))} V_{c,-\Lambda,2}(r), \qquad r \leq i_p.$$

The Bishop-Gromov volume comparison theorem of Riemannian geometry uses a lower Ricci curvature bound to control the ratio between the volume of a metric ball in the space and the metric ball of the same size in the comparison space form. First Shen proved in [45] a generalization of the Bishop-Gromov theorem, where the Ricci curvature is bounded below and the mean covariation is bounded on both sides. As in Riemannian geometry, this volume comparison theorem leads to precompactness and homotopy finiteness theorems in any class of Finsler n-manifolds with fixed bounds on Ricci curvature, diameter and S-curvature. Here we state the version of Wu ([51]), using the Laplacian comparison theorem, see Theorem 3.7 above.

Theorem 3.10 ([51]) *Let $(M, F, d\mu)$ be a complete Finsler manifold. If the Ricci curvature of m satisfies $\mathrm{Ric}_M \geq (n-1)c$ and $\|S\| \leq \Lambda$, then the function*

$$\frac{\mathrm{Vol}\,(B_p(r))}{V_{c,\Lambda,n}(r)}$$

is monotone decreasing in r. In particular, for $d\mu = dV_F$,

$$\mathrm{Vol}\,(B_p(r)) \leq V_{c,\Lambda,n}(r).$$

4 Rigidity theorems: Finsler manifolds of scalar curvature and locally symmetric Finsler metrics

4.1 Curvature rigidity

Rigidity results state that under such and such assumptions about the curvature, the underlying Finsler structure must be either Riemannian or locally Minkowskian. A famous result of this type is Akbar-Zadeh's theorem ([2]) about spaces with constant flag curvature (a generalization of the sectional curvature).

Theorem 4.1 ([2]) *Let (M, F) be a compact connected boundaryless Finsler manifold of constant flag curvature λ.*

- *If $\lambda < 0$, then (M, F) is Riemannian.*
- *If $\lambda = 0$, then (M, F) is locally Minkowskian.*

In the paper [48], Shen addressed the case of negative but not necessarily constant flag curvature. He showed that in this case, if we impose the additional hypothesis that the S-curvature be constant, then the said rigidity still holds. As in the Akbar-Zadeh result, the compact boundaryless hypothesis may be replaced by a growth condition on the Cartan tensor.

Theorem 4.2 *Let (M, F) be a complete Finsler manifold with nonpositive flag curvature. Suppose that F has constant S-curvature and bounded mean Cartan curvature. Then F is weakly Landsbergian. Moreover, F is Riemannian at points where the flag curvature is negative.*

Any Finsler metric on a closed manifold is complete with bounded Cartan torsion. Therefore one immediately obtains

Corollary 4.3 *Let (M, F) be a closed Finsler manifold with negative flag curvature. If F has constant S-curvature, then it must be Riemannian.*

Let $K(P, y)$ be the flag curvature of the Finsler space (M, F). One calls (M, F) a Finsler space of scalar curvature if $K(P, y) = K(x, y)$ is independent of the flags P (associated with any fixed flagpole y)

Mo and Shen ([35]) proved that, for any n-dimensional $(n \geq 3)$ compact negatively curved Finsler spaces (M, F) of scalar curvature, F is a Randers metric.

Theorem 4.4 *Let (M, F) be an n-dimensional complete Finsler manifold of scalar flag curvature $K = K(x, y) \leq -1$ $(n \geq 3)$. Suppose that the Matsumoto torsion grows sub-exponentially at rate $k = 1$. Then F is a Randers metric.*

Any Finsler metric on a closed manifold is complete with bounded Cartan torsion, and hence bounded Matsumoto torsion. Therefore we have

Corollary 4.5 *Let (M, F) be an n-dimensional closed Finsler manifold of scalar curvature with negative flag curvature, $n \geq 3$. Then F is a Randers metric.*

If we impose the reversibility condition on the Finsler metric, we obtain the following

Corollary 4.6 *Let F be a reversible Finsler metric on a closed manifold of dimension $n \geq 3$. Suppose that F has of scalar curvature with negative flag curvature, then it is a Riemannian metric of constant curvature.*

Furthermore, Foulon in [18] proved a rigidity theorem for compact Finsler spaces. The curvature is assumed to be covariantly constant along a distinguished vector field on the homogeneous bundle of tangent half lines. Also, the flag curvature is assumed to be strictly negative, though variable. Under these conditions, the Finsler structure is shown to be Riemannian.

4.2 Curvature obstruction for Finsler surfaces

Given a Finsler structure on a surface one can define a canonical coframing on the unit sphere bundle in TM. The structure equations of this coframing are determined by three scalar functions, usually denoted by I, K and J. The function I is called the Cartan scalar of the structure and it vanishes iff the structure is Riemannian. When the Finsler structure is Riemannian, K is the usual Gaussian curvature. Landsberg structures are those for which J vanishes identically. In [9], Bao, Chern and Shen studied Landsberg surfaces for which the function K descends to a function on M. For compact surfaces with $K \leq 0$, their main result says that if the structure is Landsberg or K descends to a function on M, then the Finsler structure is Riemannian everywhere if K does not vanish identically, or is locally Minkowskian everywhere if K is identically zero.

Paternain proved a rigidity theorem for Finsler surfaces which are real analytic ([38]). It says that if the surface is of Landsberg type or its Gaussian curvature has no directional dependence, and if the Euler characteristic is negative, then the real analytic Finsler structure in question must be Riemannian. For comparison, Bao, S. S. Chern and Z. Shen have proved in [9] that if a Finsler surface is of Landsberg type or its Gaussian curvature has no directional dependence, and if the Gaussian curvature is everywhere negative, then the Finsler structure must be Riemannian. In Paternain's theorem, real analyticity is assumed, but the Gaussian curvature only needs to be negative on average. In [9], real analyticity is not needed, but then the Gaussian curvature is required to be pointwise negative.

The next theorem of Szabó ([49]) states that strongly convex Berwald surfaces are always trivial: locally Minkowskian or Riemannian.

Theorem 4.7 *Let (M, F) be a connected Berwald surface with smoth and strongly convex F on $TM \setminus 0$. Then*

- *if the Gauss curvature K vanishes identically, then F is locally Minkowskian,*
- *if the Gauss curvature K is not identically zero, then F is Riemannian.*

The proof is based on the following fact: If (M, F) is a Landsberg surface with smooth F on $TM \setminus 0$, then the value of the Gauss curvature K at any point of the indicatric S_xM

is determined by the Cartan scalar I accoding to the following formula:

$$K(t) = K(0)e^{\int_0^t I(\tau)\,d\tau}.$$

See also [10], p. 277.

4.3 Locally symmetric Finsler manifolds

Foulon ([19]) gave a sophisticated treatment of locally symmetric spaces for the Finsler case. Here local assumptions on curvature are much more flexible than in the Riemannian case. It is clear that the flag curvature governs the Jacobi equation and the second variation of length. Due to this, there exists a description of the flag curvature in terms of the dynamics of the geodesic flow. In fact, the generator of the geodesic flow of a Finsler metric is a second order differential equation. It may be observed ([18]) that, in the much more general context of second order differential equations, there is a natural operator which plays the role of the flag curvature.

A smooth reversible Finsler metric is said to be *parallel* if and only if $D_X R = 0$, i.e. if the curvature is parallel along the flow lines. A Finsler metric is called *locally symmetric* if for any chosen point the geodesic reflection is a local isometry.

There are many examples of non-Riemannian parallel Finsler spaces; for instance, $\mathbb{R}^n$, equipped with a Banach norm satisfying the positivity condition. But these cases are flat Finsler spaces. The most famous negatively curved symmetric Finsler space was invented by Hilbert in 1894. It is known as the Hilbert geometry of bounded convex sets in $\mathbb{R}^n$.

First one can prove the following

Proposition 4.8 ([19]) *A reversible, locally symmetric, C^3 Finsler metric is parallel.*

In contrast to the Riemannian case, the converse is not true in general. For instance, D. Egloff showed that a Hilbert geometry is locally symmetric if and only if it is Riemannian. The main rigidity theorem for Finsler manifolds with negative flag curvature states:

Theorem 4.9 ([19]) *A compact Finsler space with parallel negative curvature is isometric to a Riemannian locally symmetric negatively curved space.*

Combining Theorem 4.9 with Proposition 4.8 immediately implies the following

Corollary 4.10 *A locally symmetric compact Finsler space with negative curvature is isometric to a negatively curved Riemannian locally symmetric space.*

Theorem 4.9 contains, as a particular case, compact manifolds with constant curvature, for which this result was known by a theorem of Akbar Zadeh [2].

5 Closed geodesics on Finsler manifolds, sphere theorem and the Gauss–Bonnet formula

5.1 Closed geodesics

The study on closed geodesics on spheres is a classical and important problem in both dynamical systems and differential geometry. The results of V. Bangert in 1993 and J. Franks in 1992 prove that for every Riemannian metric on S^2 there exist infinitely many geometrically distinct closed geodesics. In contrast, in 1973, A. Katok ([24]) constructed a remarkable irreversible Finsler metric on S^2 which possesses precisely two distinct prime closed geodesics. See a fine analysis about it in [52].

Closed geodesics on a compact manifold with a Finsler metric F can be characterized as the critical points of the energy functional

$$E : \Lambda M \to \mathbb{R};\ E(\gamma) = \frac{1}{2}\int_0^1 F^2(\gamma'(t))\,dt.$$

Here ΛM is the free loop space consisting of closed H^1-curves $\gamma : S^1 := [0,1]/\{0,1\} \to M$ on the manifold M. On M there is an S^1-action $(u,\gamma) \in S^1 \times \Lambda M \mapsto u.\gamma \in \Lambda M;\ u.\gamma(t) = \gamma(t+u), t \in S^1$ leaving the energy functional invariant. In addition there is the mapping $^m : \gamma \in \Lambda M \mapsto \gamma^m \in \Lambda M;\ \gamma^m(t) = \gamma(mt); t \in S^1$ and $E(\gamma^m) = m^2 E(\gamma)$. Here γ^m is the m-fold cover of γ. A closed geodesic c is called *prime* if there is no closed geodesic c_1 and no integer $m > 1$ with $c = c_1^m$. Recall that a Finsler metric F is called reversible, if for all tangent vectors X we have: $F(-X) = F(X)$. Otherwise we call the metric non-reversible. We call two closed geodesics $c_1, c_2 : S^1 \to M$ of a non-reversible Finsler metric on a differentiable manifold M geometrically equivalent if their traces $c_1(S^1) = c_2(S^1)$ coincide and if their orientations coincide. The equivalence class is also called a geometric closed geodesic. For a closed geodesic c_1 there is a prime closed geodesic c such that the set of all geometrically equivalent closed geodesics consists of $u.c^m;\ m \geq 1, u \in S^1$. Let $\theta : \Lambda M \to \Lambda M$ be the orientation reversing, i.e. $\theta(c)(t) = c(1-t)$. For a non-reversible Finsler metric this mapping in general does not leave the energy functional invariant. And in general for a closed geodesic c the curve θc is not a geodesic.

The second order behaviour of the energy functional in a neighborhood of a closed geodesic is determined by its index form $\mathcal{H}_c$ which equals the hessian $d^2E(c)$ of the energy functional by the second variational formula, cf. [39]. The index $\mathrm{ind}\,(c)$ of the closed geodesic c is the index of the index form $\mathcal{H}_c$ i.e. it is the maximal dimension of a subspace on which $\mathcal{H}_c$ is negative definite. The nullity (c) is the nullity of the index form $\mathcal{H}_c$ minus 1. This convention is used since due to the S^1-action the nullity of the index form $\mathcal{H}_c$ is at least 1. We call a Finsler metric *bumpy*, if all closed geodesics are non-degenerate, i.e. for all closed geodesics c the nullity $(c) = 0$ vanishes. We call two prime closed geodesics c and d *distinct* if there is no $\theta \in (0,1)$ such that $c(t) = d(t+\theta)$. We shall omit the word "distinct" for short when we talk about more than one prime closed geodesics. In recent years, geodesics and closed geodesics on Finsler manifolds have got more attentions. We refer readers to [11] of Bao, Robles and Shen, and [31] of Long and the references therein for recent progress in this area.

Note that by the classical theorem of Lyusternik–Fet in 1951, there exists at least one closed geodesic on every compact Riemannian manifold. Because the proof is variational, this result works also for compact Finsler manifolds. In [41] of 2005, Rademacher obtained existence of closed geodesics on n–dimensional Finsler spheres under pinching conditions which generalizes results on Riemannian manifolds. There are few results on the existence of multiple closed geodesics on Finsler spheres without pinching conditions. Read the story in [16], where the latest result in this direction states the following one, specially for bumpy irreversible Finsler rationally homological n–spheres without pinching conditions.

Theorem 5.1 ([16], [42]) *For every bumpy Finsler metric F on every rationally homological n-sphere S^n with $n \geq 2$, there exist at least two distinct prime closed geodesics.*

Note that the proof in [16] uses only the $\mathbb{Q}$-homological properties of the Finsler manifold, thus one can carry out the proof of this theorem just for n-dimensional spheres. Meanwhile, the main ingredients of the proof in [42] are the relation between the average indices of closed geodesics for metrics with only finitely many closed geodesics and a detailed analysis of the sequence of Morse indices $\mathrm{ind}\,(c^m)$ of the coverings c^m of a prime closed geodesic c using a formula due to Bott as well as a careful discussion of the Morse inequalities.

We mention that in [28] Kristály, Kozma and Varga studied the critical point theory on the loop space of closed geodesics, and proved that the energy integral satisfies the so-called Palais-Smale condition. This implies that if M_1 and M_2 are closed submanifolds of M, and (M, F) is a dominated by a Riemannian metric, then under different assumptions on M_1 and M_2, there exist infinitely many Finsler geodesics on M joining the submanifolds M_1 and M_2.

5.2 Sphere theorem

The classical sphere theorem states that a simply connected and compact manifold of dimension n with a Riemannian metric whose sectional curvature K satisfies $\frac{1}{4} < K \leq 1$ is homeomorphic to the n-sphere. In the proof the homeomorphism is constructed using the estimate for the injectivity radius $\mathrm{inj} \geq \pi$ and the Toponogov comparison theorem. W. Klingenberg showed that one can give a different proof without using the Toponogov comparison theorem: The injectivity radius estimate gives as lower bound for the length of a closed geodesic the value 2π: Then a Rauch comparison argument shows that the Morse index of a closed geodesic is at least $n-1$: From the Morse theory of the energy functional on the free loop space one can conclude, that the free loop space is $(n-2)$-connected. This implies that the manifold is homotopy equivalent to the n-sphere. P. Dazord ([15]) remarked that this proof extends to the case of a reversible Finsler metric, i. e. a Finsler metric F for which $F(-X) = F(X)$ for all tangent vectors. The flag curvature, which depends on a flag $(V; \sigma)$ consisting of a non-zero tangent vector V and a 2-plane σ in which V lies, generalizes the sectional curvature.

In [39] Rademacher considered also non-reversible Finsler metrics, introduced the reversibility $\lambda = \lambda(M, F)$ of a Finsler metric F on a compact manifold M:

$$\lambda := \max\{F(-X) \,\|\, X \in TM,\ F(X) = 1\}.$$

Obviously $\lambda \geq 1$ and $\lambda = 1$ if and only if F is reversible. The reversibility enters in the following generalization of the injectivity radius estimate for Riemannian metrics:

Theorem 5.2 ([39]) *Let (M, F) be a simply connected, compact Finsler manifold of dimension $n \geq 2$ with reversibility λ and flag curvature $\left(1 - \frac{1}{1+\lambda}\right)^2 < K \leq 1$. Then the length of a closed geodesic is at least $\pi(1 + \frac{1}{\lambda})$.*

Using a Hamiltonian description A. Katok defined in [24] a 1-parameter family $F_\epsilon; \epsilon \in [0, 1)$ of Finsler metrics on the 2-sphere. For $\epsilon = 0$ this is the standard Riemannian metric, for $\epsilon \in (0, 1)$ these metrics are nonreversible and for irrational parameter ϵ these metrics have exactly two geometrically distinct closed geodesics, cf. [39, 52]. These two geodesics differ by orientation. These examples show that the estimate for the length of a closed geodesic in Theorem 5.2 is sharp.

Using a Rauch comparison argument and the Morse theory of the energy functional on the free loop space one concludes from Theorem 5.2 the following Sphere Theorem:

Theorem 5.3 *A simply connected and compact Finsler manifold of dimension $n \geq 3$ with reversibility λ and with flag curvature $\left(1 - \frac{1}{1+\lambda}\right)^2 < K \leq 1$ is homotopy equivalent to the n-sphere.*

It remains an open problem whether one can improve the sphere theorem in the nonreversible case by choosing the lower curvature bound $1/4$ as in the reversible case.

5.3 Gauss-Bonnet-Chern theorem for Finsler manifolds

Historically the Gauss-Bonnet formula gives the relationship between curvature and angular excess. The notion of angular excess for a geodesic triangle on a surface is evident, however for more general figures in higher dimension, angular excess becomes much more complex. The modern view of the Gauss-Bonnet formula is that the curvature of a Riemannian manifold reflects the topology of the space: for a compact and oriented Riemannian manifold without boundary one has:

$$\chi(M) = \int_M \Psi(\zeta) dv(\zeta),$$

where $\chi(M)$ is the Euler characteristic of the manifold, $\Psi(\zeta)$ is the Pfaffian of the Riemannian curvature suitable normalized, dv is the Riemannian volume form.

In 1944 Chern gave a simple proof for the famous Riemannian formula using the so-called method of transgression. Extending the formula to Finsler spaces is natural; while the method of transgression requires the lifting of the curvature to the sphere bundle, for a Finsler manifold the curvature form already lives on the sphere bundle. This was done in 1996 by Bao and Chern ([7]): The Gauss-Bonnet-Chern theorem was established for Finsler manifolds under an additional assumption that the volume function $V(x) = \mathrm{Vol}(S_x M)$ of the tangent sphere $S_x M$ is a constant. This happens in the case of Landsberg manifolds. Consequently, Bao, Chern, and Shen in [8] showed how the Gauss-Bonnet integrand is simplified on 4-dimensional Landsberg spaces. Further, one can rewrite the horizontal part of the simplified Gauss-Bonnet integrand, in a way which formally generalizes an identity derived for Riemannian 4-manifolds. The integrand is written in terms of the hh-curvature tensor of the Chern-Rund connection and Cartan's tensor A_{ijk}, containing Vol(Finsler S^3).

Lackey in [29] generalized the Gauss-Bonnet formula for even-dimensional closed oriented Finsler manifolds without any additional restriction on the Finsler structure.

Theorem 5.4 *Let (M, F) be such a Finsler manifold, and let ψ be a section of its projective sphere bundle $\pi\colon SM \to M$, possibly with isolated singularities. Let (ω^i_j) be a torsion-free Finslerian connection with curvature forms (Ω^i_j). Then*

$$\int_M \psi^* \left[\frac{-1}{\mathit{Vol}(x)} (\Omega + \mathcal{F}) \right] = \chi(M),$$

where $\mathit{Vol}(x)$ is the volume of the tangent sphere $S_x M$,

$$\Omega = \frac{(-1)^{m-1}}{2^{2m-1}(m-1)!m!} \mathit{Pf}(\Omega^j_k),\ m = \frac{1}{2}\dim M,$$

is a $2m$*-form on* SM *expressed in terms of* Ω^j_k *as defined in the Riemannian case by S. S. Chern, and* $\mathcal{F}$ *is a polynomial in the entries of* $\sigma^{jk} := \omega^{jk} + \omega^{kj}$, $\Sigma^{jk} := \Omega^{jk} + \Omega^{kj}$, Ω^i_j *and* $d \ln \mathit{Vol}(x)$.

The above Gauss-Bonnet formula is established using an arbitrary torsion-free connection whose structure equations in a preferred orthogonal coframe $\{\omega^j\}$ with $\omega^n := F_{y^i} dx^i$ take the following form:

$$d\omega^j + \omega^j_k \wedge \omega^k = 0, \; \omega_{jk} + \omega_{kj} = M_{jkl}\omega^l + A_{jkl}\omega^l_n.$$

This general setting allows to select a new torsion-free connection involving $d \ln \mathrm{Vol}(x)$ such that the integrand of the Gauss-Bonnet formula consists solely of polynomials in the connection forms, curvature forms, and covariant derivatives thereof. One can make a choice of ω^i_j with

$$M_{jkl} dx^l := \frac{1}{n-1}\Big(a l_j l_k - b h_{jk}\Big) d \ln \mathrm{Vol}(x),$$

where a, b are arbitrary constants with $a + b = 1$, $l_i := F_{y^i}$ and $h_{jk} := g_{jk} - l_j l_k$.

Bibliography

[1] T. Aikou: Some remarks on the geometry of tangent bundles of Finsler spaces *Tensor, N. S.* **52** (1993) 234–242

[2] H. Akbar-Zadeh: Sur les espaces de Finsler a courbures sectionnelles constantes *Acad. Roy. Belg. Bull. Cl. Sci.* (5) **74** (10), (1988) 281–322

[3] M. Anastasiei: A historical remark on the connections of Chern and Rund *Contemp. Math.* **196** (1996) 171–176

[4] P. Antonelli, R. Ingarden and M. Matsumoto: *The theory of Sprays and Finsler Spaces with Applications in Physics and Biology* (Kluwer Academic Publishers, 1993)

[5] L. Auslander: On curvature in Finsler geometry *Trans. Amer. Math. Soc.* **79** (1955) 378–388

[6] D. Bao and S.S. Chern: On a notable connection in Finsler geometry *Houston J. Math.* **19** (1), (1993) 135–180

[7] D. Bao and S.S. Chern: A note on the Gauss-Bonnet Theorem for Finsler spaces *Ann. Math.* **143** (1996) 233–252

[8] D. Bao, S.S. Chern and Z. Shen: On the Gauss-Bonnet integrand for 4-dimensional Landsberg spaces *In: Finsler geometry (Seattle, WA, 1995)* (Contemp. Math., **196** Providence, RI: Amer. Math. Soc., 1996) 15–26

[9] D. Bao, S.S. Chern and Z. Shen: Rigidity issues on Finsler surfaces, Collection of papers in honour of Academician Radu Miron on his 70th birthday *Rev. Roumaine Math. Pures Appl.* **42** (9–10), (1997) 707–735

[10] D. Bao, S. S. Chern and Z. Shen: *An Introduction to Riemann-Finsler Geometry* (Grad. Texts Mathem. 200, Springer Verlag, 2000)

[11] David Bao, Colleen Robles and Zhongmin Shen: Zermelo navigation on Riemannian manifolds *J. Differential Geom.* **66** (3), (2004) 377–435

[12] S.S. Chern: Local equivalence and Euclidean connection in Finsler spaces *Sci. Rep. Nat. Tsing Hua Univ. Ser. A* **5** (1948) 95–121

[13] S.S. Chern, W.H. Chen and K.S. Lam: *Lectures On Differential Geometry* (World Scientific, 1999)

[14] S.S. Chern and Z. Shen: *Riemannian-Finsler Geometry* (Nankai Tracts in Mathematics, **6**, World Scientific, 2005)

[15] P. Dazord: Variétés finslériennes de dimension δ—pincées *C. R. Acad. Sc. Paris* **266** (1968) 496–498

[16] Huagui Duan and Yiming Long: Multiple closed geodesics on bumpy Finsler n-spheres *Journal of Differential Equations* **233** (2007) 221–240

[17] Daniel Egloff: Uniform Finsler Hadamard manifolds *Annales de l'institut Henri Poincaré (A) Physique théorique* **66** (3), (1997) 323–357

[18] Patrick Foulon: Locally symmetric Finsler spaces in negative curvature *C. R. Acad. Sci. Paris Sér. I Math.* **324** (10), (1997) 1127–1132

[19] Patrick Foulon: Curvature and global rigidity in Finsler manifolds; Special issue for S.S. Chern. *Houston J. Math.* **28** (2), (2002) 263–292

[20] I. Hasegawa, V. S. Sabau and H. Shimada: Randers spaces of constant flag curvature induced by almost contact metric structures *Hokkaido Math. J.* **33** (2004) 215–232

[21] Y. Ichijyō: Finsler manifolds modeled on a Minkowski space *J. Math. Kyoto Univ.* **16** (1976) 639–652

[22] Y. Ichijyō: On special Finsler connections with the vanishing hv-curvature tensor *Tensor, N. S.* **32** (1978) 149–155

[23] S. Ishihara and M. Konishi: *Differential Geometry of Fibered Manifolds* (Publications of the study group of geometry, Vol. 8, 1973)

[24] A. Katok: Ergodic properties of degenerate integrable Hamiltonian systems *Izs. Akad. Nauk SSSR Ser. Mat.* **37** (1973) 539–576; (engl. transl.) *Math. USSR Izv.* **7**, 535–572 (1973)

[25] S. Kikuchi: On the condition that a space with (α, β)-metric be locally Minkowskian *Tensor N. S.* **33** (1979) 242–246

[26] L. Kozma: On holonomy groups of Landsberg manifolds *Tensor N. S.* **62** (2000) 87–90

[27] L. Kozma and Ioan Radu Peter: Weinstein's theorem for Finsler manifolds *Journal of Mathematics of Kyoto University* **46** (2006) 377–382

[28] L. Kozma, A. Kristály and C. Varga: Critical point theorems on Finsler manifolds *Beiträge Algebra Geom.* **45** (2004) 47–59

[29] Brad Lackey: On the Gauss-Bonnet-Chern theorem in Finsler geometry *In: Handbook of Finsler geometry. Vol. 1, 2,* (Dordrecht: Kluwer Acad. Publ., 2003) 491–509

[30] J. M. Lee: *Riemannian manifolds – An Introduction to Curvature* (Springer, 1997)

[31] Yiming Long: Multiplicity and stability of closed geodesics on Finsler 2-spheres *J. Eur. Math. Soc. (JEMS)* **8** (2), (2006) 341–353

[32] M. Matsumoto: *Foundations of Finsler Geometry and Special Finsler Spaces* (Japan: Kaiseisha-Press, 1986)

[33] S. Mazur and S. Ulam: Sur les transformations isométriques d'espaces vectoriels normés *C. R. Acad. Sci. Paris* **194** (1932) 946–948

[34] X. Mo: *An Introduction to Finsler Geometry* (Peking University Series in Mathematics, **V**ol. 1, World Scientific, 2006)

[35] Xiaohuan Mo and Zhongmin Shen: On negatively curved Finsler manifolds of scalar curvature *Canad. Math. Bull.* **48** (1), (2005) 112–120

[36] Karl-Hermann Neeb: A Cartan-Hadamard theorem for Banach-Finsler manifolds *In: Proceedings of the Conference on Geometric and Combinatorial Group Theory, Part II (Haifa, 2000)* (Geom. Dedicata **9**5, 2002) 115–156

[37] T. Okada: On models of projectively flat Finsler spaces of negative curvature *Tensor, N. S.* **40** (1983) 117–124

[38] Gabriel P. Paternain: Finsler structures on surfaces with negative Euler characteristic *Houston J. Math.* **23** (3), (1997) 421–426

[39] Hans-Bert Rademacher: A sphere theorem for non-reversible Finsler metrics *Math. Ann.* **328** (3), (2004) 373–387

[40] Hans-Bert Rademacher: Nonreversible Finsler metrics of positive flag curvature *In: A sampler of Riemann-Finsler geometry* (Math. Sci. Res. Inst. Publ., 50, Cambridge: University Press, 2004) 261–302

[41] Hans-Bert Rademacher: Existence of closed geodesics on positively curved Finsler manifolds *Preprint* `arXiv:math.DG/0503153` (2005)

[42] Hans-Bert Rademacher: The second closed geodesic on Finsler spheres of dimension $n > 2$ *Preprint* `arXiv:math.DG/0608160` (2006)

[43] H. Rund: *The Differential Geometry of Finsler Spaces* (Springer, 1959)

[44] Z. Shen: On some special spaces in Finsler Geometry *Preprint* (1994)

[45] Zhongmin Shen: Volume comparison and its applications in Riemann-Finsler geometry *Adv. Math.* **128** (2), (1997) 306–328.

[46] Z. Shen: *Lectures on Finsler geometry,* (Singapore: World Scientific, 2001)

[47] Z. Shen: Two-dimensional Finsler metrics with constant curvature *Manuscr. Math.* **109** (2002) 349–366

[48] Zhongmin Shen: Finsler manifolds with nonpositive flag curvature and constant S-curvature *Math. Z.* **249** (3), (2005) 625–639

[49] Z. I. Szabó: Positive definite Berwald spaces *Tensor N. S.* **35** (1981) 25–39

[50] A. C. Thompson: *Minkowski Geometry* (Cambridge: University Press, 1996)

[51] B. Y. Wu and Y. L. Xin: Comparison theorems in Finsler geometry and their applications *Math. Ann.* **337** (1), (2007) 177–196

[52] W. Ziller: Geometry of the Katok examples *Ergod. Th. & Dynam. Sys.* **3** (1), (1983) 135–157

Tadashi Aikou

Department of Mathematics and Computer Science, Kagoshima University,
1-2-35 Korimoto, Kagoshima, 890 Japan

E-mail: `aikou@sci.kagoshima-u.ac.jp`

László Kozma

Institute of Mathematics, University of Debrecen,
H-4010 Debrecen, P O Box 12, Hungary

E-mail: `kozma@math.unideb.hu`

This page intentionally left blank

Demeter Krupka and David Saunders

Morse theory and nonlinear differential equations

Thomas Bartsch, Andrzej Szulkin
and Michel Willem

Contents

1 Introduction

Classical Morse theory's object is the relation between the topological type of critical points of a function f and the topological structure of the manifold X on which f is defined. Traditionally Morse theory deals with the case where all critical points are non-degenerate and it relates the index of the Hessian of a critical point to the homology of the manifold. A closely related more homotopy theoretical approach goes back to Lusternik and Schnirelmann allowing for degenerate critical points. Nowadays these topics are subsumed under the more general heading of critical point theory. The classical Morse theory in finite dimension is described in [42], first extensions to Hilbert manifolds in [43, 47]. In this survey we treat Morse theory on Hilbert manifolds for functions with degenerate critical points.

In section 2 the general case of a topological space X and a continuous function $f : X \to \mathbb{R}$ is considered.

The topological type of the "critical set" between a and b is described by the Morse polynomial $M_f(t; a, b)$. The relative homology of the pair (f^a, f^b) is described by the Poincaré polynomial $P_f(t; a, b)$. By definition $f^a = \{x \in X : f(x) \leq a\}$. Then there exists a polynomial $Q(t)$ with non-negative integer coefficients such that

$$M_f(t; a, b) = P_f(t; a, b) + (1 + t)Q(t).$$

The Morse inequalities follow from this relation.

In section 3 the Morse polynomial is computed in the case of isolated critical points of finite Morse index. Besides the Morse lemma, the main results are the Shifting Theorem and the Splitting Theorem, due to Gromoll and Meyer. As in [17] and [39] we consider the case of an infinite-dimensional Hilbert space.

In section 4 we give some elementary, but typical, applications to semi-linear elliptic problems. We consider in particular the Dirichlet problem

$$\begin{cases} -\Delta u = g(u), & \text{in } \Omega, \\ \quad u = 0, & \text{on } \partial\Omega. \end{cases}$$

This section contains also a general bifurcation theorem.

Another typical problem to which Morse theory can be applied is the study of existence of periodic solutions for the second order Newtonian systems of ordinary differential equations

$$-\ddot{q} = V_q(q,t), \quad q \in \mathbb{R}^N.$$

However, since we consider the more general and more difficult case of first order Hamiltonian systems in section 6 in some detail, we do not discuss second order systems here and refer the reader e.g. to [39]. We do however discuss the problem of the existence of closed geodesics on a riemannian manifold. Work on this problem was vital for the development of Morse theory, it still offers open problems, and it contains difficulties also present in the more general problem of first order Hamiltonian systems.

Section 5 contains a Morse theory for functionals of the form

$$\Phi(x) = \frac{1}{2}||x^+||^2 - \frac{1}{2}||x^-||^2 - \psi(x)$$

defined on a Hilbert space E, where $x = x^+ + x^0 + x^- \in E = E^+ \oplus E^0 \oplus E^-$, $\dim E^0 < \infty$, $\dim E^\pm = \infty$ and $\nabla\psi$ is a compact operator. The Morse indices are infinite but using a suitable cohomology theory, the Poincaré and the Morse polynomials can be defined. Theorem 5.7 contains the Morse inequalities. The local theory is also extended to this setting.

Section 6 contains applications of the results of section 5 to Hamiltonian systems. We consider the existence of periodic solutions of the system

$$\dot{z} = JH_z(z,t), \quad z \in \mathbb{R}^{2N}$$

where J is the standard symplectic matrix. The problem is delicate since the Morse indices of the corresponding critical points are infinite. The existence of nontrivial solutions of asymptotically linear sytems is considered in Theorems 6.5 and 6.6. Another application is the existence of at least 2^{2N} periodic solutions when the Hamiltonian is periodic in all variables and when all the periodic solutions are nondegenerate.

We refer to [13] and [14] for surveys on Morse theory with historical remarks. The homotopy index, introduced by C. Conley, is a generalization of the Morse index. We refer to the monograph by Rybakowski [44], and, for a Morse theory based on the Conley index, to Benci [9]. A useful survey of Morse theory in the context of nonsmooth critical point theory is due to Degiovanni [23]. Morse theory on infinite-dimensional manifolds for functions with infinite Morse index is due to Witten and Floer. Here we refer the reader to [34, 40, 45, 46].

2 Global theory

2.1 Preliminaries

Let H_* denote a homology theory with coefficients in a field $\mathbb{F}$. Thus H_* associates to a pair (X, Y) of topological spaces $Y \subset X$ a sequence of $\mathbb{F}$-vector spaces $H_n(X, Y)$, $n \in \mathbb{Z}$, and to a continuous map $f : (X, Y) \to (X', Y')$ a sequence of homomorphisms $f_* : H_*(X, Y) \to H_*(X', Y')$ satisfying the Eilenberg-Steenrod axioms. In particular, given a triple (X, Y, Z) of spaces $Z \subset Y \subset X$ there is a long exact sequence

$$\dots \xrightarrow{j_{n+1}} H_{n+1}(X,Y) \xrightarrow{\partial_{n+1}} H_n(Y,Z) \xrightarrow{i_n} H_n(X,Z) \xrightarrow{j_n} \\ \xrightarrow{j_n} H_n(X,Y) \xrightarrow{\partial_n} H_{n-1}(Y,Z) \xrightarrow{i_{n-1}} \dots$$

Given two homomorphisms $V_1 \xrightarrow{\varphi_1} V_2 \xrightarrow{\varphi_2} V_3$ with image $\varphi_1 = \ker \varphi_2$ we have $\dim V_2 = \operatorname{rank} \varphi_1 + \operatorname{rank} \varphi_2$. It follows that

$$\begin{aligned} \dim H_n(X,Z) &= \operatorname{rank} i_n + \operatorname{rank} j_n \\ &= \dim H_n(Y,Z) + \dim H_n(X,Y) - \operatorname{rank} \partial_{n+1} - \operatorname{rank} \partial_n \end{aligned} \tag{2.1}$$

holds for $n \in \mathbb{Z}$. We define the formal series

$$P(t; X, Y) := \sum_{n=0}^{\infty} [\dim H_n(X,Y)] t^n$$

and

$$Q(t; X, Y, Z) := \sum_{n=0}^{\infty} [\operatorname{rank} \partial_{n+1}] t^n.$$

It follows from (2.1) and $\partial_0 = 0$ that:

$$P(t; X, Y) + P(t; Y, Z) = P(t; X, Z) + (1+t) Q(t; X, Y, Z). \tag{2.2}$$

This is correct even when some of the coefficients are infinite so that the series lie in $\overline{\mathbb{N}}_0[[t]]$, $\overline{\mathbb{N}}_0 = \mathbb{N}_0 \cup \{\infty\}$.

Proposition 2.1 *Given a sequence $X_0 \subset X_1 \subset \ldots \subset X_m$ of topological spaces there exists a series $Q(t) \in \overline{\mathbb{N}}_0[[t]]$ such that*

$$\sum_{k=1}^{m} P(t; X_k, X_{k-1}) = P(t; X_m, X_0) + (1+t) Q(t).$$

Proof By (2.2) the result is true for $m = 2$. If it holds for $m - 1 \geq 2$ then

$$\begin{aligned} \sum_{k=1}^{m} P(t; X_k, X_{k-1}) &= P(t; X_m, X_{m-1}) + P(t; X_{m-1}, X_0) + (1+t) Q_{m-1}(t) \\ &= P(t; X_m, X_0) + (1+t) Q(t; X_m, X_{m-1}, X_0) + (1+t) Q_{m-1}(t) \\ &= P(t; X_m, X_0) + (1+t) Q_m(t). \end{aligned}$$

□

Remark 2.2 Let $m_n, p_n \in \mathbb{N}_0$ be given for $n \geq 0$, and set $M(t) := \sum_{n=0}^{\infty} m_n t^n$, $P(t) := \sum_{n=0}^{\infty} p_n t^t$. Then the following are equivalent:

(i) There exists a series $Q(t) \in \mathbb{N}_0[[t]]$ with $M(t) = P(t) + (1+t)Q(t)$

(ii) $m_n - m_{n-1} + m_{n-2} - + \ldots + (-1)^n m_0 \geq p_n - p_{n-1} + p_{n-2} - + \ldots + (-1)^n p_0$ holds for every $n \in \mathbb{N}_0$

In fact, if (i) holds with $Q(t) = \sum_{n=0}^{\infty} q_n t^n$ we have $m_n = p_n + q_n + q_{n-1}$ where $q_{-1} := 0$, hence

$$\sum_{i=0}^{n} (-1)^{n-i} m_i = \sum_{i=0}^{n} (-1)^{n-i} (p_i + q_i + q_{i-1}) = \sum_{i=0}^{n} (-1)^{n-i} p_i + q_n \geq \sum_{i=0}^{n} (-1)^{n-i} p_i.$$

On the other hand, if (ii) holds we define $q_{-1} := 0$,

$$q_n := \sum_{i=0}^{n} (-1)^{n-i} m_i - \sum_{i=0}^{n} (-1)^{n-i} p_i \in \mathbb{N}_0, \qquad n \geq 0,$$

and obtain $q_n + q_{n-1} = m_n - p_n$ for all $n \geq 0$. This yields (i).

Clearly, (i) and (ii) imply $m_n \geq p_n$ for all $n \in \mathbb{N}_0$. If $M(t) = \sum_{n=0}^{N} m_n t^n$, $P(t) = \sum_{n=0}^{N'} p_n t^n$ are polynomials and (i), (ii) hold then $N = N'$ and applying (ii) for $n = N$, $N+1$ yields the equality:

$$\sum_{i=0}^{N} (-1)^{N-i} m_i = \sum_{i=0}^{N} (-1)^{N-i} p_i.$$

As a consequence of these observations we obtain:

Corollary 2.3 *Consider a sequence $X_0 \subset X_1 \subset \ldots \subset X_m$ of topological spaces such that all Betti numbers $\dim H_n(X_k, X_{k-1})$ are finite, and set*

$$m_n := \sum_{k=1}^{m} \dim H_n(X_k, X_{k-1}), \qquad p_n := \dim H_n(X_m, X_0).$$

Then

$$\sum_{i=0}^{n} (-1)^{n-i} m_i \geq \sum_{i=0}^{n} (-1)^{n-i} p_i$$

holds for all $n \geq 0$. In particular, $m_n \geq p_n$ holds for all $n \in \mathbb{N}_0$. Moreover, if $m_n = 0$ for $n > N$ then $p_n = 0$ for $n > N$ and

$$\sum_{i=0}^{N} (-1)^{N-i} m_i = \sum_{i=0}^{N} (-1)^{N-i} p_i.$$

Remark 2.4 In this section we have only used the exact sequence of a triple for H_*. We may also work with cohomology instead of homology.

2.2 The Morse inequalities

Let X be a topological space, $f : X \to \mathbb{R}$ be continuous, $K \subset X$ closed. In the differentiable setting K will be the set of critical points of f and $f(K)$ the set of critical values. For $c \in \mathbb{R}$ we set $f^c := \{x \in X : f(x) \le c\}$ and $K_c := \{x \in K : f(x) = c\}$. For $c \le d$ we shall also use the notation $f_c^d := \{x \in X : c \le f(x) \le d\}$. Let H_* be a homology theory. For $c \in \mathbb{R}$ and $S \subset f^{-1}(c)$ we define

$$C_n(f,S) := H_n(f^c, f^c \setminus S), \quad n \in \mathbb{Z}.$$

For $x \in f^{-1}(c)$ we set

$$C_n(f,x) := C_n(f,\{x\}) = H_n(f^c, f^c \setminus \{x\}), \quad n \in \mathbb{Z}.$$

Lemma 2.5 *If $S_1, S_2 \subset f^{-1}(c)$ are closed and disjoint then $C_n(f, S_1 \cup S_2) = C_n(f,S_1) \oplus C_n(f,S_2)$.*

Proof This follows from the relative Mayer-Vietoris sequence of the triad $(f^c; f^c \setminus S_1, f^c \setminus S_2)$. Observe that $(f^c \setminus S_1) \cap (f^c \setminus S_2) = f^c \setminus (S_1 \cup S_2)$ and $(f^c \setminus S_1) \cup (f^c \setminus S_2) = f^c$. Thus we have an exact sequence

$$\begin{array}{ccccccc} H_{n+1}(f^c, f^c) & \to & H_n(f^c, f^c \setminus (S_1 \cup S_2)) & \to & H_n(f^c, f^c \setminus S_1) \oplus H_n(f^c, f^c \setminus S_2) & \to & H_n(f^c, f^c) \\ \| & & & & & & \| \\ 0 & & & & & & 0 \end{array}$$

□

Now we fix two real numbers $a < b$ and require:

(A1) $f(K) \cap [a,b] = \{c_1, \ldots, c_k\}$ is finite and
$c_0 := a < c_1 < c_2 < \ldots < c_k < c_{k+1} := b$

(A2) $H_*(f^{c_{j+1}} \setminus K_{c_j+1}, f^{c_j}) = 0$ for $j = 0, \ldots, k$, $n \in \mathbb{Z}$.

If one wants to prove the existence of one or many critical points of f using Morse theory one can assume (A1) to hold because otherwise one has already infinitely many critical values. Condition (A2) is equivalent to:

(A3) The inclusion $i_j : f^{c_j} \hookrightarrow f^{c_{j+1}} \setminus K_{c_j+1}$ induces an isomorphism $i_{j*} : H_*(f^{c_j}) \to H_*(f^{c_{j+1}} \setminus K_{c_j+1})$ in homology for all $j = 0, \ldots, k$.

In classical Morse theory, f^{c_j} will in fact be a strong deformation retract of $f^{c_{j+1}} \setminus K_{c_j+1}$. Recall that $Z \subset Y$ is a strong deformation retract of Y if there exists a continuous map $h : [0,1] \times Y \to Y$ such that $h(0,x) = x$ and $h(1,x) \in Z$ for all $x \in Y$, and $h(t,x) = x$ for all $x \in Z$, $0 \le t \le 1$. h deforms Y into Z keeping Z fixed. This implies that $Z \hookrightarrow Y$ is a homotopy equivalence, hence it induces an isomorphism in homology.

Now for $n \ge 0$ we define

$$m_n := \sum_{j=1}^{k} \dim C_n(f, K_{c_j}) \in \overline{\mathbb{N}}_0 \quad \text{and} \quad p_n := \dim H_n(f^b, f^a) \in \overline{\mathbb{N}}_0.$$

We also set

$$M_f(t;a,b) := \sum_{n=0}^{\infty} m_n t^n \in \overline{\mathbb{N}}_0[[t]] \quad \text{and} \quad P_f(t;a,b) := \sum_{n=0}^{\infty} p_n t^n \in \overline{\mathbb{N}}_0[[t]]$$

If these are polynomials with finite coefficients then $M_f(t;a,b)$ is called Morse polynomial, $P_f(t;a,b)$ Poincaré polynomial.

Theorem 2.6 *If (A1), (A2) hold then there exists a series $Q(t) \in \overline{\mathbb{N}}_0[[t]]$ such that*

$$M_f(t;a,b) = P_f(t;a,b) + (1+t)Q(t).$$

If in addition all coefficients m_n, p_n of $M_f(t;a,b)$, $P_f(t;a,b)$ are finite then the following Morse inequalities hold:

$$\sum_{i=0}^{n} (-1)^{n-i} m_i \geq \sum_{i=0}^{n} (-1)^{n-i} p_i, \quad \text{all } n \geq 0.$$

Consequently, $m_n \geq p_n$ for all $n \geq 0$ and, if $M_f(t;a,b) \in \mathbb{N}_0[t]$ is a polynomial of degree N so is $P_f(t;a,b)$. In that case, the Morse equality

$$\sum_{i=0}^{N} (-1)^{N-i} m_i = \sum_{i=0}^{N} (-1)^{N-i} p_i$$

holds.

Proof The long exact sequence of the triple $(f^{c_j}, f^{c_j} \setminus K_{c_j}, f^{c_j-1})$ and (A2) yield:

$$C_n(f, K_{c_j}) = H_n(f^{c_j}, f^{c_j} \setminus K_{c_j}) = H_n(f^{c_j}, f^{c_j-1}) \quad \text{for } j = 1, \ldots, k+1,\ n \in \mathbb{Z}.$$

The theorem follows from Proposition 2.1 and Corollary 2.3 applied to $X_j = f^{c_j}, j = 0, \ldots, k+1$. □

What we presented so far is just elementary algebraic topology. Analysis enters when proving (A2) for certain (classes of) maps $f : X \to \mathbb{R}$. As mentioned above, (A1) is usually assumed to hold. We begin with a simple and classical situation.

Proposition 2.7 *Let X be a smooth closed (i.e. compact without boundary) riemannian manifold, $f \in C^2(X, \mathbb{R})$, and let $a < b$ be regular values of f. Let $K := \{x \in X : f'(x) = 0\}$ be the set of critical points of f and assume that $K \cap f_a^b$ is finite. Then (A1) and (A2) hold.*

In fact, (A1) holds trivially true. (A2) follows from a stronger result.

Proposition 2.8 *In the situation of Proposition 2.7 let $c, d \in (a,b)$, $c < d$, be such that $f(K) \cap (c,d) = \emptyset$. Then f^c is a strong deformation retract of $f^d \setminus K_d$.*

The proof of Proposition 2.8 uses the negative gradient flow associated to f. If $\langle \cdot, \cdot \rangle$ denotes the riemannian metric on X then $\nabla f(x) \in T_x X$ is defined by $\langle \nabla f(x), v \rangle = f'(x)v$ for all $v \in T_x X$. Since f is C^2, $\nabla f : X \to TX$ is a C^1-vector field and induces a flow φ_f on X defined by

$$\begin{cases} \dfrac{d}{dt}\varphi_f(t,x) = -\nabla f(\varphi_f(t,x)) \\ \varphi_f(0,x) = x. \end{cases}$$

We only need the induced semiflow on $Y := X \setminus K$ which we simply denote by $\varphi : [0,\infty) \times Y \to Y$. We also write $\varphi^t(x) := \varphi(t,x)$ and denote the ϵ-neighborhood of K by $U_\epsilon(K)$.

Lemma 2.9 *φ has the following properties:*

($\varphi1$) *For every $\epsilon > 0$ there exists $\delta > 0$ such that for $x \in Y \cap f_a^b$, $t > 0$ there holds:*
If $\varphi^s(x) \in f_a^b \setminus U_\epsilon(K)$ for all $0 \le s \le t$ then $f(x) - f(\varphi^t(x)) \ge \delta d(x, \varphi^t(x)) > 0$.

($\varphi2$) *If $f(\varphi^t(x)) \ge a$ for all $t \ge 0$ then the orbit $\{\varphi^t(x) : t \ge 0\}$ is relatively compact in X.*

In $(\varphi1)$, $d(x,y) = \inf_\gamma \int_0^1 \|\dot\gamma(t)\| dt$ denotes the distance in X; the infimum extends over all C^1-paths $\gamma : [0,1] \to X$ with $\gamma(0) = x$, $\gamma(1) = y$. According to $(\varphi1)$, for every $x \in Y \cap f_a^b$ the level $f(\varphi^t(x))$ strictly decreases as a function of t, as long as $\varphi^t(x) \in Y \cap f_a^b$. Moreover, if $\varphi^t(x)$ stays uniformly away from the set K, the difference quotient $\big(f(x) - f(\varphi^t(x))\big)/d(x, \varphi^t(x))$ is bounded away from 0.

Proof $(\varphi2)$ is clear because X is compact. In order to prove $(\varphi1)$ fix $\epsilon > 0$ and set

$$\delta := \inf\{\|\nabla f(x)\| : \ x \in f_a^b \setminus U_\epsilon(K)\} > 0.$$

Then we have for x, t as in $(\varphi1)$:

$$\begin{aligned} f(x) - f(\varphi^t(x)) &= -\int_0^t \frac{d}{ds} f(\varphi^s(x)) ds = \int_0^t \|\nabla f(\varphi^s(x))\|^2 ds \\ &\ge \delta \int_0^t \|\nabla f(\varphi^s(x))\| ds = \delta \int_0^t \|\frac{d}{ds}\varphi^s(x)\| ds \ge \delta d(\varphi^t(x), x). \end{aligned}$$

□

For the proof of Proposition 2.8 we only use the properties $(\varphi1)$, $(\varphi2)$. We need the map

$$\tau : f^d \setminus K_d \to [0,\infty], \quad \tau(x) := \inf\{t \ge 0 : f(\varphi^t(x)) \le c\}.$$

Here $\inf \emptyset = \infty$, so $\tau(x) = \infty$ if and only if $f(\varphi^t(x)) > c$ for all $t \ge 0$. Clearly $\tau(x) = 0$ if and only if $x \in f^c$.

Lemma 2.10 *If $\tau(x) = \infty$ then $\varphi^t(x) \to \overline{x} \in K_c$ as $t \to \infty$.*

Proof By $(\varphi2)$ the ω-limit set

$$\omega(x) := \bigcap_{t \ge 0} \operatorname{clos}\{\varphi^s(x) : s \ge t\} = \{y \in X : \text{ there exists } t_n \to \infty \text{ with } \varphi^{t_n}(x) \to y\}$$

is a compact connected nonempty subset of X. $(\varphi1)$ implies that $\omega(x) \subset K$, hence $\omega(x) \subset K_c$ because $f(K) \cap (c,d) = \emptyset$ by assumption. Finally, since K_c is finite we obtain $\omega(x) = \{\overline{x}\}$ for some $\overline{x} \in K_c$. □

Lemma 2.11 *τ is continuous.*

Proof Fix $x \in f^d \setminus K_d$ and $t < \tau(x)$. Then $f(\varphi^t(x)) > c$, hence there exists a neighborhood N of x with $f(\varphi^t(y)) > c$ for every $y \in N$. This implies $\tau(y) > t$ for $y \in N$, so τ is lower semi-continuous. Analogously one shows that τ is upper semi-continuous. □

Proof of Proposition 2.8 Clearly, the flow φ deforms any $x \in Z := f^d \setminus K_d$ within the time $\tau(x)$ to a point in f^c. In order to define the deformation $h : [0,1] \times Z \to Z$ required in Proposition 2.8 we just need to rescale the interval $[0,1]$ to $[0,\tau(x)]$. This is achieved, for instance, by the map

$$\chi(s,t) := \begin{cases} \dfrac{st}{1+t-st} & \text{for } 0 \le s \le 1,\ 0 \le t < \infty; \\ \dfrac{s}{1-s} & \text{for } 0 \le s < 1,\ t = \infty. \end{cases}$$

For fixed $t < \infty$ we have a reparametrization $\chi(\cdot,t) : [0,1] \to [0,t]$, while for $t = \infty$ we have a reparametrization $\chi(\cdot,\infty) : [0,1) \to [0,\infty)$. Clearly χ is continuous.

Now we define

$$h : [0,1] \times Z \to Z, \quad h(s,x) := \begin{cases} \lim_{t\to\infty} \varphi^t(x) & \text{if } s = 1,\ \tau(x) = \infty; \\ \varphi(\chi(s,\tau(x)),x) & \text{else.} \end{cases}$$

Then $h(0,x) = \varphi(0,x) = x$, $h(1,x) = \varphi(\tau(x),x) \in f^c$ if $\tau(x) < \infty$ and $h(1,x) = \lim_{t\to\infty} \varphi^t(x) \in K_c \subset f^c$ if $\tau(x) = \infty$. Moreover, if $x \in f^c$ then $\tau(x) = 0$ and $h(t,x) = \varphi(0,x) = x$ for all $t \in [0,1]$.

It remains to prove that h is continuous. Since φ, χ, τ are continuous we only need to consider the continuity at points $(1,x)$ with $\tau(x) = \infty$. We first show that $f \circ h$ is continuous at $(1,x)$. For $\varepsilon > 0$ there exists $t_\varepsilon \ge 0$ with $f(\varphi^{t_\varepsilon}(x)) < c + \varepsilon$. There also exists a neighborhood N_ε of x with $f(\varphi^{t_\varepsilon}(y)) < c + \varepsilon$ for all $y \in N_\varepsilon$, hence $c < f(\varphi^t(y)) < c + \varepsilon$ for all $y \in N_\varepsilon$, $t \in [t_\varepsilon, \tau(y)]$. This implies $c \le f(h(s,y)) < c + \varepsilon$ for all $y \in N_\varepsilon$, all $s \ge \chi(\cdot,\tau(y))^{-1}(t_\varepsilon) =: s_{\varepsilon,y}$. By the continuity of τ we may assume that $\tau(y) \ge t_\varepsilon + 1$ for all $y \in N_\varepsilon$. This implies that $s_\varepsilon := \sup_{y\in N_\varepsilon} s_{\varepsilon,y} < 1$, and $c \le f(h(s,y)) < c + \varepsilon$ for all $s \in [s_\varepsilon, 1]$, $y \in N_\varepsilon$. Thus $f \circ h$ is continuous.

Set $\bar{x} := h(1,x)$ and suppose there exist sequences $s_n \to 1$, $x_n \to x$ such that $h(s_n,x_n) \notin \overline{U}_{2\varepsilon}(\bar{x})$ for some $\varepsilon > 0$. Since $\varphi^k(x) \to \bar{x}$ as $k \to \infty$ and since φ is continuous, there exists a subsequence x_{n_k} with $\varphi^k(x_{n_k}) \to \bar{x}$. Setting $t_{n_k} := \chi(\cdot,\tau(x_{n_k}))^{-1}(k)$ we have $h(t_{n_k},x_{n_k}) \to \bar{x}$ and $t_{n_k} \to 1$. Thus, after passing to a subsequence, we may assume that $h(t_n,x_n) \to \bar{x}$ for some sequence $t_n \to 1$. We may also assume that $U_{3\varepsilon}(\bar{x}) \cap K = \{\bar{x}\}$. Then between the times s_n and t_n, the orbit $\varphi^t(x)$ passes through $\overline{U}_{2\varepsilon}(\bar{x}) \setminus \overline{U}_\varepsilon(\bar{x}) \subset f_a^b \setminus U_\varepsilon(K)$, so ($\varphi$1) yields $\delta > 0$ such that $|f(h(s_n,x_n)) - f(h(t_n,x_n))| \ge \delta\varepsilon/2$ for all $n \in \mathbb{N}$. On the other hand, since $f \circ h$ is continuous we obtain $f(h(s_n,x_n)) - f(h(t_n,x_n)) \to 0$, a contradiction. □

If, in the smooth case of Proposition 2.7 all critical points in $K \cap f_a^b$ are nondegenerate then $\dim C_n(f^c, f^c \setminus K_c)$ is precisely the number of critical points in K_c with Morse index n. This will be proved in section 3.2 below. Here x is a nondegenerate critical point of f if the hessian $f''(x) : T_xX \times T_xX \to \mathbb{R}$ is a nondegenerate quadratic form. The Morse index of x is the maximal dimension of a subspace of T_xX on which $f''(x)$ is negative definite. Combining Theorem 3.6 with Theorem 2.6 and Proposition 2.7 we obtain

Theorem 2.12 *Let X be a smooth closed riemannian manifold, $f \in C^2(X,\mathbb{R})$, and let $a < b$ be regular values of f. Let $K = \{x \in X : f'(x) = 0\}$ and suppose that all critical* *points in $K \cap f_a^b$ are nondegenerate. For $i \in \mathbb{N}_0$ let $m_i \in \mathbb{N}_0$ be the number of critical points in $K \cap f_a^b$ with Morse index i, and let $p_i := \dim H_i(f^b, f^a)$. Then there exists a polynomial $Q(t) \in \mathbb{N}_0[t]$ such that*

$$\sum_{i=0}^{\dim X} m_i t^i = \sum_{i=0}^{\dim X} p_i t^i + (1+t)Q(t).$$

Equivalently, the Morse inequalities

$$\sum_{i=0}^{n} (-1)^{n-i} m_i \geq \sum_{i=0}^{n} (-1)^{n-i} p_i, \quad n \geq 0$$

hold.

In the special case $a < \min f$, $b > \max f$ we have that $p_i = \dim H_i(X)$ is the i-th Betti number of X and m_i is the number of all critical points of f with Morse index i. Then we obtain the classical Morse inequalities.

In applications to boundary value problems for ordinary or partial differential equations, X is an infinite-dimensional Hilbert space or Hilbert manifold and $f : X \to \mathbb{R}$ is often only of class C^1. In that case one needs a replacement for the negative gradient flow. This is being achieved by considering pseudo-gradient vector fields. A vector $v \in X$ (or $v \in T_x X$) is said to be a pseudo-gradient vector for f at x if the following two conditions are satisfied:

(pg1) $\|v\| < 2\|f'(x)\|$

(pg2) $f'(x)v > \frac{1}{2}\|f'(x)\|^2$

A pseudo-gradient vector field for f on $Y \subset X$ is a locally Lipschitz continuous vector field $V : Y \to TX$ such that $V(x)$ is a pseudo-gradient vector for f at x. Using partitions of unity it is easy to construct a pseudo-gradient vector field for f on $X \setminus K$; see [17, Lemma I.3.1]. The conditions ($\varphi 1$), ($\varphi 2$) do not hold in general, however. They do hold if the following Palais-Smale condition is satisfied for $c \in \mathbb{R}$.

(PS)$_c$ Every Palais-Smale sequence x_n in X, i.e. a sequence such that $f'(x_n) \to 0$ and $f(x_n) \to c$, has a convergent subsequence.

Proposition 2.13 *Let X be a complete C^2-Hilbert manifold (without boundary), $f : X \to \mathbb{R}$ be C^1 and let K be the set of critical points of f. Let $a < b$ be given and suppose that the Palais-Smale condition (PS)$_c$ holds for every $c \in [a, b]$. Then there exists a flow φ on $Y := X \setminus K$ with the properties ($\varphi 1$), ($\varphi 2$) from Lemma 2.9.*

Proof One constructs a pseudo-gradient vector field V for f on Y and takes φ to be the semi-flow induced by $-V$. The conditions ($\varphi 1$), ($\varphi 2$) follow easily since $\|f'(x)\|$ is bounded away from 0 for $x \in f_a^b \setminus U_\varepsilon(K)$. We cheated a bit because φ is not defined on $Y \times [0, \infty)$, in general. This can be remedied however by considering the vector field $-\chi V$ with an appropriately chosen cut-off function $\chi : X \to [0, 1]$. Alternatively, one may rewrite the proof of Proposition 2.8 for not globally defined semiflows. □

As a consequence we obtain the Morse inequalities also in the infinite-dimensional setting. In the setting of Proposition 2.13 we call a critical point x of f topologically nondegenerate with Morse index $\mu \in \mathbb{N}_0$ if it is an isolated critical point and $\dim C_k(f,x) = \delta_{k,\mu}$.

Theorem 2.14 *Let X be a C^2-Hilbert manifold, $f : X \to \mathbb{R}$ be C^1 and let K be the set of critical points of f. Let $a < b$ be given and suppose that the Palais-Smale condition (PS)$_c$ holds for every $c \in [a,b]$. Suppose moreover that $f(K) \cap [a,b] = \{c_1 < \cdots < c_k\} \subset (a,b)$ is finite. Then the Morse inequalities from Theorem 2.6 hold. If in addition all critical points in f_a^b are topologically nondegenerate and have finite Morse index then the Morse inequalities from Theorem 2.12 hold.*

If f is C^2 then the concepts of nondegeneracy and Morse index are as in the finite-dimensional setting, and a nondegenerate critical point is topologically non-degenerate; see Definition 3.3 and Theorem 3.6.

Remark 2.15 For a number of applications it would be useful to develop Morse theory on Banach manifolds. This is rather delicate, however. For instance, if X is a Banach space and $f : X \to \mathbb{R}$ is C^2, then the existence of a nondegenerate critical point of f with finite Morse index implies the existence of an equivalent Hilbert space structure on X. Extensions of Morse theory to the Banach space setting are still a topic of research; see [16, 19, 41, 52, 53].

Remark 2.16 The Morse theory developed here can be refined to localize critical points.

a) Suppose the semi-flow associated to a pseudo-gradient vector field leaves a subset $Z \subset Y$ positively invariant, that is if $x \in Z$ then $\varphi^t(x) \in Z$ for all $t \geq 0$. In order to find critical points in Z one may replace X by Z and K by $K \cap Z$. The conditions ($\varphi 1$), ($\varphi 2$) continue to hold so that one obtains the Morse inequalities constrained to Z. It is important to note however that the coefficients in the Morse and the Poincaré polynomials depend on Z. This idea can be used for instance, to find positive (or negative) solutions of semilinear elliptic boundary value problems, that is, solutions lying in the cone of positive functions – provided one can construct a flow φ as above leaving this cone positively invariant.

b) One can also localize critical points outside of a positive invariant set Z. It can also be used to find sign-changing solutions of elliptic boundary value problems, that is solutions lying outside of the cone of positive or negative functions; see [18] for such an application. A first idea is to use the inverse flow $\varphi_-(t,x) = \varphi(-t,x)$ corresponding to $-f$. Observe that $X \setminus Z$ is positive invariant with respect to φ_- if Z is positive invariant with respect to φ. In the infinite-dimensional setting this does not work so easily because the theory developed so far yields nontrivial results only if the Morse indices are finite. If X is an infinite-dimensional manifold and $x \in X$ is a critical point of f of finite Morse index (and finite nullity) then it has infinite Morse index considered as a critical point of $-f$. Instead one can set up a relative Morse theory replacing X by the pair (X,Z) and K by $K \cap (X \setminus Z)$. The coefficients m_n, p_n are now defined as $m_n = \sum_{j=1}^k \dim C_n(f, K_{c_j} \cap (X \setminus Z))$ and $p_n = \dim H_n(f^b, f^a \cup Z)$.

3 Local theory

3.1 Morse lemma

The Morse lemma is the basic tool for the computation of the critical groups of a nondegenerate critical point. For degenerate isolated critical points, the splitting theorem gives the appropriate representation. Since the theory is local, we consider a Hilbert space E. We require in this section

(M) U is an open neighborhood of 0 in the Hilbert space E, 0 is the only critical point of $f \in \mathcal{C}^2(U, \mathbb{R})$, $L = f''(0)$ is invertible or 0 is an isolated point of $\sigma(L)$.

Here, using the scalar product of E, the Hessian $f''(0) : E \times E \to \mathbb{R}$ of f at 0 corresponds to a linear map $L : E \to E$. By abuse of notation we write $f''(0)$ for both maps.

Theorem 3.1 (Morse Lemma.) *Let L be invertible. Then there exists an open ball B_δ and local diffeomorphism $g : B_\delta \to E$ such that $g(0) = 0$ and, on B_δ,*

$$f \circ g(u) = \frac{1}{2}(Lu, u).$$

If 0 is an isolated point of $\sigma(L)$, E is the orthogonal sum of $R(L)$, the range of L, and $N(L)$, the kernel of L. Let $u = v + w$ be the corresponding decomposition of $u \in E$.

Theorem 3.2 (Splitting Theorem.) *Let* 0 *be an isolated point of* $\sigma(L)$*. Then there exists an open ball B_δ, a local homeomorphism $g : B_\delta \to E$ such that $g(0) = 0$ and a $\mathcal{C}^1$ mapping $h : B_\delta \cap N(L) \to R(L)$ such that, on B_δ,*

$$f \circ g(v + w) = \frac{1}{2}(Lv, v) + f(h(w) + w) = \frac{1}{2}(Lv, v) + \hat{f}(w).$$

For the proofs of Theorems 3.1 and 3.2, we refer to [17] and to [39].

3.2 Critical groups

In this section, we denote by U an open subset of the Hilbert space E. Let us recall a definition of section 2.2 in this setting.

Definition 3.3 Let x be an isolated critical point of $f \in \mathcal{C}^1(U, \mathbb{R})$. The critical groups of x are defined by

$$C_n(f, x) = H_n(f^c, f^c \backslash \{x\}), n \in \mathbb{Z},$$

where $c = f(x)$.

Remark 3.4 By excision, the critical groups depend only on the restriction of f to an arbitrary neighborhood of x in U. The critical groups of a critical point of $f \in \mathcal{C}^2(X, \mathbb{R})$, where X is a $\mathcal{C}^2$-Hilbert manifold are defined in the same way.

Definition 3.5 Let x be a critical point of $f \in \mathcal{C}^2(U, \mathbb{R})$. The Morse index of x is the supremum of the dimensions of the subspaces of E on which $f''(x)$ is negative definite. The critical point x is nondegenerate if $f''(x)$ is invertible. The nullity of x is the dimension of the kernel of $f''(x)$.

Theorem 3.6 *Let x be a nondegenerate critical point of $f \in \mathcal{C}^2(U, \mathbb{R})$ with finite Morse index m. Then*

$$\dim C_n(f, x) = \delta_n^m.$$

Proof We can assume that $x = 0$ and $f(x) = 0$. By Theorem 3.1, there exists a local diffeomorphism $g : B_\delta \to E$ such that $g(0) = 0$ and, on B_δ,

$$f \circ g(u) = \frac{1}{2}(Lu, u) := \psi(u).$$

We have, for $m \geq 1$,

$$\begin{aligned} C_n(f, 0) &= H_n(f^0 \cap g(B_\delta), f^0 \cap g(B_\delta)\backslash\{0\}) \\ &\cong H_n(\psi^0 \cap B_\delta, \psi^0 \cap B_\delta\backslash\{0\}) \\ &\cong H_n(B^m, S^{m-1}) \end{aligned}$$

and

$$\dim H_n(B^m, S^{m-1}) = \delta_n^m.$$

We have also, for $m = 0$,

$$C_n(f, 0) \cong H_n(\{0\}, \emptyset)$$

and

$$\dim H_n(\{0\}, \emptyset\} = \delta_n^0.$$

□

Theorem 3.7 (Shifting Theorem.) *Let U be an open neighborhood of 0 in the Hilbert space E. Assume that 0 is the only critical point of $f \in \mathcal{C}^2(U, \mathbb{R})$, that the Palais-Smale condition is satisfied over a closed ball in U and that the Morse index m of 0 is finite. Then*

$$C_n(f, 0) \cong C_{n-m}(\hat{f}, 0)$$

where $\hat{f}$ is defined in Theorem 3.2.

For the proof we refer to Theorem 8.4 in [39].

Corollary 3.8 *Assume moreover that the nullity k of 0 is finite. Then*

a) $C_n(f, 0) \cong 0$ *for* $n \leq m - 1$ *and* $n \geq m + k + 1$*;*

b) 0 *is a local minimum of* $\hat{f}$ *iff* $\dim C_m(f, 0) = \delta_n^m$;

c) 0 *is a local maximum of* $\hat{f}$ *iff* $\dim C_m(f, 0) = \delta_n^{m+k}$.

For the proof see Corollary 8.4 in [39].

4 Applications

4.1 A three critical points theorem

Let E be a Hilbert space and $f \in \mathcal{C}^1(E,\mathbb{R})$. As in section 2.2, f satisfies the $(\mathrm{PS})_c$ condition, $c \in \mathbb{R}$, if every sequence (x_n) in X with $f'(x_n) \to 0$ and $f(x_n) \to c$ has a convergent subsequence. The function f satisfies the (PS) condition if, for every $c \in \mathbb{R}$, the $(\mathrm{PS})_c$ condition is satisfied.

Theorem 4.1 *Let $f \in \mathcal{C}^2(E,\mathbb{R})$ be bounded from below. Assume that f satisfies the (PS)-condition and that x_1 is a nondegenerate nonminimum critical point of f with finite Morse index. Then f has at least three critical points.*

Proof Since f is bounded from below and satisfies (PS), there exists a minimizer x_0 of f on E. Let us assume that x_0 and x_1 are the only critical points of f. Corollary 3.8 and Theorem 3.6 yield

$$\dim C_n(f,x_0) = \delta_n^0, \quad \dim C_n(f,x_1) = \delta_n^m,$$

where m is the Morse index of x_1. Theorem 2.14, applied to $a = f(x_0) - 1$ and $b = f(x_1) + 1$ yields a polynomial $Q \in \mathbb{N}_0[t]$

$$1 + t^m = 1 + (1+t)Q(t),$$

a contradiction. We have used the fact that f^b is a deformation retract of E so that, for $n \in \mathbb{Z}$,

$$H_n(f^b, f^a) = H_n(E,\emptyset) = H_n(E).$$

□

Let us consider the Dirichlet problem

$$\begin{cases} -\Delta u = g(u), & \text{in } \Omega, \\ \\ u = 0, & \text{on } \partial\Omega, \end{cases} \tag{4.1}$$

where $\Omega \subset \mathbb{R}^N$ is a smooth bounded domain. We assume that

(g_1) $g \in \mathcal{C}^1(\mathbb{R})$ and

$$|g'(t)| \le c_1(1 + |t|^{p-2}), 2 \le p < 2^*,$$

with $2^* = +\infty$ if $N = 2$, $2^* = 2N/(N-2)$ if $N \ge 3$.

Let $\lambda_1 < \lambda_2 \le \dots$ be the eigenvalues of $-\Delta$ with the Dirichlet boundary condition on Ω. We assume also that

(g_2) $G(t) \le c_2(1+t^2)$, $c_2 < \lambda_1/2$, where $G(t) = \int_0^t g(s)ds$,

(g_3) $g(0) = 0$, $\lambda_j < g'(0) < \lambda_{j+1}$, $j \ge 1$.

It follows from (g_1) that the solutions of (4.1) are the critical points of the $\mathcal{C}^2$-functional

$$f(u) = \frac{1}{2}\int_\Omega |\nabla u|^2 dx - \int_\Omega G(u)dx,$$

defined on $H_0^1(\Omega)$. The space $H_0^1(\Omega)$ is the subspace of functions of

$$H^1(\Omega) = \left\{u \in L^2(\Omega) : \frac{\partial u}{\partial x_k} \in L^2(\Omega), 1 \leq k \leq N\right\}$$

satisfying the Dirichlet boundary condition in the sense of traces.

Theorem 4.2 *Under assumptions* (g_{1-2-3})*, problem* (4.1) *has at least 3 solutions.*

Proof It follows from (g_2) that

$$f(u) \to +\infty \quad \text{as } \|u\| = \|\nabla u\|_{L^2} \to \infty, \tag{4.2}$$

on $H_0^1(\Omega)$. By Rellich's theorem, the imbedding $H_0^1(\Omega) \subset L^p(\Omega)$ is compact for $p < 2^*$. Using (g_1), it is then not difficult to verify the (PS)-condition. In particular, by (4.2), there exists a minimizer x_0. By assumption (g_3), 0 is a nondegenerate critical point of f with Morse index $j \geq 1$. In particular $x_0 \neq 0$. By Theorem 4.1, f has at least 3 critical points. □

4.2 Asymptotically linear problems

We consider again problem (4.1) under the assumptions (g_1) and

(g_4) $g(0) = 0$, $\lambda_j < g'(0) < \lambda_{j+1}$, $j \geq 0$, where $\lambda_0 = -\infty$,

(g_5) $g(t) = \lambda t + o(t)$, $|t| \to +\infty$, $\lambda_k < \lambda < \lambda_{k+1}$, $0 \leq k \neq j$.

Theorem 4.3 *Under assumptions* (g_{1-4-5})*, problem* (4.1) *has at least 2 solutions.*

Proof Let us assume that 0 is the only solution and, hence, the only critical point of f. As a consequence of (g_1) and (g_5) the (PS)-condition holds (cf. [17]). Moreover, for $b > 0$ and $a < 0$,

$$\dim H_n(f^b, f^a) = \delta_n^k.$$

By Theorem 3.6,

$$\dim C_n(f, 0) = \delta_n^j.$$

It follows from Theorem 2.14 that $j = k$, contrary to our assumptions. □

More solutions exist if

(g_6) g is odd.

Theorem 4.4 *Under assumptions* $(g_{1-4-5-6})$*, problem* (4.1) *has at least* $m = |j - k|$ *pairs* $\pm u_1, \ldots, \pm u_m$ *of solutions in addition to the trivial solution 0.*

Proof We sketch the proof in the case $j > k$. Let $G = \{\pm 1\}$ denote the group of 2 elements acting on $E := H^1(\Omega)$ via the antipodal map $u \mapsto -u$. Observe that f is even by (g_6), hence sublevel sets f^a are invariant under G. Moreover, if u is a critical point of f then so is $-u$ because f' is odd: $f'(u) = -f'(-u)$. We consider the Borel cohomology $H_G^*(f^a)$ of the sublevel set f^a for $a < 0$. Since 0 is the only fixed point of the action of G on E and $0 \notin f^a$ for $a < 0$ one has $H_G^*(f^a) \cong H^*(f^a/G; \mathbb{F}_2)$ where $f^a/G = f^a/u \sim -u$ is the quotient space and $\mathbb{F}_2 = \{0, 1\}$ is the field of 2 elements. The cohomology group $H_G^*(f^a)$ is a module over the ring $H_G^*(E \setminus \{0\}) \cong H^*(\mathbb{R}P^\infty) \cong \mathbb{F}_2[w]$.

As a consequence of (g_5), equation (4.1) is asymptotically linear and f is asymptotically quadratic. More precisely, let e_i, $i \in \mathbb{N}$, be an othogonal basis of $H^1(\Omega)$ consisting of eigenfunctions of $-\Delta$ corresponding to the eigenvalues λ_i. Then

$$f(u) \to -\infty \qquad \text{for } u \in E_\infty := \operatorname{span}\{e_1, \ldots, e_k\},\ \|u\| \to \infty$$

and

$$f(u) \to +\infty \qquad \text{for } u \in E_\infty^\perp = \operatorname{span}\{e_i : i \geq k+1\},\ \|u\| \to \infty.$$

by (g_5). Thus for $b \ll 0$ and $R \gg 0$ we have inclusions

$$S_R E_\infty := \{u \in E_\infty : \|u\| = R\} \hookrightarrow f^b \hookrightarrow E \setminus E_\infty^\perp \simeq S_R E_\infty.$$

These are in fact homotopy equivalences which implies that

$$H_G^*(f^b) \cong H_G^*(S_R E_\infty) \cong H^*(\mathbb{R}P^k) \cong \mathbb{F}_2[w]/w^{k+1}.$$

On the other hand, $f(u) < 0$ for $u \in E_0 := \operatorname{span}\{e_1, \ldots, e_j\}$ with $\|u\| \leq r$, $r > 0$ small, by (g_4). Thus for $a < 0$ close to 0 we have the inclusion $S_r E_0 \hookrightarrow f^a \hookrightarrow E \setminus \{0\}$. On the cohomology level this yields homomorphisms

$$\mathbb{F}_2[w] \to H_G^*(f^a) \to \mathbb{F}_2[w]/w^j$$

whose composition is surjective. Consequently, $H_G^i(f^a) \neq 0 = H_G^i(f^b)$ for $i = k+1, \ldots, j$, and therefore $H_G^i(f^a, f^b) \neq 0$ for $i = k+1, \ldots, j$. If all critical points are nondegenerate we immediately obtain critical points $u_1, \ldots, u_m$, $m = j - k$, with Morse indices $m(u_i) = i + k$. In the degenerate case the argument is more complicated and one has to use the structure of $H_G^*(f^a, f^b)$ as a module over $\mathbb{F}_2[w]$. □

4.3 Bifurcation theory

In this section we consider nontrivial solutions of the parameter dependent problem

$$\nabla f_\lambda(u) = 0 \tag{4.2}$$

under the assumption that

$$\nabla f_\lambda(0) = 0$$

holds for all parameters λ.

Definition 4.5 Let U be an open neighborhood of 0 in the Hilbert space E, let Λ be an open interval and let $f \in \mathcal{C}^1(\Lambda \times U, \mathbb{R})$ be such that $\nabla f_\lambda(0) = 0$ for every $\lambda \in \Lambda$ when $f_\lambda = f(\lambda, \cdot)$. A point $(\lambda_0, 0) \in \Lambda \times U$ is a bifurcation point for equation (4.2) if every neighborhood of $(\lambda_0, 0)$ in $\Lambda \times U$ contains at least one solution (λ, u) of (4.2) such that $u \neq 0$.

Theorem 4.6 *Let $f \in \mathcal{C}^1(\Lambda \times U, \mathbb{R})$ be as in Definition 4.5. Assume that the following conditions are satisfied :*

i) 0 *is an isolated critical point of f_a and f_b for some reals $a < b$ in Λ,*

ii) *for every $a < \lambda < b$, f_λ satisfies the Palais-Smale condition over a closed ball $B[0, R] \subset U$, that is, every (PS)-sequence in $B[0, R]$ has a convergent subsequence.*

iii) *there exists $n \in \mathbb{N}$ such that $C_n(f_a, 0) \ncong C_n(f_b, 0)$.*

Then there exists a bifurcation point $(\lambda_0, 0) \in [a, b] \times \{0\}$ for equation (4.2).

Theorem 4.6 is due to Mawhin and Willem (see [39]).

Let us consider a variant of problem (4.1) :

$$\begin{aligned} -\Delta u &= \lambda g(u), && \text{in } \Omega, \\ u &= 0, && \text{on } \partial\Omega. \end{aligned} \tag{4.3}$$

We assume that g satisfies (g_1) and

(g_6) $g(0) = 0, g'(0) = 1.$

The corresponding functional is defined on $H_0^1(\Omega)$ by

$$f(\lambda, u) = \frac{1}{2} \int_\Omega |\nabla u|^2 dx - \lambda \int_\Omega G(u) dx.$$

Theorem 4.7 *Under assumptions (g_1) and (g_6), $(\lambda, 0)$ is a bifurcation point for problem* (4.3) *if and only if λ is an eigenvalue of $-\Delta$ with the Dirichlet boundary condition.*

For the proof we refer to [56].

4.4 Closed geodesics

Let (M, g) be a compact riemannian manifold without boundary. A geodesic is a curve $c : I \to M$, $I \subset \mathbb{R}$ an interval, satisfying the differential equation $\nabla_{\dot{c}}\, \dot{c} = 0$, that is the tangent field $\dot{c}$ is tangent along c. A periodic geodesic $c : \mathbb{R} \to M$ is said to be closed.

Closed geodesics are critical points of the energy functional

$$f : H^1(S^1, M) \to \mathbb{R}, \quad f(c) = \int_0^1 \|\dot{c}\|^2 dt$$

where $S^1 = \mathbb{R}/\mathbb{Z}$. They are also critical points of the lenght functional $L(c) = \int_0^1 \|\dot{c}\| dt$. However, this functional is invariant under reparametrizations which implies that given a nonconstant closed geodesic c, all reparametrizations $c \circ \sigma$, $\sigma : \mathbb{R} \to \mathbb{R}$ a strictly increasing C^1-map such that $\sigma(t + 1) = \sigma(t) + 1$, are also critical points of L at the same level

$L(c) = L(c \circ \sigma)$. It follows that the Palais-Smale condition cannot hold for L. On the contrary, a critical point of f is automatically parametrized proportional to arc-length. The functional f does satisfy the Palais-Smale condition.

Although being a classical problem of Morse theory, the problem of the existence of closed geodesics has new features not present in our discussion so far. First of all observe that although f is not invariant under reparametrizations, it is invariant under time shifts. Given a 1-periodic H^1-function $c : \mathbb{R} \to M$ and given $\tau \in \mathbb{R}$, we define $c_\tau(t) := c(t+\tau)$. This defines an action of $S^1 = \mathbb{R}/\mathbb{Z}$ on $H^1(S^1, M)$ and clearly f is invariant under this action: $f(c_\tau) = f(c)$. As a consequence, if c is a critical point of f, so is c_τ and therefore f does not have isolated critical points. Instead each critical point corresponding to a nonconstant closed geodesic yields a manifold c_τ, $\tau \in \mathbb{R}/\mathbb{Z}$, of critical points. The local theory developed in Section 3 can be extended to cover manifolds of critical points. This goes back to the work of Bott [11]; see also [39, Chapter 10] for a presentation with applications to differential equations.

The S^1-invariance of f can be used very successfully to obtain "many" critical points. We have seen already in section 4.2 that symmetry implies the existence of multiple critical points. For the closed geodesic problem yet another difficulty appears. If c is a critical point of f then $c^m(t) := c(mt)$ is also a critical point of f, any $m \in \mathbb{N}$. Geometrically, c and c^m, $m > 1$, describe the same closed geodesic and should not be counted separately. A closed geodesic is said to be prime if it is not the m-th iterate, $m > 1$, of a geodesic. In order to understand the contribution of c and its iterates c^m Bott [12] developed an iteration theory for closed geodesics.

A discussion of the Morse theory for the closed geodesics problem goes far beyond the scope of this survey. We state two important results where Morse theory played a decisive role.

Theorem 4.8 *If M is simply connected and if the Betti numbers of the free loop space $H^1(S^1, M) \simeq C^0(S^1, M)$ with respect to some field of coefficients form an unbounded sequence then there exist infinitely many, geometrically different prime closed geodesics on M. The hypothesis on the free loop space is satisfied if the cohomology algebra $H^*(M; \mathbb{Q})$ is not generated (as a $\mathbb{Q}$-algebra with unit) by a single element.*

Since f is bounded below and satisfies the Palais-Smale condition it has infinitely many critical points provided there are infinitely many non-zero Betti numbers. The condition that the Betti numbers form an unbounded sequence can be used to show that the infinitely many critical points are not just the multiples of only finitely prime closed geodesics. The first statement of Theorem 4.8 is due to Gromoll and Meyer [29], the second purely topological result to Vigué-Poirrier and Sullivan [54]. It applies to many manifolds, in particular to products $M = M_1 \times M_2$ of compact simply connected manifolds. It does not apply to spheres, for instance. The reader may consult the paper [20] for further results, references and a discussion of topological features of the problem.

Theorem 4.9 *On the 2-sphere (S^2, g), any metric g, there are always infinitely many, geometrically different prime closed geodesics.*

This theorem is due to Franks [27] and Bangert [6], covering separate cases. The proof of Bangert is via Morse theory, the proof of Franks uses dynamical systems methods. A Morse theory proof of this part can be found in [33].

Similar problems as in the closed geodesic problem appear for the search of periodic solutions of autonomous Hamiltonian systems. In particular one has an S^1-symmetry, critical points are not isolated, and one has the problem that iterates of critical points are also critical points. We refer the reader to the book [38] for a presentation of the Morse index theory and index formulas for iterated curves for periodic solutions of Hamiltonian systems. In additon to the above mentioned difficulties the corresponding functional is strongly indefinite, Morse indices of critical points are infinite. The next section is devoted to Morse theory for strongly indefinite functionals.

5 Strongly indefinite Morse theory

5.1 Cohomology and relative Morse index

Let E be a real Hilbert space and consider the functional

$$\Phi(x) = \frac{1}{2}\|x^+\|^2 - \frac{1}{2}\|x^-\|^2,$$

where $x^\pm \in E^\pm$ and $E = E^+ \oplus E^-$ is an orthogonal decomposition. Then $\nabla\Phi(x) = x^+ - x^-$, the critical set $K = \{0\}$, the Morse index $M^-(\Phi''(0)) = \dim E^-$ and $C_q(\Phi, 0) = \mathbb{F}$ if $q = \dim E^-$ and 0 otherwise. Hence, if $\dim E^- = +\infty$, then $C_*(\Phi, 0) = 0$, and the critical point 0 will not be seen by the Morse theory developed in sections 2 and 3. In what follows we will be concerned with functionals which are of the form $\Phi(x) = \frac{1}{2}\|x^+\|^2 - \frac{1}{2}\|x^-\|^2 - \psi(x)$, where $E = E^+ \oplus E^0 \oplus E^-$, $\dim E^\pm = +\infty$ and $\nabla\psi$ is compact. If $\Phi \in C^2(E, \mathbb{R})$, then it is easy to see that $M^-(\pm\Phi''(x)) = +\infty$ for any $x \in K$, and therefore the Morse theory developed so far becomes useless. It is also easy to see that functionals of the type described here are strongly indefinite in the sense that they are unbounded below and above on any subspace of finite codimension.

In order to establish a Morse theory which is useful for strongly indefinite functionals we first introduce a suitable cohomology theory [36]. Let $(E_n)_{n=1}^\infty$ be a sequence of closed subspaces of E such that $E_n \subset E_{n+1}$ for all n and $\bigcup_{n=1}^\infty E_n$ is dense in E. We shall call (E_n) a *filtration* of E. For a closed set $X \subset E$ we use the shorthand notation

$$X_n := X \cap E_n.$$

To each E_n we assign a nonnegative integer d_n and we write

$$\mathcal{E} := (E_n, d_n)_{n=1}^\infty.$$

Next, if $(\mathcal{G}_n)_{n=1}^\infty$ is a sequence of abelian groups, then we define the asymptotic group

$$[(\mathcal{G}_n)_{n=1}^\infty] := \prod_{n=1}^\infty \mathcal{G}_n \Big/ \bigoplus_{n=1}^\infty \mathcal{G}_n,$$

i.e., $[(\mathcal{G}_n)_{n=1}^\infty] = \prod_{n=1}^\infty \mathcal{G}_n/\sim$, where $(g_n)_{n=1}^\infty \sim (g'_n)_{n=1}^\infty$ if and only if $g_n = g'_n$ for all n large enough. In what follows we shall use the shorter notation

$$[\mathcal{G}_n] = [(\mathcal{G}_n)_{n=1}^\infty] \text{ and } [\mathcal{G}_n] = \mathcal{G} \text{ if } \mathcal{G}_n = \mathcal{G} \text{ for almost all } n.$$

Let (X, A) be a pair of *closed* subsets of E such that $A \subset X$ and denote the Čech cohomology with coefficients in a field $\mathbb{F}$ by H^*. If $\mathcal{E}$ is as above, then for each $q \in \mathbb{Z}$ we define

$$H^q_{\mathcal{E}}(X, A) := [H^{q+d_n}(X_n, A_n)].$$

Observe that $H^*_{\mathcal{E}} = H^*$ if $E_n = E$ and $d_n = 0$ for almost all n, and in general $H^q_{\mathcal{E}}$ need not be 0 for all $q < 0$. As morphisms in the category of closed pairs we take continuous mappings $f : (X, A) \to (Y, B)$ which preserve the filtration, i.e., $f(X_n) \subset E_n$ for almost all n. Such f will be called *admissible*. If $f_n := f|_{X_n}$, then f induces a homomorphism $f^* : H^q_{\mathcal{E}}(Y, B) \to H^q_{\mathcal{E}}(X, A)$ given by $f^* = [f^*_n]$, where $f^*_n : H^{*+d_n}(Y_n, B_n) \to H^{*+d_n}(X_n, A_n)$. Similarly, the coboundary operator $\delta^* : H^*_{\mathcal{E}}(A) \to H^{*+1}_{\mathcal{E}}(X, A)$ is given by $\delta^* = [\delta^*_n]$. We also define admissible homotopies $G : [0, 1] \times (X, A) \to (Y, B)$ by requiring that $G([0, 1] \times X_n) \subset E_n$ for almost all n. It is easy to see from the definitions and the properties of H^* that $H^*_{\mathcal{E}}$ satisfies the usual Eilenberg-Steenrod axioms for cohomology except the dimension axiom. Moreover, since H^* satisfies the strong excision property, so does $H^*_{\mathcal{E}}$. More precisely, this means that if A, B are closed subsets of E, then there is an isomorphism

$$H^*_{\mathcal{E}}(A, A \cap B) \cong H^*_{\mathcal{E}}(A \cup B, B).$$

The need for strong excision was in fact our reason for using the Čech cohomology.

Let $\tilde{L} : E \to E$ be a linear selfadjoint Fredholm operator such that $\tilde{L}(E_n) \subset E_n$ for almost all n. Then $E = E^+(\tilde{L}) \oplus N(\tilde{L}) \oplus E^-(\tilde{L})$, where $N(\tilde{L})$ is the nullspace (of finite dimension) and $E^\pm(\tilde{L})$ are the positive and the negative space of $\tilde{L}$. Since

$$\langle \tilde{L}x^+, y^+\rangle - \langle \tilde{L}x^-, y^-\rangle + \langle x^0, y^0\rangle, \quad x^\pm, y^\pm \in E^\pm(\tilde{L}),\ x^0, y^0 \in N(\tilde{L})$$

is an equivalent inner product, we may (and will) assume without loss of generality that $\tilde{L}x = x^+ - x^-$. We also take

$$d_n := M^-(\tilde{L}|_{E_n}) + d_0 \equiv \dim E^-(\tilde{L})_n + d_0,$$

where $E^-(\tilde{L})_n = E^-(\tilde{L}) \cap E_n$ and d_0 is a convenient normalization constant to be chosen. Let $B : E \to E$ be linear, compact and selfadjoint, $L := \tilde{L} - B$, denote the orthogonal projector on E_n by P_n and the orthogonal projector from the range $R(L)$ of L on $R(L)_n$ by Q_n, and let

$$M^-_{\mathcal{E}}(L) := \lim_{n\to\infty} \left(M^-(Q_n L|_{R(L)_n}) - d_n\right). \tag{5.1}$$

Then $M^-_{\mathcal{E}}(L)$ is a well-defined (and not necessarily positive) integer [36, Proposition 5.2]. Note that $M^-(Q_n L|_{R(L)_n})$ is the Morse index of the quadratic form $\langle Lx, x\rangle$ restricted to $R(L)_n$. It is not too difficult to show that if $N(L) \subset E_{n_0}$ for some n_0, then

$$M^-_{\mathcal{E}}(L) = \lim_{n\to\infty} \left(M^-(P_n L|_{E_n}) - d_n\right) = \lim_{n\to\infty} \left(M^-(P_n L|_{E_n}) - M^-(\tilde{L}|_{E_n})\right) - d_0 \tag{5.2}$$

(see [36, Remark 5.1]). Hence $M^-_{\mathcal{E}}(L)$ is a relative Morse index in the sense that it measures the difference between the Morse indices of the operators L and $\tilde{L}$ restricted to E_n (n large). If $N(L) \not\subset E_n$ for any n, then $M^-_{\mathcal{E}}(L)$ may not be equal to the limit in (5.2). A justification why (5.1) and not (5.2) is used as the definition of $M^-_{\mathcal{E}}(L)$ may be found in [36].

Let

$$D := \{x \in E^-(L) : \|x\| \le 1\}, \quad S := \{x \in E^-(L) : \|x\| = 1\}.$$

The connection between $H^*_{\mathcal{E}}$ and $M^-_{\mathcal{E}}(L)$ is expressed in the following

Proposition 5.1 *Suppose L satisfies the conditions formulated above, $N(L) \subset E_n$ and $L(E_n) \subset E_n$ for almost all n. Then $H^q_{\mathcal{E}}(D,S) = \mathbb{F}$ if $q = M^-_{\mathcal{E}}(L)$ and $H^q_{\mathcal{E}}(D,S) = 0$ otherwise.*

Proof Since $P_nL = LP_n$, the negative space of $L|_{E_n}$ is $E^-(L)_n$. Hence $D_n = \{x \in E^-(L)_n : \|x\| \le 1\}$, so $\dim D_n = M^-_{\mathcal{E}}(L) + d_n$ for almost all n. It follows that $H^{q+d_n}(D_n, S_n) = \mathbb{F}$ if $q = M^-_{\mathcal{E}}(L)$ and 0 otherwise. □

Note in particular that if $\dim E^-(L) = +\infty$, then $H^*(D,S)$ is trivial while $H^*_{\mathcal{E}}(D,S)$ is not.

Remark 5.2 Different though related to $H^*_{\mathcal{E}}$ cohomology theories may be found in [1, 49]. Both theories are constructed by a suitable modification of an infinite-dimensional cohomology of Gęba and Granas [28]. In [3] the reader may find an infinite-dimensional *homology* theory constructed in a very different way, with the aid of a Morse-Witten complex.

5.2 Critical groups and Morse inequalities

The critical groups of an isolated critical point x were defined in section 2.2 by setting $C_q(\Phi, x) = H_q(\Phi^c, \Phi^c \setminus \{x\})$. It may seem natural to define $C^q_{\mathcal{E}}$ here by simply replacing H_q with $H^q_{\mathcal{E}}$. However, this is not possible because $H^q_{\mathcal{E}}(X, A)$ has been defined for closed sets only, and moreover, if x is an isolated critical point of Φ, it is not clear in general how the critical set for $\Phi|_{E_n}$ looks like in small neighborhoods of x, not even if $x \in E_{n_0}$ for some n_0.

We shall construct a Morse theory for functionals which are of the form

$$\Phi(x) = \frac{1}{2}\langle \tilde{L}x, x\rangle - \psi(x) \equiv \frac{1}{2}\|x^+\|^2 - \frac{1}{2}\|x^-\|^2 - \psi(x), \tag{5.3}$$

where $x = x^+ + x^0 + x^- \in E = E^+ \oplus E^0 \oplus E^-$, $\dim E^0 < \infty$, and $\nabla\psi$ is a compact operator. Although it is possible to allow a larger class of Φ (see [36]), there can be no useful Morse theory which includes all smooth Φ such that $M^-(\pm\Phi''(x)) = +\infty$ whenever $x \in K$. This is a consequence of a result by Abbondandolo and Majer [4].

In order to avoid the problem with the critical set of $\Phi|_{E_n}$ mentioned above we shall use an approach which goes back to Gromoll and Meyer (see e.g. [17]). Its main feature is that to each isolated critical point x one assigns a pair (W, W^-) of closed sets such that x is in the interior of W, W^- is the exit set (from W) for the flow of $-\nabla\Phi$, and then one defines the critical groups of x by setting $C_*(\Phi, x) = H_*(W, W^-)$ (in homology theory) and $C^*(\Phi, x) = H^*(W, W^-)$ (in cohomology theory). Since $\nabla\Phi$ is not admissible for $H^*_{\mathcal{E}}$, we shall need a class of mappings which are related to $\nabla\Phi$ and admissible. The results we summarize below may be found, in a more general form, in [36].

Let E be a real Hilbert space, (E_n) a filtration and suppose $\Phi \in C^1(E, \mathbb{R})$. A sequence (x_j) is said to be a $(PS)^*$*-sequence* (with respect to (E_n)) if $\Phi(x_j)$ is bounded, $x_j \in E_{n_j}$ for some n_j, $n_j \to \infty$ and $P_{n_j}\nabla\Phi(x_j) \to 0$ as $j \to \infty$. If each $(PS)^*$-sequence has a convergent subsequence, then Φ is said to satisfy the $(PS)^*$*-condition*. Using the density of $\bigcup_{n=1}^{\infty} E_n$ it is easy to show that convergent subsequences of (x_j) tend to critical points and $(PS)^*$ implies the Palais-Smale condition.

Lemma 5.3 *If Φ is of the form (5.3), then each bounded $(PS)^*$-sequence has a convergent subsequence.*

Proof Let (x_j) be a bounded $(PS)^*$-sequence. Then $x_j \rightharpoonup \bar{x}$ and $\nabla\psi(x_j) \to w$ for some $w \in E$ after passing to a subsequence. If $x_j = y_j + x_j^0$, $y_j \in R(\tilde{L})$, $x_j^0 \in E^0 \equiv N(\tilde{L})$, then $\tilde{L}y_j - P_{n_j}\nabla\psi(x_j) \to 0$, and since E^0 is finite-dimensional and $\tilde{L}|_{R(\tilde{L})}$ invertible, $x_j \to \bar{x}$ after passing to a subsequence once more. □

In order to construct flows which are admissible mappings for the cohomology theory $H^*_{\mathcal{E}}$ we need to modify the notion of pseudogradient. Let $Y \subset E \setminus K$. A mapping $V : Y \to E$ is said to be a *gradient-like* vector field for Φ on Y if V is locally Lipschitz continuous, $\|V(x)\| \le 1$ for all $x \in Y$ and there is a function $\beta : Y \to \mathbb{R}^+$ such that $\langle\nabla\Phi(x), V(x)\rangle \ge \beta(x)$ for all $x \in Y$ and $\inf_{x\in Z}\beta(x) > 0$ for any set $Z \subset Y$ which is bounded away from K and such that $\sup_{x\in Z}|\Phi(x)| < \infty$.

Lemma 5.4 ([36], Lemma 2.2) *If U is an open subset of E and Φ satisfies (5.3) and $(PS)^*$, then there exists a gradient-like and filtration-preserving vector field V on $U \setminus K$.*

Suppose now A is an isolated compact subset of K. A pair (W, W^-) of closed subsets of E is said to be an *admissible pair* for Φ and A with respect to $\mathcal{E}$ if: (i) W is bounded away from $K \setminus A$, $W^- \subset \mathrm{bd}(W)$ and $A \subset \mathrm{int}(W)$ (bd and int respectively denote the boundary and the interior), (ii) $\Phi|_W$ is bounded, (iii) there exist a neighborhood N of W and a filtration-preserving gradient-like vector field V for Φ on $N \setminus A$, (iv) W^- is the union of finitely many (possibly intersecting) closed sets each lying on a C^1-manifold of codimension 1, $V|_{W^-}$ is transversal to these manifolds, the flow φ of $-V$ can leave W only through W^- and if $x \in W^-$, then $\varphi^t(x) \notin W$ for any $t > 0$.

Since W is bounded away from $K \setminus A$, it is easy to see that for each neighborhood $U \subset W$ of A the critical points of $\Phi|_{W_n}$ are contained in U_n provided n is large enough. The following two results are basic for our Gromoll-Meyer type approach to Morse theory:

Proposition 5.5 ([36], Proposition 2.5) *Suppose Φ satisfies (5.3), $(PS)^*$ and $\Phi(K) \subset (a, b)$. Then (Φ_a^b, Φ_a^a) is an admissible pair for Φ and K.*

Proposition 5.6 ([36], Propositions 2.6 and 2.7) *Suppose Φ satisfies (5.3), $(PS)^*$ and p is an isolated critical point. Then for each open neighborhood U of p there exists an admissible pair (W, W^-) for Φ and p such that $W \subset U$. Moreover, if $(\tilde{W}, \tilde{W}^-)$ is another admissible pair, then $H^*_{\mathcal{E}}(W, W^-) \cong H^*_{\mathcal{E}}(\tilde{W}, \tilde{W}^-)$.*

The existence part of this proposition is shown by considering the flow defined by

$$\frac{d\varphi}{dt} = -\chi(\varphi)V(\varphi), \quad \varphi(0, x) = x,$$

where χ is a cutoff function at $x = p$. Choosing a small $\varepsilon > 0$ and a sufficiently small ball $B_\delta(p) \subset \Phi_{c-\varepsilon}$, where $c = \Phi(p)$, we can take

$$W = \{\varphi^t(x) : t \ge 0,\ x \in \bar{B}_\delta(p),\ \varphi^t(x) \in \Phi_{c-\varepsilon}\} \text{ and } W^- = W \cap \Phi_{c-\varepsilon}^{c-\varepsilon}.$$

The second part of the proposition is also shown by cutting off the flow of V in a suitable way. The proof is rather technical and the strong excision property of $H^*_{\mathcal{E}}$ comes to an essential use here.

If x is an isolated critical point and (W, W^-) an admissible pair for Φ and x, we set

$$C^q_{\mathcal{E}}(\Phi, x) := H^q_{\mathcal{E}}(W, W^-), \quad q \in \mathbb{Z}.$$

According to Proposition 5.6 the critical groups $C^q_{\mathcal{E}}(\Phi, x)$ are well defined by the above formula. If the set K_c is finite and isolated in K for some c, then for each $x_i \in K_c$ there exists an admissible pair (W_i, W_i^-), and we may assume $W_i \cap W_j = \emptyset$ whenever $i \neq j$. So in an obvious notation and by the argument of Lemma 2.5,

$$C^q_{\mathcal{E}}(\Phi, K_c) = \bigoplus_{x \in K_c} C^q_{\mathcal{E}}(\Phi, x), \quad q \in \mathbb{Z}.$$

Let

$$\dim_{\mathcal{E}} H^q_{\mathcal{E}}(X, A) := [\dim H^{q+d_n}(X, A)] \in [\mathbb{Z}] = \prod_{n=1}^{\infty} \mathbb{Z} \Big/ \bigoplus_{n=1}^{\infty} \mathbb{Z}$$

(to be more precise, $[\dim H^{q+d_n}(X, A)] \in [\mathbb{N}_0]$ because all the dimensions are of course nonnegative). If $\dim H^{q+d_n}(X, A) = d$ for almost all n, according to our earlier convention we write $\dim_{\mathcal{E}} H^q_{\mathcal{E}}(X, A) = d$. Suppose K is finite and $\Phi(K) \subset (a, b)$. As in section 2.2, we may define

$$m^q_{\mathcal{E}} := \sum_j \dim_{\mathcal{E}} C^q_{\mathcal{E}}(\Phi, K_{c_j}), \qquad p^q_{\mathcal{E}} := \dim_{\mathcal{E}} H^q_{\mathcal{E}}(\Phi^b_a, \Phi^a_a)$$

and

$$M^{\Phi}_{\mathcal{E}}(t; a, b) := \sum_{q \in \mathbb{Z}} m^q_{\mathcal{E}} t^q, \qquad P^{\Phi}_{\mathcal{E}}(t; a, b) := \sum_{q \in \mathbb{Z}} p^q_{\mathcal{E}} t^q$$

(note that (Φ^b_a, Φ^a_a) is an admissible pair for Φ and K according to Proposition 5.5). These are formal Laurent series with coefficients in $[\mathbb{N}_0]$. If $m^q_{\mathcal{E}}$ and $p^q_{\mathcal{E}}$ are 0 for all but finitely many q, then $M^{\Phi}_{\mathcal{E}}(t; a, b), P^{\Phi}_{\mathcal{E}}(t; a, b) \in [\mathbb{N}_0]\,[t, t^{-1}]$, where $[\mathbb{N}_0][t, t^{-1}]$ is the set of Laurent polynomials with coefficients in $[\mathbb{N}_0]$. If the coefficients are the same for almost all n, we write $M^{\Phi}_{\mathcal{E}}(t; a, b), P^{\Phi}_{\mathcal{E}}(t; a, b) \in \mathbb{N}_0[t, t^{-1}]$.

Theorem 5.7 (Morse inequalities, [36], Theorem 3.1 and Corollary 3.3) *Suppose Φ satisfies (5.3), $(PS)^*$, K is finite, $\Phi(K) \subset (a, b)$ and $m^q_{\mathcal{E}} = 0$ for almost all $q \in \mathbb{Z}$. Then $p^q_{\mathcal{E}} = 0$ for almost all $q \in \mathbb{Z}$ and there exists $Q_{\mathcal{E}} \in \big[\mathbb{N}_0[t, t^{-1}]\big]$ such that*

$$M^{\Phi}_{\mathcal{E}}(t; a, b) = P^{\Phi}_{\mathcal{E}}(t; a, b) + (1 + t) Q_{\mathcal{E}}(t).$$

The proof is rather similar to that of Theorem 2.6; however, in lack of (A2) more complicated sets than Φ^{c_j} and $\Phi^{c_j} \setminus K_{c_j}$ need to be used.

In order to apply the Morse inequalities we need to be able to perform local computations.

Theorem 5.8 ([36], Theorem 5.3) *Suppose Φ satisfies (5.3), $(PS)^*$, p is an isolated critical point of Φ and*

$$\Phi(x) = \Phi(p) + \frac{1}{2}\langle L(x - p), x - p\rangle - \tilde{\psi}(x), \tag{5.4}$$

where L is invertible and $\nabla\tilde{\psi}(x) = o(\|x - p\|)$ as $x \to p$. Then $C^q_{\mathcal{E}}(\Phi, p) = \mathbb{F}$ for $q = M^-_{\mathcal{E}}(L)$ and 0 otherwise.

We note that $M_{\mathcal{E}}^{-}(L)$ is well defined and finite because $L - \tilde{L}$ is compact according to (5.3).

Suppose now $\Phi \in C^2(U, \mathbb{R})$, where U is a neighborhood of an isolated critical point p. Then (5.4) holds, $\dim N(L) < \infty$, $\nabla\Phi(p) = 0$, $\Phi''(p) = L$ and

$$\nabla\Phi(p + z + y) = Ly - \nabla\tilde{\psi}(p + z + y),$$

where $x = p + z + y$, $z \in N(L)$, $y \in R(L)$. Denote the orthogonal projector on $R(L)$ by Q. Since $L|_{R(L)}$ is invertible, we may use the implicit function theorem in order to obtain $\delta > 0$ and a C^1-mapping $\alpha : B_\delta(0) \cap N(L) \to R(L)$ such that $\alpha(0) = 0$, $\alpha'(0) = 0$ and

$$Q\nabla\Phi(p + z + \alpha(z)) \equiv 0.$$

Letting

$$g(z) := \Phi(p + z + \alpha(z)) - \Phi(p) = \frac{1}{2}\langle L\alpha(z), \alpha(z)\rangle - \tilde{\psi}(p + z + \alpha(z)),$$

one readily verifies that 0 is an isolated critical point of g, hence the critical groups $C^q(g, 0)$ of section 2.2 are well defined (however, we use cohomology instead of homology here).

Theorem 5.9 (Shifting theorem, [36], Theorem 5.4) *Suppose Φ satisfies (5.3), $(PS)^*$ and $\Phi \in C^2(U, \mathbb{R})$, where U is a neighborhood of an isolated critical point p. Then*

$$C_{\mathcal{E}}^q(\Phi, p) = C^{q - M_{\mathcal{E}}^{-}(L)}(g, 0), \quad q \in \mathbb{Z}.$$

Remark 5.10 *Other Morse theories for strongly indefinite functionals may be found e.g. in the papers [1, 3, 49] already mentioned and also in [2, 31, 32]. With an exception of [3] they are similar in essence but differ by the way the technical issues have been resolved and by the range of applications. Each of them also has certain advantages and disadvantages. We would also like to mention that we have neither touched upon equivariant Morse theory for strongly indefinite functionals [35] nor upon the Floer and Floer-Conley theories (see e.g. [5], [40], [46]).*

6 Strongly indefinite variational problems

6.1 Hamiltonian systems

As an application of Morse theory for strongly indefinite functionals we shall consider the problem of existence of periodic solutions to Hamiltonian systems

$$\dot{z} = JH_z(z, t), \qquad z \in \mathbb{R}^{2N}, \tag{6.1}$$

where

$$J := \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$$

is the standard symplectic matrix. We shall need the following assumptions on H:

(H_1) $H \in C(\mathbb{R}^{2N} \times \mathbb{R}, \mathbb{R})$, $H_z \in C(\mathbb{R}^{2N} \times \mathbb{R}, \mathbb{R}^{2N})$ and $H(0, t) \equiv 0$;

(H_2) H is 2π-periodic in the t-variable;

(H_3) $|H_z(z,t)| \le c(1+|z|^{s-1})$ for some $c>0$ and $s\in(2,\infty)$;

(H_4) $H_{zz}\in C(\mathbb{R}^{2N}\times\mathbb{R},\mathbb{R}^{4N^2})$;

(H_5) $|H_{zz}(z,t)| \le d(1+|z|^{s-2})$ for some $d>0$ and $s\in(2,\infty)$.

When (H_2) is satisfied, the natural period for solutions of (6.1) is 2π. It is clear that the assumption $H(0,t)\equiv 0$ in (H_1) causes no loss of generality and if (H_5) is assumed, then (H_3) necessarily holds. We also remark that any period T in (H_2) may be normalized to 2π by a simple change of the t-variable.

Below we give a short account of a variational setup for periodic solutions of (6.1). We follow [8] where more details and references may be found. Let $E := H^{1/2}(S^1,\mathbb{R}^{2N})$ be the Sobolev space of 2π-periodic $\mathbb{R}^{2N}$-valued functions

$$z(t) = a_0 + \sum_{k=1}^{\infty} a_k \cos kt + b_k \sin kt, \qquad a_0, a_k, b_k \in \mathbb{R}^{2N} \tag{6.2}$$

such that $\sum_{k=1}^{\infty} k(|a_k|^2+|b_k|^2) < \infty$. Then E is a Hilbert space with an inner product

$$\langle z, w\rangle := 2\pi a_0\cdot a_0' + \pi\sum_{k=1}^{\infty} k(a_k\cdot a_k' + b_k\cdot b_k')$$

(a_k', b_k' are the Fourier coefficients of w). It is well known that the Sobolev embedding $E \hookrightarrow L^q(S^1,\mathbb{R}^{2N})$ is compact for any $q\in[1,\infty)$ but $E\not\subset L^\infty(S^1,\mathbb{R}^{2N})$. Let

$$\Phi(z) := \frac{1}{2}\int_0^{2\pi}(-J\dot z\cdot z)\,dt - \int_0^{2\pi} H(z,t)\,dt \equiv \langle \tilde L z, z\rangle - \psi(z). \tag{6.3}$$

Proposition 6.1 ([8], Proposition 2.1) *If H satisfies (H_{1-2-3}), then $\Phi\in C^1(E,\mathbb{R})$ and $\nabla\Phi(z)=0$ if and only if z is a 2π-periodic solution of (6.1). Moreover, $\nabla\psi$ is completely continuous in the sense that $\nabla\psi(z_j)\to\nabla\psi(z)$ whenever $z_j\rightharpoonup z$. If, in addition, H satisfies (H_4) and (H_5), then $\Phi\in C^2(E,\mathbb{R})$ and $\psi''(z)$ is a compact linear operator for all $z\in E$.*

Suppose $z(t) = a_k\cos kt \pm Ja_k \sin kt$. Then

$$\langle \tilde L z, z\rangle = \int_0^{2\pi}(-J\dot z\cdot z)\,dt = \pm 2\pi k|a_k|^2 = \pm\|z\|^2 \tag{6.4}$$

by a simple computation and it follows that E has the orthogonal decomposition $E = E^+\oplus E^0\oplus E^-$, where

$$E^0 = \{z\in E : z = a_0\in\mathbb{R}^{2N}\},$$
$$E^\pm = \left\{ z\in E : z(t) = \sum_{k=1}^{\infty} a_k\cos kt \pm Ja_k\sin kt,\ a_k\in\mathbb{R}^{2N}\right\}.$$

According to (6.4),

$$\langle \tilde L z, z\rangle = \|z^+\|^2 - \|z^-\|^2 \qquad (z = z^+ + z^0 + z^- \in E^+\oplus E^0\oplus E^-),$$

and since $E^{\pm}$ are infinite-dimensional, Φ is strongly indefinite. Let

$$E_n := \left\{ z \in E : z(t) = a_0 + \sum_{k=1}^{n} a_k \cos kt + b_k \sin kt \right\},$$

then (E_n) is a filtration of E and $\tilde{L}(E_n) \subset E_n$ for all n. Set

$$d_n := N(1+2n) \equiv M^-(\tilde{L}|_{E_n}) + N \tag{6.5}$$

(hence $d_0 = N$ in the notation of section 5.1).

Suppose A is a symmetric $2N \times 2N$ constant matrix and let

$$\langle Bz, w \rangle := \int_0^{2\pi} Az \cdot w \, dt. \tag{6.6}$$

Then B is a selfadjoint operator on E and it follows from the compactness of the embedding $E \hookrightarrow L^2(S^1, \mathbb{R}^{2N})$ that B is compact. A simple computation using (6.2) shows that setting $L := \tilde{L} - B$, we have

$$\langle Lz, z \rangle = -2\pi A a_0 \cdot a_0 + \pi \sum_{k=1}^{\infty} k \left((-Jb_k - \frac{1}{k} A a_k) \cdot a_k + (J a_k - \frac{1}{k} A b_k) \cdot b_k \right). \tag{6.7}$$

Note in particular that $L(E_n) \subset E_n$ for all n. The restriction of the quadratic form $\langle Lz, z \rangle$ to a subspace corresponding to a fixed $k \geq 1$ is represented by the $4N \times 4N$ matrix $\pi k T_k(A)$, where

$$T_k(A) := \begin{pmatrix} -\frac{1}{k} A & -J \\ J & -\frac{1}{k} A \end{pmatrix}.$$

For a symmetric matrix C, set $M^+(C) := M^-(-C)$ and let $M^0(C)$ be the nullity of C. The matrix

$$\begin{pmatrix} 0 & -J \\ J & 0 \end{pmatrix}$$

has the eigenvalues ± 1, both of multiplicity $2N$, hence $M^{\pm}(T_k(A)) = 2N$ for all k large enough. Therefore

$$i^-(A) := M^+(A) - N + \sum_{k=1}^{\infty} (M^-(T_k(A)) - 2N),$$

$$i^+(A) := M^-(A) - N + \sum_{k=1}^{\infty} (M^+(T_k(A)) - 2N),$$

$$i^0(A) := M^0(A) + \sum_{k=1}^{\infty} M^0(T_k(A))$$

are well defined finite numbers and it is not difficult to see that $i^-(A) + i^0(A) + i^+(A) = 0$. Again, we refer to [8] for more details and references. It follows using (6.7) that $i^0(A) =$

$\dim N(L)$ is the number of linearly independent 2π-periodic solutions of the linear system $\dot{z} = JAz$ (so in particular, $\dim N(L) \leq 2N$) and $N(L) \subset E_n$ for almost all n. It can be further seen that L is invertible if and only if $\sigma(JA) \cap i\mathbb{Z} = \emptyset$, where σ denotes the spectrum (this is called the nonresonance condition because the linear system above has $z = 0$ as the only 2π-periodic solution). Also, using (6.7) again,

$$\dim E_n^-(L) = M^+(A) + \sum_{k=1}^{n} M^-(T_k(A)),$$

hence by (5.2) and (6.5), $M_{\mathcal{E}}^-(L) = i^-(A)$. Similarly, $M_{\mathcal{E}}^+(L) := M_{\mathcal{E}}^-(-L) = i^+(A)$. We have sketched a proof of the following

Proposition 6.2 ([36], Proposition 7.1) $\dim N(L) = i^0(A)$ *and* $M_{\mathcal{E}}^{\pm}(L) = i^{\pm}(A)$.

Remark 6.3 *The number $i^-(A)$ (and thus $M_{\mathcal{E}}^-(L)$) equals the Maslov index of the fundamental solution of the system $\dot{z} = JAz$. For comments and references, see [8, Remark 2.8] and [36, Remark 7.2].*

Assume that H_z satisfies the following asymptotic linearity conditions at 0 and infinity:

$$H(z,t) = \frac{1}{2}A_\infty z \cdot z + G_\infty(z,t), \text{ where } (G_\infty)_z(z,t) = o(|z|) \text{ uniformly in } t \text{ as } |z| \to \infty \tag{6.8}$$

and

$$H(z,t) = \frac{1}{2}A_0 z \cdot z + G_0(z,t), \text{ where } (G_0)_z(z,t) = o(z) \text{ uniformly in } t \text{ as } z \to 0. \tag{6.9}$$

Here A_∞ and A_0 are constant symmetric $2N \times 2N$ matrices. It is clear that (6.8) implies (H_3) for any $s > 2$. We shall use the notation $L_\infty = \tilde{L} - B_\infty$ and $L_0 = \tilde{L} - B_0$, where B_∞, B_0 are the operators defined in (6.6), with A replaced by A_∞ and A_0 respectively. We also set

$$\psi_\infty(z) := \int_0^{2\pi} G_\infty(z,t)\,dt \quad \text{and} \quad \psi_0(z) := \int_0^{2\pi} G_0(z,t)\,dt.$$

It is easy to show [8, Lemma 2.4] that $\nabla\psi_\infty(z) = o(\|z\|)$ as $\|z\| \to \infty$ and $\nabla\psi_0(z) = o(\|z\|)$ as $z \to 0$.

Lemma 6.4 *Suppose H satisfies (H_{1-2}) and (6.8). If $\sigma(JA_\infty) \cap i\mathbb{Z} = \emptyset$, then the functional Φ satisfies the $(PS)^*$-condition.*

Proof Let (z_j) be a $(PS)^*$-sequence. Then

$$P_{n_j}\nabla\Phi(z_j) = L_\infty z_j - P_{n_j}\nabla\psi_\infty(z_j) \to 0,$$

so (z_j) is bounded because L_∞ is invertible and $\nabla\psi_\infty(z_j)/\|z_j\| \to 0$ if $\|z_j\| \to \infty$. Now it remains to invoke Lemma 5.3 (with $\tilde{L}$ and ψ given by (6.3)). □

A 2π-periodic solution z_0 of (6.1) is said to be *nondegenerate* if $w = 0$ is the only 2π-periodic solution of the system $\dot{w} = JH_{zz}(z_0(t),t)w$. It is easy to see that z_0 is nondegenerate if and only if $\Phi''(z_0)$ is invertible.

Theorem 6.5 ([36], Theorem 7.4 and Remark 7.7) *Suppose H satisfies (H_{1-2}), (6.8), (6.9) and $\sigma(JA_\infty)\cap i\mathbb{Z} = \sigma(JA_0)\cap i\mathbb{Z} = \emptyset$. If $i^-(A_\infty) \neq i^-(A_0)$, then (6.1) has a nontrivial 2π-periodic solution z_0. Moreover, if H satisfies (H_{4-5}) and z_0 is nondegenerate, then (6.1) has a second nontrivial 2π-periodic solution.*

Proof It follows from our earlier considerations and from the hypotheses that $L_\infty(E_n) \subset E_n$, $L_0(E_n) \subset E_n$, $N(L_\infty) = N(L_0) = \{0\}$, $M^-_{\mathcal{E}}(L_\infty) = i^-(A_\infty)$ and $M^-_{\mathcal{E}}(L_0) = i^-(A_0)$. Suppose 0 is the only 2π-periodic solution of (6.1). Consider the functional $I(z) := \frac{1}{2}\langle L_\infty z, z\rangle$ whose only critical point is 0, and for $R_0 > 0$ let

$$W := \{z = w^+ + w^- \in E^+(L_\infty) \oplus E^-(L_\infty) : \langle \pm L_\infty w^\pm, w^\pm\rangle \le R_0\} \qquad (6.10)$$
$$W^- := \{z \in W : \langle L_\infty w^-, w^-\rangle = -R_0\}.$$

It is easy to see that the mapping $V_1(z) = \|L_\infty z\|^{-1} L_\infty z$ is a gradient-like vector field for I on $E \setminus \{0\}$, it preserves the filtration and (W, W^-) is an admissible pair for I and 0. Hence by Theorem 5.8,

$$H^q_{\mathcal{E}}(W, W^-) = \delta^q_{i^-(A_\infty)}\mathbb{F}. \qquad (6.11)$$

Since $\nabla\psi_\infty(z) = o(\|z\|)$ as $\|z\| \to \infty$, V_1 is also gradient-like for Φ on $E \setminus B_R(0)$ provided R is large enough. By Lemma 5.4 there exists a gradient-like and filtration-preserving vector field V_2 on $B_{R+1}(0) \setminus \{0\}$. Setting $V = \chi_1 V_1 + \chi_2 V_2$, where $\{\chi_1, \chi_2\}$ is a partition of unity subordinated to the cover $\{E \setminus \bar{B}_R(0), B_{R+1}(0) \setminus \{0\}\}$ of $E \setminus \{0\}$, one verifies using this V that (W, W^-) is an admissible pair for Φ and 0 if R_0 is sufficiently large. Therefore by Proposition 5.6, $H^q_{\mathcal{E}}(W, W^-) = C^q_{\mathcal{E}}(\Phi, 0) = \delta^q_{i^-(A_0)}\mathbb{F}$, a contradiction to (6.11).

It is clear that (W, W^-) is also an admissible pair for Φ and $\{0, z_0\}$, possibly after taking larger R and R_0. If z_0 is nondegenerate, then $C^{q_0}_{\mathcal{E}}(\Phi, z_0) = \delta^q_{q_0}\mathbb{F}$ for some $q_0 \in \mathbb{Z}$, hence choosing $t = -1$ in Theorem 5.7 we obtain

$$(-1)^{i^-(A_0)} + (-1)^{q_0} = (-1)^{i^-(A_\infty)},$$

a contradiction again. □

It is in fact not necessary to assume (H_5). It has been shown in [36] that (H_4) implies the existence (but not continuity) of $\Phi''(z_0)$. We also remark that the existence of one nontrivial solution can be shown without using Morse theory, with the aid of a linking argument [8]. And since it is in general not possible to verify whether the solution z_0 is nondegenerate, we can only make a heuristic statement that "usually" this will be the case. Below we give a sufficient condition for (6.1) to have two nontrivial solutions regardless of any nondegeneracy assumption.

Theorem 6.6 ([10], Theorem 2.3, [36], Theorem 7.8) *Suppose H satisfies $(H_{1-2-4-5})$, (6.8), (6.9) and $\sigma(JA_\infty)\cap i\mathbb{Z} = \sigma(JA_0)\cap i\mathbb{Z} = \emptyset$. If $|i^-(A_\infty) - i^-(A_0)| \ge 2N$, then (6.1) has at least 2 nontrivial 2π-periodic solutions.*

Proof Let z_0 be the nontrivial solution obtained in the preceding theorem and suppose there are no other ones. By Theorem 5.9,

$$C^q_{\mathcal{E}}(\Phi, z_0) = C^{q-r_0}(g, 0),$$

where $r_0 = M^-_{\mathcal{E}}(\Phi''(z_0))$ and g is defined in an open neighborhood of 0 in $N(\Phi''(z_0))$. Since $\dim N(\Phi''(z_0)) \le 2N$, $C^q_{\mathcal{E}}(\Phi, z_0) = 0$ for $q - r_0 < 0$ and $q - r_0 > 2N$. Moreover, if g has a local minimum at 0, then $C^{q-r_0}(g, 0) \neq 0$ if and only if $q - r_0 = 0$, and if g has a local maximum there, $C^{q-r_0}(g, 0) \neq 0$ if and only if $q - r_0 = \dim N(\Phi''(z_0))$. Otherwise $C^0(g, 0) = C^{2N}(g, 0) = 0$ (cf. Corollary 3.8). Therefore there exists $q_0 \in \mathbb{Z}$ such that $C^q_{\mathcal{E}}(\Phi, z_0) = 0$ whenever $q < q_0$ and $q > q_0 + 2N - 2$. Hence by (6.11) and the Morse inequalities,

$$t^{i^-(A_0)} + \sum_{q=q_0}^{q_0+2N-2} b_q t^q = t^{i^-(A_\infty)} + (1+t)Q_{\mathcal{E}}(t),$$

where $b_q \in [\mathbb{N}_0]$. Since there is an exponent $i^-(A_\infty)$ on the right-hand side, we must have $q_0 \le i^-(A_\infty) \le q_0 + 2N - 2$. Moreover, $q_0 - 1 \le i^-(A_0) \le q_0 + 2N - 1$. To see this, suppose $i^-(A_0) \le q_0 - 2$ (the other case is similar). Then on the left-hand side there is an exponent $i^-(A_0)$ but no exponents $i^-(A_0) \pm 1$ which is impossible for the right-hand side. Now combining the inequalities above we obtain $|i^-(A_\infty) - i^-(A_0)| \le 2N - 1$, a contradiction. □

As our final application we consider the system (6.1) with H being 2π-periodic in all variables. It is clear that if z_0 is a 2π-periodic solution of (6.1), so is $z_k(t) = z_0(t) + 2\pi k$, $k \in \mathbb{Z}^{2N}$. We shall call two solutions z_1, z_2 *geometrically distinct* if $z_1 - z_2 \not\equiv 2\pi k$ for any $k \in \mathbb{Z}^{2N}$. Let $z = x + v$, $x \in \tilde{E} := E^+ \oplus E^-$, $v \in E^0$. Since $N(\tilde{L}) = E^0 \equiv \mathbb{R}^{2N}$, we may redefine Φ by setting

$$\Phi(x, v) = \frac{1}{2}\int_0^{2\pi} (-J\dot{x} \cdot x)\, dt - \int_0^{2\pi} H(x + v, t)\, dt = \frac{1}{2}\langle \tilde{L}x, x\rangle - \psi(x, v).$$

The periodicity of H with respect to $z_1, \dots, z_{2N}$ implies $\Phi(x, v_1) = \Phi(x, v_2)$ whenever $v_1 \equiv v_2 \pmod{2\pi}$. Therefore v may be regarded as an element of the torus $T^{2N} = \mathbb{R}^{2N}/2\pi\mathbb{Z}^{2N}$ and $\Phi \in C^1(M, \mathbb{R})$, where $M := \tilde{E} \times T^{2N}$. The advantage of such representation of Φ is that distinct critical points of Φ on M correspond to geometrically distinct solutions of (6.1).

Theorem 6.7 *Suppose H is 2π-periodic in all variables and satisfies (H_{1-4}). If all 2π-periodic solutions of (6.1) are nondegenerate, then the number of geometrically distinct ones is at least 2^{2N}.*

This is the second part of the celebrated result by Conley and Zehnder on Arnold's conjecture [21]. The first part asserts that without the nondegeneracy assumption the number of geometrically distinct 2π-periodic solutions is at least $2N + 1$ (see [8, section 2.6] for a sketch of a proof).

Proof We outline the argument. Since the periodicity implies (H_5), $\Phi \in C^2(M, \mathbb{R})$. Let $\mathcal{E} := (M_n, d_n)$, where $M_n = \tilde{E}_n \times T^{2N}$ and $d_n = 2nN$. According to [36, Remark 2.15], the theory developed in section 5.2 still applies. Suppose Φ has only finitely many critical points and define $\tilde{I}(x) = \frac{1}{2}\langle \tilde{L}x, x\rangle$. Then $\tilde{I} : \tilde{E} \to \tilde{E}$, 0 is the only critical point of $\tilde{I}$ and $(\tilde{W}, \tilde{W}^-)$ is an admissible pair for $\tilde{I}$ and 0, where $(\tilde{W}, \tilde{W}^-)$ is defined in the same way as (W, W^-) in (6.10), but with L_∞ replaced by $\tilde{L}$. We see as in the proof of Theorem 6.5 that $(\tilde{W}, \tilde{W}^-) \times T^{2N}$ is an admissible pair for Φ and K provided R and R_0

are large enough. Since $M^-(\tilde{L}|_{E_n}) = d_n$, $C^{q+d_n}(\tilde{I},0) = \delta_0^q\,\mathbb{F}$ for all n and therefore $H^q_{\mathcal{E}}(\tilde{W},\tilde{W}^-) = C^q_{\mathcal{E}}(\tilde{I},0) = \delta_0^q\,\mathbb{F}$. By Künneth's formula [17, p. 8],

$$H^*_{\mathcal{E}}((\tilde{W},\tilde{W}^-)\times T^{2N}) = H^*_{\mathcal{E}}(\tilde{W},\tilde{W}^-)\otimes H^*(T^{2N}) = H^*(T^{2N}).$$

Since $H^*(T^{2N}) = H^*(S^1)\otimes\cdots\otimes H^*(S^1)$ ($2N$ times), $H^q(T^{2N})$ is the direct sum of $\binom{2N}{q}$ copies of $\mathbb{F}$ if $0\le q\le 2N$ and is 0 otherwise (cf. [17, p. 6]). Thus

$$p^q_{\mathcal{E}} = \dim_{\mathcal{E}} H^q_{\mathcal{E}}((\tilde{W},\tilde{W}^-)\times T^{2N}) = \binom{2N}{q}, \quad 0\le q\le 2N$$

and $p^q_{\mathcal{E}} = 0$, $q\notin[0,2N]$. As the coefficients of $Q_{\mathcal{E}}$ are in $[\mathbb{N}_0]$, it follows from Theorem 5.7 that $m^q_{\mathcal{E}} \ge p^q_{\mathcal{E}}$. Denoting the cardinality of the critical set K by $\#K$ and using Theorem 5.8 we obtain

$$\#K = \sum_{q\in\mathbb{Z}} m^q_{\mathcal{E}} \ge \sum_{q\in\mathbb{Z}} p^q_{\mathcal{E}} = \sum_{q=0}^{2N}\binom{2N}{q} = 2^{2N}.$$

□

6.2 Concluding remarks

In the preceding subsection we have assumed that A_∞ and A_0 are constant matrices which satisfy the nonresonance condition $\sigma(JA_\infty)\cap i\mathbb{Z} = \sigma(JA_0)\cap i\mathbb{Z} = \emptyset$. More generally, one can admit t-dependent matrices with 2π-periodic entries and replace the nonresonance condition by certain conditions on G_∞ and G_0. However, the proofs become more technical. See e.g. [30, 36, 51].

Existence of multiple periodic solutions in the setting of Theorem 6.7 has also been studied under the assumption that H is periodic in some (but not necessarily all) z-variables, see e.g. [17, 36, 49]. The result of Conley and Zehnder [21] described in Theorem 6.7 was a starting point for Floer's work on Arnold's conjectures and on what became known as the Floer homology and cohomology, see e.g. the already mentioned references [40, 46], the papers [26, 37] for a solution of the Arnold conjecture, and the recent survey [15].

In [50] the Morse theory of [49] has been applied in order to study bifurcation of nonconstant periodic solutions of small amplitude for the autonomous Hamiltonian system $\dot{z} = JH'(z)$, $z\in\mathbb{R}^{2N}$. By a change of the independent variable one can equivalently look for bifurcation of 2π-periodic solutions for the system $\dot{z} = \lambda H'(z)$, and the results of [50] assert that if $\Phi_\lambda(z) := \frac{1}{2}\int_0^{2\pi}(-J\dot{z}\cdot z)\,dt - \lambda\int_0^{2\pi}H(z)\,dt$ and the Morse index $M^-_{\mathcal{E}}(\Phi''_\lambda(0))$ changes as λ crosses λ_0, then $(0,\lambda_0)$ is a bifurcation point (this can be translated into a statement concerning the original system and the change of index may be expressed in terms of the properties of the matrix $H''(0)$). See also [7] where a more precise result has been obtained by employing a finite-dimensional reduction and equivariant Conley index theory. The above result is local. It has been shown in [22] that if a suitable S^1-degree changes at λ_0, then the bifurcation from $(0,\lambda_0)$ is in fact global.

Other problems where critical point theory for strongly indefinite functionals has been employed are the wave equation of vibrating string type, the beam equation and certain elliptic systems of partial differential equations in bounded domains. For more information on these problems we respectively refer to [36, 58], [58, 59], [36, 50, 59] and the references therein.

If H is convex in z or $H(z,t) = \frac{1}{2}Az \cdot z + G(z,t)$, where G is convex in z, it is possible to replace Φ by a dual functional to which the Morse theory of sections 2 and 3 can be applied. See e.g. [24, 25, 38, 39, 55]. In particular, this approach has turned out to be successful when studying the number of geometrically distinct periodic solutions for autonomous Hamiltonian systems on a prescribed energy surface $H(z) = c$ bounding a convex set in $\mathbb{R}^{2N}$.

Bibliography

[1] A. Abbondandolo: A new cohomology for the Morse theory of strongly indefinite functionals on Hilbert spaces *Topol. Meth. Nonl. Anal.* **9** (1997) 325–382

[2] A. Abbondandolo: *Morse Theory for Hamiltonian Systems* (Chapman & Hall/CRC Research Notes in Mathematics 425, Boca Raton FL, 2001)

[3] A. Abbondandolo and P. Majer: Morse homology on Hilbert spaces *Comm. Pure Appl. Math.* **54** (2001) 689–760

[4] A. Abbondandolo and P. Majer: When the Morse index is infinite *Int. Math. Res. Not.* **71** (2004) 3839–3854

[5] S. Angenent, R. van der Vorst: A superquadratic indefinite elliptic system and its Morse-Conley-Floer homology *Math. Z.* **231** (1999) 203–248

[6] V. Bangert: On the existence of closed geodesics on two-spheres *Internat. J. Math.* **4(1)** (1993) 1–10

[7] T. Bartsch: A generalization of the Weinstein-Moser theorems on periodic orbits of a Hamiltonian system near an equilibrium *Ann. Inst. H. Poincaré, Anal. non lin.* **14** (1997) 691–718

[8] T. Bartsch and A. Szulkin: Hamiltonian systems: periodic and homoclinic solutions by variational methods *In: Handbook of Differential Equations: Ordinary Differential Equations* (Vol. 2, A. Cañada, P. Drábek and A. Fonda eds., Elsevier, Amsterdam, 2005) 77–146

[9] V. Benci: Introduction to Morse theory: a new approach *In: Topological nonlinear analysis* (Progr. Nonl. Diff. Eq. Appl. 15, Birkhäuser, Boston, 1995) 37–177

[10] M. L. Bertotti: Forced oscillations of asymptotically linear Hamiltonian systems *Boll. Un. Mat. Ital.* **7(1-B)** (1987) 729–740

[11] R. Bott: Nondegenerate critical manifolds *Ann. Math.* **60** (1954) 248-261

[12] R. Bott: On the iteration of closed geodesics and the Sturm intersection theory *Comm. Pure Appl. Math.* **9** (1956) 171-206

[13] R. Bott: Marston Morse and his mathematical works *Bull. Amer. Math. Soc.* **3** (1980) 907–950

[14] R. Bott: Lectures on Morse theory, old and new *Bull. Amer. Math. Soc.* **7** (1982) 331–358

[15] A. Cannas da Silva: Symplectic geometry *In: Handbook of Differential Geometry* (Vol. II, F. Dillen and L. Verstraelen (eds.), Elsevier/North-Holland, Amsterdam, 2006) 79–188

[16] K. C. Chang: Morse Theory on Banach spaces and its applications *Chin. Ann. Math. Ser.* **4B** (1983) 381-399

[17] K. C. Chang: *Infinite Dimensional Morse Theory and Multiple Solution Problem* (Birkhäuser, Boston, 1993)

[18] K. C. Chang and M. Jiang: Dirichlet problem with indefinite nonlinearities *Calc. Var. Part. Diff. Eq.* **20** (2004) 257–282

[19] S. Cingolani and G. Vannella: Critical groups computations on a class of Sobolev Banach spaces via Morse index *Ann. Inst. H. Poincaré, Anal. non lin.* **20** (2003) 271–292

[20] M. Clapp and D. Puppe: Critical point theory of symmetric functions and closed geodesics *Diff. Geom. Appl.* **6** (1996) 367-396

[21] C. Conley and E. Zehnder: The Birkhoff-Lewis fixed point theorem and a conjecture of V. Arnold *Invent. Math.* **73** (1983) 33–49

[22] E. N. Dancer and S. Rybicki: A note on periodic solutions of autonomous Hamiltonian systems emanating from degenerate stationary solutions *Diff. Int. Eq.* **12** (1999) 147–160

[23] M. Degiovanni: On Morse theory for continuous functionals *Conf. Sem. Mat. Univ. Bari* **290** (2003) 1–22

[24] I. Ekeland: Une théorie de Morse pour les systèmes hamiltoniens convexes *Ann. Inst. H. Poincaré, Anal. non lin.* **1** (1984) 19–78

[25] I. Ekeland: *Convexity Methods in Hamiltonian Mechanics* (Springer-Verlag, Berlin, 1990)

[26] K. Fukaya and K. Ono: Arnold conjecture and Gromov-Witten invariant *Topology* **38** (1999) 933–1048

[27] J. Franks: Geodesics on S^2 and periodic points of annulus homeomorphisms *Invent. Math.* **108** (1992) 403–418

[28] K. Gęba and A. Granas: Infinite dimensional cohomology theories *J. Math. Pures Appl.* **52** (1973) 145–270

[29] D. Gromoll and W. Meyer: Periodic geodesics on compact riemannian manifolds *J. Diff. Geom.* **3** (1969) 493–510

[30] Y. X. Guo: Nontrivial periodic solutions for asymptotically linear Hamiltonian systems with resonance *J. Diff. Eq.* **175** (2001) 71–87

[31] Y. X. Guo and J. Q. Liu: Morse theory for strongly indefinite functional *Nonl. Anal.* **48** (2002) 831–851

[32] Y. X. Guo, J. Q. Liu and P. A. Zeng: A new morse index theory for strongly indefinite functionals *Nonl. Anal.* **57** (2004) 485–504

[33] N. Hingston: On the growth of the number of closed geodesics on the two-sphere *Internat. Math. Res. Notices* **9** (1993) 253–262

[34] H. Hofer and E. Zehnder: *Symplectic Invariants and Hamiltonian Dynamics* (Birkhäuser, Basel, 1994)

[35] M. Izydorek: A cohomological Conley index in Hilbert spaces and applications to strongly indefinite problems *J. Diff. Eq.* **170** (2001) 22–50

[36] W. Kryszewski and A. Szulkin: An infinite dimensional Morse theory with applications *Trans. Amer. Math. Soc.* **349** (1987) 3181–3234

[37] G. Liu and G. Tian: Floer homology and Arnold conjecture *J. Diff. Geom.* **49** (1998) 1–74
[38] Y. Long: *Index Theory for Symplectic Paths with Applications* (Birkhäuser, Basel, 2002)
[39] J. Mawhin and M. Willem: *Critical Point Theory and Hamiltonian Systems* (Springer-Verlag, New York, 1989)
[40] D. McDuff and D. Salamon: *Introduction to Symplectic Topology* (Oxford University Press, New York, 1998)
[41] F. Mercuri and G. Palmieri: Problems in extending Morse theory to Banach spaces *Boll. U.M.I.* **12(4)** (1975) 397–401
[42] J. W. Milnor: *Morse Theory* (Princeton Univ. Press, Princeton, 1963)
[43] R. S. Palais: Morse theory on Hilbert manifolds *Topology* **2** (1963) 299–340
[44] K. Rybakowski: *The Homotopy Index and Partial Differential Equations* (Springer, Berlin, 1987)
[45] D. Salamon: Morse theory, the Conley index and Floer homology *Bull. London Math. Soc.* **22** (1990) 113–140
[46] M. Schwarz: *Morse Homology* (Birkhäuser, Basel, 1993)
[47] S. Smale: Morse theory and a non-linear generalization of the Dirichlet problem *Ann. Math.* **80** (1964) 382–396
[48] A. Szulkin: Morse theory and existence of periodic solutions of convex Hamiltonian systems *Bull. Soc. Math. France* **116** (1988) 171–197
[49] A. Szulkin: Cohomology and Morse theory for strongly indefinite functionals *Math. Z.* **209** (1992) 375–418
[50] A. Szulkin: Bifurcation for strongly indefinite functionals and a Liapunov type theorem for Hamiltonian systems *Diff. Int. Eq.* **7** (1994) 217–234
[51] A. Szulkin and W. Zou: Infinite dimensional cohomology groups and periodic solutions of asymptotically linear Hamiltonian systems *J. Diff. Eq.* **174** (2001) 369–391
[52] A. J. Tromba: A general approach to Morse theory *J. Diff. Geom.* **12** (1977) 47–85
[53] K. Uhlenbeck: Morse theory on Banach manifolds *J. Funct. Anal.* **10** (1972) 430–445
[54] M. Vigué-Poirrier and D. Sullivan: The homology theory of the closed geodesic problem *J. Diff. Geom.* **11** (1976) 633–644
[55] C. Viterbo: Equivariant Morse theory for starshaped Hamiltonian systems *Trans. Amer. Math. Soc.* **311** (1989) 621–655
[56] M. Willem: Bifurcation, symmetry and Morse theory *Boll. U.M.I.* **3-B** (1989) 17–27
[57] M. Willem: *Minimax Theorems* (Birkhäuser, Boston, 1996)
[58] P. A. Zeng, J. Q. Liu and Y. X. Guo: Computations of critical groups and applications to asymptotically linear wave equation and beam equation *J. Math. Anal. Appl.* **300** (2004) 102–128
[59] W. Zou: Computations of the cohomology groups with the applications to asymptotically linear beam equations and noncooperative elliptic systems *Comm. Part. Diff. Eq.* **27** (2002) 115–147

THOMAS BARTSCH

Mathematisches Institut, Universität Giessen, Arndtstr. 2, 35392 Giessen, Germany

E-mail: `Thomas.Bartsch@math.uni-giessen.de`

ANDRZEJ SZULKIN

Department of Mathematics, Stockholm University, 106 91 Stockholm, Sweden

E-mail: `andrzejs@math.su.se`

MICHEL WILLEM

Institut de Mathématique Pure et Appliquée, Université Catholique de Louvain,
Chemin du Cyclotron, 2, 1348 Louvain-La-Neuve, Belgium

E-mail: `willem@math.ucl.ac.be`

This page intentionally left blank

Demeter Krupka and David Saunders

Index theory

David Bleecker

Contents

1 Introduction and some history

When first considering infinite dimensional linear spaces, there is the immediate realization that there are injective or surjective linear endomorphisms which are not isomorphisms, and more generally the dimension of the kernel minus that of the cokernel (i.e., the index) could be any integer. However, in the classical theory of Fredholm integral operators which goes back at least to the early 1900s (see [22]), one is dealing with perturbations of the identity and the index is zero. Several sources point to Fritz Noether (in his study [38] of singular integral operators, published in 1921), as the first to encounter the phenomenon of a nonzero index for operators naturally arising in analysis *and* to give a formula for the index in terms a winding number constructed from data defining the operator. Over some decades, this result was generalized in various directions by I.N. Vekua and others (see [45]). Meanwhile, many working mainly in abstract functional analysis were producing results, such as the stability of the index of a Fredholm operator under perturbations by compact operators or bounded operators of sufficiently small operator norm (e.g., first J.A. Dieudonne [21], followed by F.W. Atkinson [14], B. Yood [48], I.Z. Gohberg and M.G. Krein [28], etc.).

Around 1960, the time was ripe for I.M. Gelfand (see [23]) to propose that the index of an elliptic differential operator (with suitable boundary conditions in the presence of a boundary) should be expressible in terms of the coefficients of highest order part (i.e., the principal symbol) of the operator, since the lower order parts provide only compact perturbations which do not change the index. Indeed, a continuous, ellipticity-preserving deformation of the symbol should not affect the index, and so Gelfand noted that the index should only depend on a suitably defined homotopy class of the principal symbol. The hope was that the index of an elliptic operator could be computed by means of a formula involving only the topology of the underlying domain (or manifold), the bundles involved, and the symbol of the operator. In early 1962, M.F. Atiyah and I.M. Singer discovered the Dirac operator in the context of Riemannian geometry and were busy working at Oxford on a proof that the $\hat{A}$-genus of a spin manifold is the index of this Dirac operator. At that time Stephen Smale happened to pass through Oxford and turned their attention to Gelfand's general program described in [23]. Drawing on the foundational and case work of analysts (e.g., M.S. Agranovic, A.D. Dynin, L. Nirenberg, R.T. Seeley and A.I. Volpert), particularly that involving pseudo-differential operators, Atiyah and Singer discovered and proved the desired index formula at Harvard in the Fall of 1962. Moreover, the Riemannian Dirac operator played a major role in establishing the general case. The details of this original proof involving cobordism actually first appeared in [41]. A K-theoretic embedding proof was given in [12], the first in a series of five papers. This proof was more direct and susceptible to generalization (e.g., to G-equivariant elliptic operators [10] and families of elliptic operators [13]).

The approach to proving the Index Theorem in [12] is based on the following clever strategy. The invariance of the index under homotopy implies that the index (say, the *analytic index*) of an elliptic operator is stable under rather dramatic, but continuous, changes in its principal symbol while maintaining ellipticity. Using this fact, one finds (after considerable effort) that the index (i.e., the *analytic index*) of an elliptic operator transforms predictably under various global operations such as embedding and extension. Using K-theory and Bott periodicity, a topological invariant (say, the *topological index*) with the same transformation properties under these global operations is constructed from the symbol of the elliptic operator. One then verifies that a general index function having these properties is unique, subject to normalization. To deduce the Atiyah–Singer Index Theorem (i.e., *analytic index* = *topological index*), it then suffices to check that the two indices are the same in the trivial case where the base manifold is just a single point. A particularly nice exposition of this approach for twisted Dirac operators over even-dimensional manifolds (avoiding many complications of the general case) is found in E. Guentner's article [30] following an argument of P. Baum.

Not long after the K-theoretical embedding proof (and its variants), there emerged a fundamentally different means of proving the Atiyah–Singer Index Theorem, namely the *heat kernel method*. This is outlined here (see subsection 5.4) in the important case of the chiral half $\mathcal{D}^+$ of a twisted Dirac operator $\mathcal{D}$. In the index theory of closed manifolds, one usually studies the index of a chiral half $\mathcal{D}^+$ instead of the total Dirac operator $\mathcal{D}$, since $\mathcal{D}$ is symmetric for compatible connections and then $\operatorname{index} \mathcal{D} = 0$. The heat kernel method had its origins in the late 1960s (e.g., in [35]) and was pioneered in the works [42], [26], [6]. The method is exhibited with a high degree of virtuosity in book [15]. In the final analysis, it is debatable as to whether this method is really much shorter or better. That depends on the background and tastes of the beholder. Geometers and analysts (as opposed to topologists) are likely to find the heat kernel method appealing. The method not only applies to geometric operators which are expressible in terms of twisted Dirac operators, but also largely for more general elliptic pseudo-differential operators, as R.B. Melrose has done in [36]. Moreover, the heat kernel method gives the index of a "geometric" elliptic differential operator naturally as the integral of a characteristic form (a polynomial of curvature forms) which is expressed solely in terms of the geometry of the operator itself (e.g., curvatures of metric tensors and connections). One does not destroy the geometry of the operator by using ellipticity-preserving deformations. Rather, in the heat kernel approach, the invariance of the index under changes in the geometry of the operator is a consequence of the index formula itself rather than a means of proof. However, considerable analysis and effort are needed to obtain the heat kernel for $e^{-t\mathcal{D}^2}$ and to establish its asymptotic expansion as $t \to 0^+$. Moreover, in [34], we are cautioned that the index theorem for families (in its strong form) generally involves torsion elements in K-theory that are not detectable by cohomological means, and hence are not computable in terms of local densities produced by heat asymptotics. Nevertheless, when this difficulty does not arise, the K-theoretical expression for the topological index may be less appealing than the integral of a characteristic form, particularly for those who already understand and appreciate the geometrical formulation of characteristic classes. More importantly, the heat kernel approach exhibits the index as just one of a whole sequence of spectral invariants appearing as coefficients of terms of the asymptotic expansion (as $t \to 0^+$) of the trace of the relevant heat kernel (see P.B. Gilkey's contribution to this volume). All disputes aside, the student who learns *both* approaches and formulations to the index formula will be more accomplished (and probably older).

Insofar as the coverage of topics in this article is concerned, we hope the above table of contents needs no elaboration, except to say that space limitations prevented the inclusion of some important topics (e.g., the index theorem for families, and index theory for manifolds with boundary, other than the A-P-S Theorem). However, we now provide some guidance for further study. A fairly complete exposition, by Atiyah himself, of the history of index theory from 1963 to 1984 is found in Volume 3 of [1] and duplicated in Volume 4. Volumes 3, 4 and 5 contain many unsurpassed articles written by Atiyah and collaborators on index theory and its application to gauge theory. We all owe a debt of gratitude to Herbert Schröder for the definitive guide to the literature on index theory (and its roots and offshoots) through 1994 in Chapter 5 of the excellent book [27] of P.B. Gilkey. The present author has benefited greatly not only from this book, but also from the marvelous work [34] of H.B. Lawson and M.L. Michelsohn. In that book, there are proofs of index formulas in various contexts, and numerous beautiful applications illustrating the power of Dirac operators, Clifford algebras and spinors in the geometrical analysis of manifolds, immersions, vector fields, and much more. The classic book [44] of P. Shanahan is also a masterful, elegant exposition of not only the standard index theorem, but also the G-index theorem and its numerous applications. A fundamental source on index theory for certain open manifolds and manifolds with boundary is the authoritative book [36] of R.B. Melrose. In the case of boundary-value problems for Dirac operators, there is also the carefully written and very readable book [20] of B. Booss–Bavnbek and K.P. Wojciechowski.

2 Fredholm operators – theory and examples

2.1 Definition of Fredholm operator and index

Let H_1 and H_2 be separable Hilbert spaces, and let $F : H_1 \to H_2$ be a bounded (continuous) linear transformation. The **kernel** of F is the closed subspace $\operatorname{Ker} F := \{v \in H_1 \mid F(v) = 0\} = F^{-1}(\{0\})$ and the **cokernel** of F is the vector space $\operatorname{Coker} F := H_2/F(H_1)$. Suppose that $\operatorname{Ker}(F)$ and $\operatorname{Coker}(F)$ have finite dimension. Then F is a **Fredholm operator**. The concept of "Fredholm operator" can be extended to encompass certain bounded (and unbounded) operators between Banach spaces, but we will not do so. It might be argued that we could further simplify matters by taking $H_2 = H_1$, but in applications F is often a Sobolev extension of a differential operator with H_1 and H_2 *different* Sobolev spaces, which is desirable to ensure that F is bounded.

It is often required that the range of F be closed, but this is implied by $\dim \operatorname{Coker}(F) < \infty$. Indeed, since $\operatorname{Ker}(F)$ is closed, $H_1/\operatorname{Ker}(F)$ is Hilbertable and F induces a continuous injection

$$G : H_1/\operatorname{Ker} F \to H_2$$

of Hilbert spaces. Let H_3 be an algebraic complement of $F(H_1)$ in H_2. Since $\dim H_3 = \dim \operatorname{Coker}(F) < \infty$, H_3 is a Hilbert space, as is $(H_1/\operatorname{Ker} F) \oplus H_3$. Define

$$\widetilde{G} : (H_1/\operatorname{Ker} F) \oplus H_3 \to H_2 \text{ by } \widetilde{G}(v, w) = G(v) + w.$$

Note that $\widetilde{G}$ is injective, and bounded since

$$\begin{aligned} ||\widetilde{G}(v,w)|| &= \|G(v) + w\| \le \|G(v)\| + \|w\| \le \|G\| \, \|v\| + \|w\| \\ &\le (\|G\| + 1)\,(\|v\| + \|w\|) \le 2\,(\|G\| + 1)\,\|(v,w)\|\,. \end{aligned}$$

By the open mapping theorem, $\widetilde{G}$ is a topological isomorphism, and hence

$$F(H_1) = G(H_1/\operatorname{Ker} F) = \widetilde{G}((H_1/\operatorname{Ker} F) \oplus \{0\})$$

is closed.

The **adjoint** of F is $F^* : H_2 \to H_1$ which is uniquely determined by the property:

$$\langle F^* w, v\rangle = \langle w, Fv\rangle \text{ for all } v \in H_1 \text{ and } w \in H_2.$$

We have $\|F^*\| \le \|F\|$, since

$$\begin{aligned} \|F^* w\|^2 &= |\langle F^* w, F^* w\rangle| = |\langle w, FF^* w\rangle| \le \|w\| \, \|F\| \, \|F^* w\| \\ &\Rightarrow \|F^* w\| \le \|F\| \, \|w\|\,. \end{aligned}$$

Since $F^{**} = F$, we have $\|F\| = \|F^*\|$. Clearly

$$\operatorname{Ker} F^* = F(H_1)^\perp := \{w \in H_2 \mid \langle F(v), w\rangle = 0 \text{ for all } v \in H_1\}\,.$$

As $F(H_1)$ is closed, $H_2 = F(H_1) \oplus F(H_1)^\perp$, and hence

$$\dim \operatorname{Coker}(F) = \dim F(H_1)^\perp = \dim \operatorname{Ker} F^*.$$

Note that F^* is Fredholm since $\dim \operatorname{Ker} F^* = \dim F(H_1)^\perp < \infty$ and $F^*(H_2) = (\ker F)^\perp$.

The **index** of F is defined by

$$\begin{aligned} \operatorname{index} F &:= \dim \operatorname{Ker} F - \dim \operatorname{Coker} F \\ &= \dim \operatorname{Ker} F - \dim F(H_1)^\perp = \dim \operatorname{Ker} F - \dim \operatorname{Ker} F^*. \end{aligned}$$

Plainly, $\operatorname{index} F^* = -\operatorname{index} F$. If H_1 and H_2 are finite-dimensional, then $\operatorname{index} F$ is independent of F, since

$$\begin{aligned}\operatorname{index} F &= \dim \operatorname{Ker}(F) - \dim F(H_1)^\perp = \dim H_1 - \dim F(H_1) - \dim F(H_1)^\perp \\ &= \dim H_1 - \big(\dim F(H_1) + \dim F(H_1)^\perp\big) = \dim H_1 - \dim H_2.\end{aligned}$$

In general $\dim \operatorname{Ker}(F)$ and $\dim \operatorname{Ker}(F^*)$ (if positive) can both decrease under a perturbation of F, but we might suspect that the difference (i.e., $\operatorname{index} F$) is invariant under suitable perturbations or deformations, as it is in finite dimensions. Before exploring such stability properties, we now look at some more examples.

2.2 Elementary examples

Example 2.1 (The classical Fredholm Alternative) Let $-\infty \leq a < b \leq \infty$ and suppose $K \in L^2([a,b] \times [a,b])$ (i.e., $K : [a,b] \times [a,b] \to \mathbb{C}$ is measurable with $\int_a^b \int_a^b |K(x,y)|^2 \, dxdy < \infty$. Let $F : L^2([a,b]) \to L^2([a,b])$ be given by

$$F[u](x) = u(x) + \int_a^b K(x,y)\, u(y)\, dx$$

This is a Fredholm operator and its adjoint is given by

$$F^*[v](y) = v(y) + \int_a^b \overline{K(x,y)}\, v(x)\, dx$$

In this case, the Fredholm alternative says that $\dim \ker F = \dim \ker F^*$ (i.e., $\operatorname{index} F = 0$). In other words, if $v_1, \ldots, v_k$ form a basis of $\ker F^*$, then $F[u] = v$ has a solution u if and only if $\langle v_1, v\rangle = \cdots = \langle v_k, v\rangle = 0$. Moreover, two solutions differ by an element of $\ker F$ which also has dimension k.

Example 2.2 (shift operators) Take $H_1 = H_2 = L^2(\mathbb{Z}_+)$ be the Hilbert space of sequences $(c_0, c_1, c_2, \ldots)$ of complex numbers for which $\sum_{n=0}^\infty |c_n|^2 < \infty$, and define the left and right shift operators S_L and S_R on $L^2(\mathbb{Z}_+)$ via

$$S_L(c_0, c_1, c_2, \ldots) := (c_1, c_2, \ldots) \text{ and } S_R(c_0, c_1, c_2, \ldots) := (0, c_0, c_1, \ldots).$$

Clearly, $\operatorname{index} S_L = 1$ and $\operatorname{index} S_R = -1$. By considering compositions of S_L and S_R, we can achieve any integer for the index.

Example 2.3 (Toeplitz operators) Let

$$S^1 := \{z \in \mathbb{C} \mid |z| = 1\} = \left\{e^{i\theta} \in \mathbb{C} \mid \theta \in \mathbb{R}\right\} \cong \mathbb{R}/(2\pi\mathbb{Z})$$

denote the unit circle. Recall that the functions $e_n(z) := z^n/\sqrt{2\pi}$ $(n \in \mathbb{Z})$ form an orthonormal basis (complete orthonormal set) of $L^2(S^1)$. For $u \in L^2(S^1)$,

$$u = \sum\nolimits_{n=-\infty}^{\infty} \langle u, e_n\rangle\, e_n = \lim_{N\to\infty} \sum\nolimits_{n=-N}^{N} \langle u, e_n\rangle\, e_n \text{ in } L^2(S^1).$$

Let $P : L^2(S^1) \to L^2(S^1)$ be the projection onto the closed subspace

$$H_0 := \overline{\operatorname{span}\{e_n \mid n \geq 0\}}.$$

This is the subspace of functions $u \in L^2(S^1)$ with Fourier coefficients $c_n(u) := \frac{1}{2\pi}\int_{-\pi}^{\pi} u(e^{i\theta})e^{-in\theta}d\theta = \frac{1}{\sqrt{2\pi}}\langle u, e_n\rangle$ which are 0 for $n < 0$. Suppose that $f : S^1 \to \mathbb{C}^* := \mathbb{C}\setminus\{0\}$ is continuous. Let $M_f : L^2(S^1) \to L^2(S^1)$ be the multiplication operator $M_f(u) = fu$. Define $T_f : H_0 \to H_0$ by $T_f = P \circ M_f|_{H_0}$; i.e., $T_f(u) = P(M_f(u))$ for $u \in H_0$; T_f is known as a Toeplitz operator. Note that for $f(z) = z^m$ (or simply $f = z^m$, where z is the identity function on S^1), we obtain

$$T_{z^m}(e_n) = P(e_{n+m}) = \begin{cases} e_{n+m} & \text{for } n \geq -m \\ 0 & \text{for } n < -m. \end{cases}$$

Thus, for $m \geq 0$, $\operatorname{Ker} T_{z^m} = \{0\}$ and $\operatorname{Coker} T_{z^m} \cong \operatorname{span}(\{e_0, \ldots, e_{m-1}\})$, while for $m < 0$, we get $\operatorname{Ker} T_{z^m} = \operatorname{span}(\{e_0, \ldots, e_{|m|-1}\})$. In either case, $\operatorname{index} T_{z^m} = -m$. It can be shown that $T_{(\cdot)} : C^0(S^1) \to \mathcal{B}(H_0)$ is continuous. Moreover T_f is Fredholm for f with values in $\mathbb{C}^*$. As a consequence of the invariance of the index under continuous deformation in $\mathcal{F}(H_0)$ (see below), we get that for $f : S^1 \to \mathbb{C}^*$, $\operatorname{index} T_f$ is minus the winding number of f about 0.

Example 2.4 (oblique Neumann problem) Let $U := \{z \in \mathbb{C} \mid |z| < 1\}$ Consider a C^∞ vector field $\nu : S^1 \to \mathbb{C}$ on the boundary $\partial U = S^1 := \{z \mid |z| = 1\}$. For $f \in C^\infty(U, \mathbb{C})$, $z \in S^1$, and $\nu(z) = \alpha(z) + i\beta(z)$, one defines the "directional derivative" of f relative to the vector field ν at the point z to be

$$\frac{\partial f}{\partial \nu}(z) := \alpha(z)\frac{\partial f}{\partial x}(z) + \beta(z)\frac{\partial f}{\partial y}(z) \in \mathbb{C}.$$

Let $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = 4\frac{\partial^2}{\partial z \partial \bar{z}} : C^\infty(U, \mathbb{C}) \to C^\infty(U, \mathbb{C})$ be the Laplace operator. The pair $(\Delta, \frac{\partial}{\partial \nu})$ defines a linear operator

$$(\Delta, \tfrac{\partial}{\partial \nu}) : C^\infty(U, \mathbb{C}) \to C^\infty(U, \mathbb{C}) \oplus C^\infty(S^1, \mathbb{C}) \text{ given by } f \mapsto (\Delta f, \tfrac{\partial f}{\partial \nu}).$$

The following result is due to I. N. Vekua (see [45]).

Theorem 2.5 *For $p \in \mathbb{Z}$ and $\nu(z) := z^p$, we have that $(\Delta, \frac{\partial}{\partial \nu})$ is an operator with finite-dimensional kernel and cokernel, and*

$$\operatorname{index}(\Delta, \tfrac{\partial}{\partial \nu}) = 2(1-p).$$

Remark 2.6 The theorem of Vekua remains true, if we replace z^p by any nonvanishing "vector field" $\nu : S^1 \to \mathbb{C}\setminus\{0\}$ with "winding number" p. This is because of the fact that such a vector field can be deformed without zeros on S^1 to z^p, and the index is invariant under the induced deformation of the operator $(\Delta, \frac{\partial}{\partial \nu})$. In the next subsection, we explore the invariance properties of the index under suitable deformations or perturbations.

2.3 Basic properties

Let $\mathcal{B}(H_1, H_2)$ be the Banach space of all bounded linear maps from H_1 to H_2. Recall that $K \in \mathcal{B}(H_1, H_2)$ is **compact** if K maps bounded subsets of H_1 into relatively compact subsets of H_2. Define $\mathcal{K}(H_1, H_2)$ to be the subset of compact elements of $\mathcal{B}(H_1, H_2)$. If $\dim K(H_1) < \infty$ (i.e., K has **finite rank**), then K is compact, but there are $K \in \mathcal{K}(H_1, H_2)$ with $\dim K(H_1) = \infty$ and $K(H_1)$ not closed. Clearly, $\mathcal{K}(H_1, H_2)$ is a linear subspace of $\mathcal{B}(H_1, H_2)$. Also, it is immediate that for $T \in \mathcal{B}(H_1, H_2)$ and $K_i \in \mathcal{K}(H_i) := \mathcal{K}(H_i, H_i)$ $(i = 1, 2)$, we have $K_2 \circ T$ and $T \circ K_1 \in \mathcal{K}(H_1, H_2)$. In particular, for a single Hilbert space H, $\mathcal{K}(H)$ is a two-sided ideal in $\mathcal{B}(H) := \mathcal{B}(H, H)$. Let $\mathcal{B}_f(H_1, H_2) := \{T \in \mathcal{B}(H_1, H_2) \mid \dim T(H_1) < \infty\}$, the linear subspace of operators in $\mathcal{B}(H_1, H_2)$ with finite rank.

Theorem 2.7 *$\mathcal{K}(H_1, H_2)$ is the closure of the subset $\mathcal{B}_f(H_1, H_2)$ of $\mathcal{B}(H_1, H_2)$ in the strong topology induced by the operator norm.*

Proof We first show that $\mathcal{K}(H_1, H_2)$ is closed. Let $T \in \mathcal{B}(H_1, H_2)$ be an operator in the closure of $\mathcal{K}(H_1, H_2)$. To prove that $T \in \mathcal{K}(H_1, H_2)$, we verify that the image $T(B_{H_1})$ of the closed unit ball B_{H_1} in H_1 is precompact. Let $v_1, v_2, \ldots$ be a bounded sequence in B_{H_1} and let $T_1, T_2, \ldots$ be a sequence in $\mathcal{K}(H_1, H_2)$ converging to T. By a diagonal argument, we can find a subsequence $u_1, u_2, \ldots$ of $v_1, v_2, \ldots$, such that $T_n(v_1), T_n(v_2), \ldots$ converges for each $n = 1, 2, \ldots$. Given $\varepsilon > 0$, we can choose n large enough so that $\|T - T_n\| < \varepsilon/3$ and $\|T_n(u_j) - T_n(u_k)\| < \varepsilon/3$ for $j, k \geq M$. Then for $j, k \geq M$

$$\begin{aligned}
&\|T(u_j) - T(u_k)\| \\
&< \|T(u_j) - T_n(u_j)\| + \|T_n(u_j) - T_n(u_k)\| + \|T_n(u_k) - T(u_k)\| \\
&< \varepsilon/3\,\|u_j\| + \varepsilon/3 + \varepsilon/3\,\|u_k\| \leq \varepsilon
\end{aligned}$$

Thus, $T(u_1), T(u_2), \ldots$ is a Cauchy sequence, converging by the completeness of H_2. We now show that each $K \in \mathcal{K}(H_1, H_2)$ is the limit of a sequence of operators in $\mathcal{B}_f(H_1, H_2)$. Let $e_1, e_2, \ldots$ be a complete orthonormal system in H_2, and let $Q_n : H_2 \to \mathrm{span}(e_1, \ldots, e_n)$ be the orthogonal projection from H_2 to the linear span of $e_1, \ldots, e_n$. It remains to prove that, for each $K \in \mathcal{K}$, the sequence $(Q_n K)_1^\infty$ converges to K in $\mathcal{B}(H_1, H_2)$. For this, choose (for each $\varepsilon > 0$) a finite covering of the precompact set $K(B_{H_1})$ by balls of radius $\varepsilon/3$, say with centers at $K(u_1), \ldots, K(u_m)$. Then choose some $n \in \mathbb{N}$, so large that $\|Ku_i - Q_n K u_i\| < \varepsilon/3$ for $i = 1, \ldots, m$. This is no problem, since $(Q_n)_1^\infty$converges pointwise to the identity. For each $u \in B_{H_1}$, we then have (noting that $\|Q_n\| = 1$ for the last term)

$$\|Ku - Q_n K u\| \leq \|Ku - Ku_i\| + \|Ku_i - Q_n K u_i\| + \|Q_n K u_i - Q_n K u\| < 3(\varepsilon/3)$$

for some $i \in \{1, \ldots, m\}$. Thus, we have proven $\|K - Q_n K\| < \varepsilon$ for all sufficiently large n, depending on ε. Thus, $K \subset \overline{\mathcal{B}_f(H_1, H_2)}$ and $\mathcal{B}_f(H_1, H_2) \subset K \Rightarrow \overline{\mathcal{B}_f(H_1, H_2)} \subset \overline{K} = K$, and so $K = \overline{\mathcal{B}_f(H_1, H_2)}$. □

Proposition 2.8 *If $K \in \mathcal{K}(H_1, H_2)$, then $K^* \in \mathcal{K}(H_2, H_1)$.*

Proof If T has finite rank, then $T = \sum_{i=1}^n \langle \cdot, u_i \rangle v_i$, where $u_1, \ldots, v_n \in H$. Note that $T^* = \sum_{i=1}^n \langle \cdot, v_i \rangle u_i$, since

$$\begin{aligned}
\langle T(v), w \rangle &= \left\langle \sum\nolimits_{i=1}^n \langle v, u_i \rangle v_i, w \right\rangle = \sum\nolimits_{i=1}^n \langle v, u_i \rangle \langle v_i, w \rangle \\
&= \left\langle v, \sum\nolimits_{i=1}^n \overline{\langle v_i, w \rangle} u_i \right\rangle = \left\langle v, \sum\nolimits_{i=1}^n \langle w, v_i \rangle u_i \right\rangle.
\end{aligned}$$

Thus, T^* also has finite rank. Now, we can reduce the general case $K \in \mathcal{K}$ to the finite-rank case. Namely, $K \in \mathcal{K}(H_1, H_2)$ is the limit of a sequence $(T_n)_1^\infty$ in $\mathcal{B}_f(H_1, H_2)$ and $T_n^* \to K^*$, since $\|T_n^* - K^*\| = \left\|(T_n - K)^*\right\| = \|T_n - K\|$. Hence, $K^* \in \mathcal{K}(H_2, H_1)$. □

Let $\mathcal{F}(H_1, H_2)$ be the set of Fredholm operators from H_1 to H_2.

Theorem 2.9 *For $F \in \mathcal{B}(H_1, H_2)$, we have $F \in \mathcal{F}(H_1, H_2)$ if and only if there is $G \in \mathcal{B}(H_2, H_1)$ such that*

$$GF - I_{H_1} \in \mathcal{K}(H_1, H_1) \text{ and } FG - I_{H_2} \in \mathcal{K}(H_2, H_2).$$

In other words, the Fredholm operators are precisely those which are invertible modulo compact operators.

Proof If $F \in \mathcal{F}(H_1, H_2)$, then $H_1 = \operatorname{Ker} F \oplus (\operatorname{Ker} F)^\perp$ and $H_2 = F(H_1) \oplus F(H_1)^\perp$. Then $F|_{(\operatorname{Ker} F)^\perp} : (\operatorname{Ker} F)^\perp \to F(H_1)$ is a continuous linear isomorphism, and $(F|_{(\operatorname{Ker} F)^\perp})^{-1} : F(H_1) \to (\operatorname{Ker} F)^\perp$ is then continuous (by open mapping theorem). Take

$$G := (F|_{(\operatorname{Ker} F)^\perp})^{-1} \oplus 0 : F(H_1) \oplus F(H_1)^\perp \to H_1.$$

Then, $GF = 0 \oplus I_{(\operatorname{Ker} F)^\perp} : \operatorname{Ker} F \oplus (\operatorname{Ker} F)^\perp \to (\operatorname{Ker} F)^\perp$ and

$$GF - I_{H_1} = -I_{\operatorname{Ker} F} \oplus 0 : \operatorname{Ker} F \oplus (\operatorname{Ker} F)^\perp \to \operatorname{Ker} F$$

is of finite rank. Moreover,

$$FG = I_{F(H_1)} \oplus 0 : F(H_1) \oplus F(H_1)^\perp \to F(H_1) \text{ and}$$
$$FG - I_{H_2} = -0 \oplus I_{F(H_1)^\perp} : F(H_1) \oplus F(H_1)^\perp \to F(H_1)^\perp,$$

which is also of finite rank. Conversely, suppose that $GF - I_{H_1} \in \mathcal{K}(H_1, H_1)$ and $FG - I_{H_2} \in \mathcal{K}(H_2, H_2)$. Then $\dim \operatorname{Ker} F < \infty$, since otherwise $(GF - I_{H_1})_{|\operatorname{Ker} F} = I_{\operatorname{Ker} F}$, and $GF - I_{H_1}$ would not be compact. Moreover,

$$FG - I_{H_2} \in \mathcal{K}(H_2, H_2) \Rightarrow G^* F^* - I_{H_2} \in \mathcal{K}(H_2, H_2)$$

and hence $\dim \operatorname{Ker} F^* < \infty$ by the same reasoning. Although $\dim \operatorname{Ker} F^* < \infty$, we still need to show that $F(H_1)$ is closed. Suppose that $F(v_k) \to w \in H_2$ for $v_k \in (\operatorname{Ker} F)^\perp$. We need to show that by taking a suitable subsequence $v_k \to v$, so that $F(v) = w \in F(H_1)$. We first prove that v_k is bounded. If not we may assume by taking a subsequence that $\|v_k\| \to \infty$. Then, $\|F(v_k/\|v_k\|)\| = \frac{1}{\|v_k\|}\|F(v_k)\| \to 0$, since $\|F(v_k)\| < \|w\| + 1$ for k sufficiently large, and so $F(\frac{v_k}{\|v_k\|}) \to 0$. Now, since $I_{H_1} - GF$ is compact, by taking a subsequence, we may assume $(I_{H_1} - GF)(v_k/\|v_k\|) \to u$, for some $u \in H_1$. Since $G_L F(v_k/\|v_k\|) \to 0$, we then have $v_k/\|v_k\| \to u \in (\operatorname{Ker} F)^\perp$, and

$$F(u) = F(\lim_{k\to\infty} v_k/\|v_k\|) = \lim_{k\to\infty} F(v_k/\|v_k\|) = 0.$$

As $u \in (\operatorname{Ker} F)^\perp$, we have $u = 0$, but $v_k/\|v_k\| \to u \Rightarrow 1 = \|v_k/\|v_k\|\| \to \|u\| \Rightarrow \|u\| = 1$. Thus, v_k is bounded. Since $I_{H_1} - GF$ is compact, by passing to a subsequence we may assume that $(I_{H_1} - G_L F)(v_k)$ converges, say to v'. Then, as required,

$$v_k = (I_{H_1} - GF)(v_k) + GF(v_k) \to v' + GF(v_k) \to v' + Gw.$$

□

Theorem 2.10 *For $T \in \mathcal{F}(H_1, H_2)$ and $S \in \mathcal{F}(H_2, H_3)$, we have $ST \in \mathcal{F}(H_1, H_3)$ and*

$$\operatorname{index}(ST) = \operatorname{index} S + \operatorname{index} T.$$

Proof Since $\dim \operatorname{Ker}(ST) \le \dim \operatorname{Ker} T + \dim \operatorname{Ker} S < \infty$, and

$$\dim \operatorname{Coker}(ST) \le \dim \operatorname{Coker} S + \dim \operatorname{Coker} T < \infty,$$

we have $ST \in \mathcal{F}(H_1, H_3)$. The exact sequence

$$0 \to T^{-1}(0) \hookrightarrow \operatorname{Ker}\left(T^{-1}S^{-1}(0)\right) \xrightarrow{T} S^{-1}(0) \cap T(H_1) \to 0$$

can be written as

$$0 \to \operatorname{Ker} T \hookrightarrow \operatorname{Ker}(ST) \xrightarrow{T} \operatorname{Ker}(S) \cap T(H_1) \to 0,$$

from which we get

$$\begin{aligned}\dim \operatorname{Ker}(ST) &= \dim \operatorname{Ker} T + \dim(\operatorname{Ker} S \cap T(H_1)) \\ &= \dim \operatorname{Ker} T + \dim\left(\operatorname{Ker} S \cap (\operatorname{Ker} T^*)^{\perp}\right) \\ &= \dim \operatorname{Ker} T + \dim \operatorname{Ker} S - \dim(\operatorname{Ker} S \cap \operatorname{Ker} T^*).\end{aligned}$$

Subtracting result (obtained from the above by switching S and T^*)

$$\dim \operatorname{Ker}\left((ST)^*\right) = \dim \operatorname{Ker} S^* + \dim \operatorname{Ker} T^* - \dim(\operatorname{Ker} T^* \cap \operatorname{Ker} S),$$

we get the desired result. □

Theorem 2.11 *$\mathcal{F}(H_1, H_2)$ is open in $\mathcal{B}(H_1, H_2)$ and the index function* index $:$ *$\mathcal{F}(H_1, H_2) \to \mathbb{Z}$ is constant on each connected component of $\mathcal{F}(H_1, H_2)$ and surjective.*

Proof For $F_0 \in \mathcal{F}(H_1, H_2)$, let

$$H_3 := H_1 \oplus F_0(H_1)^{\perp} = H_1 \oplus \operatorname{Ker} F_0^* \ \text{ and } \ H_4 := H_2 \oplus \operatorname{Ker} F_0,$$

For $A \in \mathcal{B}(H_1, H_2)$, define

$$\overline{A} : H_3 \to H_4 \text{ by } \overline{A}(v, w) = (Av - w, \pi_{\operatorname{Ker} F_0} v).$$

Note that $\overline{A}$ is bounded, since $\left\|\overline{A}\right\|^2 \le \|A\|^2 + 1$:

$$\begin{aligned}\left|\overline{A}(v, w)\right|^2 &= |Av - w|^2 + |\pi_{\operatorname{Ker} F_0} v|^2 \le |Av|^2 + |w|^2 + |v|^2 \\ &\le \left(\|A\|^2 + 1\right)\left(|v|^2 + |w|^2\right) = \left(\|A\|^2 + 1\right)\|(v, w)\|^2.\end{aligned}$$

For the obvious inclusion $i : H_1 \to H_1 \oplus \operatorname{Ker} F_0^*$ and projection $\pi : H_2 \oplus \operatorname{Ker} F_0 \to H_2$, we have $A = \pi \circ \overline{A} \circ i$. For $A, B \in \mathcal{B}(H_1, H_2)$ we have $\left\|\overline{A} - \overline{B}\right\| \le \|A - B\|$, since

$$\begin{aligned}\left|\left(\overline{A} - \overline{B}\right)(v, w)\right| &= |(Av - w, \pi_{\operatorname{Ker} F_0} v) - (Bv - w, \pi_{\operatorname{Ker} F_0} v)| \\ &= |(A - B)v| \le \|A - B\|\,|v| \le \|A - B\|\,|(v, w)|.\end{aligned}$$

Observe that $\overline{F}_0 \in \mathrm{Iso}(H_1, H_2) := \{C \in \mathcal{B}(H_1, H_2) : C^{-1} \in \mathcal{B}(H_2, H_1)\}$. Indeed, $\mathrm{Ker}(\overline{F_0}) = \{(0,0)\}$, since

$$\overline{F_0}(v,w) = (F_0 v - w, \pi_{\mathrm{Ker}\, F_0} v) = (0,0)$$
$$\Rightarrow\ F_0 v = w \in F_0(H_1)^{\perp} \text{ and } v \in (\mathrm{Ker}\, F_0)^{\perp} \Rightarrow (v,w) = (0,0)\,,$$

Also $\overline{F_0}$ is onto, since given $(x,y) \in H_2 \oplus \mathrm{Ker}\, F_0$, we have $x = F_0(u) + w$, where $u \in (\mathrm{Ker}\, F_0)^{\perp}$ and $w \in F_0(H_1)^{\perp}$, whence

$$\overline{F_0}(u+y, -w) = (F_0(u+y) + w, \pi_{\mathrm{Ker}\, F_0}(u+y)) = (x,y).$$

Then $\overline{F}_0 \in \mathrm{Iso}(H_1, H_2)$ by the Open Mapping Theorem. Since $\|\overline{A} - \overline{B}\| \le \|A - B\|$, we have $\|\overline{F} - \overline{F}_0\| < \|\overline{F}_0\|$ for $\|F - F_0\| < \|\overline{F}_0\|$, in which case $\overline{F} \in \mathrm{Iso}(H_3, H_4)$. This implies that $F = \pi \circ \overline{F} \circ i \in \mathcal{F}(H_1, H_2)$, since each of $\pi, \overline{F}$ and i are Fredholm. Moreover,

$$\begin{aligned}\mathrm{index}\, F &= \mathrm{index}\, \pi + \mathrm{index}\, \overline{F} + \mathrm{index}\, i \\ &= \dim \mathrm{Ker}\, F_0 + 0 - \dim \mathrm{Ker}\, F_0^*(H_1) = \mathrm{index}\, F_0,\end{aligned}$$

which shows that $\mathrm{index} : \mathcal{F}(H_1, H_2) \to \mathbb{Z}$ is locally constant, and hence constant on each connected component of $\mathcal{F}(H_1, H_2)$. By composing F_0 with shift operators on H_1 or H_2, we see that $\mathrm{index} : \mathcal{F}(H_1, H_2) \to \mathbb{Z}$ is surjective. □

Corollary 2.12 *If $F \in \mathcal{F}(H_1, H_2)$ and $K \in \mathcal{K}(H_1, H_2)$, then $F + K \in \mathcal{F}(H_1, H_2)$ and* $\mathrm{index}\,(F+K) = \mathrm{index}\,(F)$. *Thus, if $F \in \mathrm{Iso}(H_1, H_2)$, then* $\mathrm{index}\,(F+K) = 0$.

Proof Note that $t \mapsto F + tK$, $t \in [0,1]$ connects F with $F+K$ in $\mathcal{F}(H_1, H_2)$ by Theorem 2.9, and then $\mathrm{index}\,(F+K) = \mathrm{index}\,(F)$ by Theorem 2.11. □

Corollary 2.13 *About every $F_0 \in \mathcal{F}(H_1, H_2)$, there is an open neighborhood $U_0 \subseteq \mathcal{B}(H_1, H_2)$, such that $F \in U_0$ implies*

- (i) $(\mathrm{Ker}\, F_0)^{\perp} \cap \mathrm{Ker}\, F = \{0\}\,,$
- (ii) $F((\mathrm{Ker}\, F_0)^{\perp}) \oplus F_0(H_1)^{\perp} = H_2$, *and hence* $\frac{H_2}{F((\mathrm{Ker}\, F_0)^{\perp})} \cong F_0(H_1)^{\perp}$, *and*
- (iii) *for any $W \subseteq \mathrm{Ker}\, F_0$ of finite codimension, we have* $F(W) \oplus F_0(W)^{\perp} = H_2$, *and hence* $\frac{H_2}{F(W)} \cong F_0(W)^{\perp}$.

Proof Recall from the proof of Theorem 2.11 that for $F \in \mathcal{B}(H_1, H_2)$, $\overline{F} : H_1 \oplus F_0(H_1)^{\perp} \to H_2 \oplus \mathrm{Ker}\, F_0$ is defined by

$$\overline{F}(v,w) = (Fv - w, \pi_{\mathrm{Ker}\, F_0} v).$$

Moreover, for $F \in U_0 := \{G \in \mathcal{B}(H_1, H_2) : \|G - F_0\| < \|\overline{F}_0\|\}$ we have shown that $\overline{F} \in \mathrm{Iso}(H_3, H_4)$. Note that the injectivity of $\overline{F}$ yields (i), since

$$v \in (\mathrm{Ker}\, F_0)^{\perp} \cap \mathrm{Ker}\, F \Rightarrow \overline{F}(v,0) = (Fv, \pi_{\mathrm{Ker}\, F_0} v) = (0,0) \Rightarrow v = 0.$$

For (ii), it remains to show that

$$\overline{F} \in \mathrm{Iso}(H_3, H_4) \Rightarrow F((\mathrm{Ker}\, F_0)^{\perp}) \oplus F_0(H_1)^{\perp} = H_2.$$

First we show that $F((\operatorname{Ker} F_0)^{\perp}) \cap F_0(H_1)^{\perp} = \{0\}$. Indeed, suppose that $F(v) = w$ for some $v \in (\operatorname{Ker} F_0)^{\perp}$ and some $w \in F_0(H_1)^{\perp}$. Then $\overline{F}(v, w) = (Fv - w, \pi_{\operatorname{Ker} F_0} v) = (0, 0)$ in which case $(v, w) = (0, 0)$ by the injectivity of $\overline{F}$. Moreover, since $\overline{F}$ is onto, for any $(u, f_0) \in H_2 \oplus \operatorname{Ker} F_0$, there is $(v, w) \in H_1 \oplus F_0(H_1)^{\perp}$, with $u = Fv + w$ and $\pi_{\operatorname{Ker} F_0} v = f_0$. Note that $v - f_0 \in (\operatorname{Ker} F_0)^{\perp}$, and $u = Fv + w = F(v - f_0) + w$. Thus, $H_2 = F((\operatorname{Ker} F_0)^{\perp}) + F_0(H_1)^{\perp}$. Combining the above yields $F((\operatorname{Ker} F_0)^{\perp}) \oplus F_0(H_1)^{\perp} = H_2$. For (iii), we first show that $F(W) \cap F_0(W)^{\perp} = \{0\}$. Indeed, suppose that $F(v) = w$ for some $v \in W$ and some $w \in F_0(W)^{\perp}$. Then $\overline{F}(v, w) = (Fv - w, \pi_{\operatorname{Ker} F_0} v) = (0, 0)$ and so $(v, w) = (0, 0)$. Now

$$F(W) \oplus F_0(W)^{\perp} = H_2 = F(W) \oplus F(W)^{\perp},$$

provided that $\dim F_0(W)^{\perp} = \dim F(W)^{\perp}$ (i.e., $\operatorname{codim} F_0(W) = \operatorname{codim} F(W)$). Note that

$$\begin{aligned}
&F((\operatorname{Ker} F_0)^{\perp}) \oplus F_0(H_1)^{\perp} = H_2 \Rightarrow \\
&\operatorname{codim} F_0((\operatorname{Ker} F_0)^{\perp}) = \dim F_0(H_1)^{\perp} = \operatorname{codim} F\left((\operatorname{Ker} F_0)^{\perp}\right), \text{ and so} \\
&\operatorname{codim} F_0(W) = \dim \frac{F_0((\operatorname{Ker} F_0)^{\perp})}{F_0(W)} + \operatorname{codim} F_0((\operatorname{Ker} F_0)^{\perp}) \\
&= \dim \frac{F((\operatorname{Ker} F_0)^{\perp})}{F(W)} + \operatorname{codim} F((\operatorname{Ker} F_0)^{\perp})) = \operatorname{codim} F(W).
\end{aligned}$$

□

Proposition 2.14 *The group* $\operatorname{Iso}(H)$ *of invertible bounded linear operators on a Hilbert space* H *is pathwise connected.*

Proof Each $T \in \operatorname{Iso}(H)$ can be factored as $T = UB$ where U is unitary, and $B = \sqrt{T^*T}$ is in the convex set of self-adjoint, positive elements of $\operatorname{Iso}(H)$. Using the Spectral Theorem, U can be written in the form $U = e^{iA}$ where A is a self-adjoint operator. Let

$$t \mapsto U_t := e^{itA} \text{ and } B_t := (1 - t)\operatorname{Id} + tB,\ t \in [0, 1].$$

Then $t \mapsto U_t B_t$ is a continuous path in $\operatorname{Iso}(H)$ from Id to T. □

Let $\pi_0(\mathcal{F}(H_1, H_2))$ denote the connected components of $\mathcal{F}(H_1, H_2)$. Since an open subset of a locally path-connected space is connected if and only if it is path-connected, the connected components of $\mathcal{F}(H_1, H_2)$ are the same as the path components. In the case $H_1 = H_2 =: H$, composition in $\mathcal{F}(H) := \mathcal{F}(H, H)$ induces a well-defined group operation on $\pi_0(\mathcal{F}(H))$. Namely, if $[F] \in \pi_0(\mathcal{F}(H))$ denotes the component of F in $\mathcal{F}(H)$, then $[F] = [F']$ and $[G] = [G']$ imply $[FG] = [F'G']$. The identity is $[I]$ and the inverse of $[F]$ is $[G]$, where G satisfies $FG = I + K$ for $K \in \mathcal{K}(H) = \mathcal{K}(H, H)$. Note that $[I + K] = [I]$ since $I + tK$ connects I to $I + K$ in $\mathcal{F}(H)$. Since

$$\operatorname{index}(FG) = \operatorname{index} F + \operatorname{index} G = \operatorname{index} F' + \operatorname{index} G' = \operatorname{index}(F'G'),$$

index induces a well-defined homomorphism $\operatorname{Index} : \pi_0(\mathcal{F}(H)) \to \mathbb{Z}$.

Theorem 2.15 $\operatorname{Index} : \pi_0(\mathcal{F}(H)) \to \mathbb{Z}$ *is a group isomorphism.*

Proof By Example 2.2, we have surjectivity of Index. For injectivity, it remains to show that there is a path in $\mathcal{F}(H)$ from any $F \in \mathcal{F}(H)$ with $\operatorname{index} F = 0$ to some operator in $\operatorname{Iso}(H)$. Since $\operatorname{index} F = 0$, we can find an isomorphism $L : \operatorname{Ker} F \to F(H)^{\perp}$, and let $\pi : H \to \operatorname{Ker} F$ be orthogonal projection. Then $t \mapsto F + t\,(L \circ \pi)$ (for $t \in [0,1]$) connects F with $F + (L \circ \pi)$ which is in $\operatorname{Iso}(H)$ since $L \circ \pi$ has finite rank. Indeed, for $v \in \operatorname{Ker} F$ and $w \in (\operatorname{Ker} F)^{\perp}$, we have

$$(F + (L \circ \pi))\,(v + w) = F(w) + L(v) = 0 \Rightarrow F(w) = -L(v) \in F(H) \cap F(H)^{\perp}$$
$$\Rightarrow F(w) = L(v) = 0 \Rightarrow w = v = 0, \text{ and}$$

$$(F + (L \circ \pi))\,(H) = F(H) + (L \circ \pi)\,(H) = F(H) + F(H)^{\perp} = H.$$

□

3 The space of Fredholm operators and K-theory

There is a far-reaching generalization of Theorem 2.15, namely the Atiyah-Jänich Theorem which exhibits the close relation of Fredholm operators, index theory and K-theory. This Theorem is stated and finally proved in subsection 3.4, after some preliminary subsections. We start with a brief introduction to K-theory.

3.1 An introduction to K-theory

Let X be a compact topological space. We denote the set of all isomorphism classes of complex vector bundles $E \to X$ by $\operatorname{Vect}(X)$. Note that $\operatorname{Vect}(X)$ has an abelian semigroup structure $\dot{+} : \operatorname{Vect}(X) \times \operatorname{Vect}(X) \to \operatorname{Vect}(X)$ induced by direct sum of vector bundles, namely $[E] \dot{+} [F] := [E \oplus F]$. The class of the zero vector bundle is an identity, and so $\operatorname{Vect}(X)$ is an abelian monoid (commutative semigroup with identity). There is a standard way (known as the Grothendieck construction) to produce an abelian group from an abelian monoid. In the case of $\operatorname{Vect}(X)$, the abelian group is $K(X)$. Explicitly, $K(X)$ consists of equivalence classes of pairs $([E]\,,[F])$, where

$$([E]\,,[F]) \equiv ([E']\,,[F']) \Leftrightarrow ([E] \dot{+} [G]\,,[F] \dot{+} [G]) = ([E'] \dot{+} [G']\,,[F'] \dot{+} [G'])\,,$$

for some $[G]\,,[G'] \in \operatorname{Vect}(X)$. In other words, $K(X)$ is the quotient semigroup of $\operatorname{Vect}(X) \times \operatorname{Vect}(X)$ by its diagonal:

$$K(X) = \frac{\operatorname{Vect}(X) \times \operatorname{Vect}(X)}{\Delta\left(\operatorname{Vect}(X) \times \operatorname{Vect}(X)\right)}.$$

Then $K(X)$ is a group since $[([E]\,,[F])]$ and $[([F]\,,[E])]$ are additive inverses:

$$[([E]\,,[F])] + [([F]\,,[E])] = [([E] \dot{+} [F]\,,[F] \dot{+} [E])] = [([E \oplus F]\,,[E \oplus F])] \equiv [(0,0)]\,.$$

If we apply this construction to the semigroup $\mathbb{Z}^{+} = \{0,1,2,\ldots\}$, we obtain $\frac{\mathbb{Z}^{+} \times \mathbb{Z}^{+}}{\Delta(\mathbb{Z}^{+} \times \mathbb{Z}^{+})}$ which is isomorphic to $\mathbb{Z}$ via $[(m,n)] \mapsto m - n$ and the natural embedding $\mathbb{Z}^{+} \to \mathbb{Z}$ factors through the semigroup monomorphism $\mathbb{Z}^{+} \to \frac{\mathbb{Z}^{+} \times \mathbb{Z}^{+}}{\Delta(\mathbb{Z}^{+} \times \mathbb{Z}^{+})}$ given by $m \mapsto [(m,0)]$.

However, in general the homomorphism $\mathrm{Vect}(X) \to K(X)$, given by $[E] \mapsto [([E], [0])]$ is not injective. For example, since

$$\pi_4(GL(2,\mathbb{C})) \cong \pi_4(U(2)) \cong \pi_4(S^3 \times S^1) \cong \pi_4(S^3) \cong \mathbb{Z}_2,$$

there is a nontrivial two-dimensional complex vector bundle E over S^5, but $K(S^5) = \mathbb{Z}$, which implies that $[E] = [X \times \mathbb{C}^2]$ in $K(S^5)$, even though E is not trivial. Nevertheless, in general it is customary to write $[([E], [0])] \in K(X)$ simply as $[E]$, even though $[E] = [F]$ in $K(X)$ may not imply that E and F are equivalent vector bundles. Since

$$[([E], [F])] = [([E], [0])] + [([0], [F])] = [([E], [0])] + (-[([F], [0])]),$$

it is convenient to use the notation $[E] - [F] := [([E], [F])]$, and to refer to elements of $K(X)$ as *virtual bundles*. If we assume that X is not only compact, but also Hausdorff, then for any complex vector bundle $\pi : F \to X$, there is another F', such that $[F \oplus F'] = [n]$, where $[n]$ denotes the class of the trivial bundle $\mathbb{C}^n \times X \to X$ and $n = \dim F + \dim F'$. To show this, we may cover X by a finite number of open sets $U_1, \ldots, U_k$, such that $\varphi_i : F_{|U_i} \cong U_i \times \mathbb{C}^m$ ($m = \dim F$). Then (owing to the Hausdorff assumption), there is a partition of unity $\{\rho_i : X \to [0,1]\}_{i=1}^m$ subordinate to $\{U_1, \ldots, U_k\}$, so that (for $\pi_i : U_i \times \mathbb{C}^m \to \mathbb{C}^m$ the projection onto the second factor) we have an injective vector bundle morphism

$$\Phi : \pi \times (\pi_1\varphi_1\rho_1 \oplus \cdots \oplus \pi_k\varphi_k\rho_k) : F \to X \times (\mathbb{C}^m \oplus \cdots \oplus \mathbb{C}^m) = X \times \mathbb{C}^{km}.$$

We then take F' to be the subbundle of $X \times \mathbb{C}^{km}$ whose fiber F'_x at any point $x \in X$ is the orthogonal complement of $\Phi(F_x)$, relative to a Hermitian metric on $\mathbb{C}^{km}$. As a consequence, any $[E] - [F] \in K(X)$ can be written in the form $[H] - [n]$ for some $n \geq 0$, since

$$[E] - [F] = [E] + [F'] - ([F] + [F']) = [E \oplus F'] - [F \oplus F'] = [E \oplus F'] - [n].$$

Moreover, while by definition $[E] = [F]$ in $K(X) \Leftrightarrow E \oplus G \cong F \oplus G$ for some bundle G, note that

$$E \oplus G \cong F \oplus G \Rightarrow E \oplus G \oplus G' \cong F \oplus G \oplus G' \Rightarrow E \oplus \left(X \times \mathbb{C}^N\right) \cong F \oplus \left(X \times \mathbb{C}^N\right),$$

where (as we have just seen) G' can be so that $G \oplus G' \cong X \times \mathbb{C}^N$. Thus, $[E] = [F]$ in $K(X) \Leftrightarrow E \oplus G \cong F \oplus G$ for some trivial bundle G; i.e., $[E] = [F]$ iff E and F are *stably equivalent*.

If $E \to X$ and $F \to X$ are complex vector bundles, then the tensor product $E \otimes F := \cup_x (E_x \otimes F_x) \to X$ is a complex vector bundle and this induces a multiplication on $K(X)$ making it a ring. For a base point $x_0 \in X$, the inclusion $i : \{x_0\} \to X$ induces $i^* : K(X) \to K(\{x_0\}) \cong \mathbb{Z}$. The *reduced* K-ring of X is $\widetilde{K}(X) := \mathrm{Ker}\, i^* \subset K(X)$. For a closed nonvoid subset $Y \subset X$, we define $K(X,Y) = \widetilde{K}(X/Y)$, where Y/Y is the base point. For a locally compact space X, one defines $K(X) := \widetilde{K}(X^+)$, where X^+ is the one-point compactification of X and the point at infinity is the base point; if X is already compact, then X^+ is just X with a disjoint base point, so that $\widetilde{K}(X^+) = K(X)$.

3.2 The index bundle

Let X be a compact, Hausdorff topological space, and let $[X, \mathcal{F}(H)]$ denote the set of homotopy classes of continuous maps $X \to \mathcal{F}(H)$. For two such maps f and g we define a product $fg : X \to \mathcal{F}(H)$ via $(fg)(x) = f(x) \circ g(x)$, and this induces a product on $[X, \mathcal{F}(H)]$, giving it the structure of an associative semigroup with identity $[\mathrm{Id}]$ where $\mathrm{Id} : X \to \{\mathrm{Id}\} \subset \mathcal{F}(H)$ is the constant map. We will show that there is a natural homomorphism

$$\mathrm{Index} : [X, \mathcal{F}(H)] \to K(X).$$

The fact that this is an isomorphism is the Atiyah-Jänich Theorem (see subsection 3.4 below). In the case where $X = \{p\}$ is a singleton, $[X, \mathcal{F}(H)]$ is simply $\pi_0(\mathcal{F}(H))$, $K(X) = K(\{p\}) \cong \mathbb{Z}$, and the map $\mathrm{Index} : [X, \mathcal{F}(H)] \to K(X) \cong \mathbb{Z}$, induced by $f \mapsto \mathrm{Index}(f(p))$, is an isomorphism by Theorem 2.15.

Let A be a compact subset of $\mathcal{F}(H)$. We will show that there is a canonical element $i_A \in K(A)$. If $\mathrm{Ker}_A := \cup_{F \in A} \{\mathrm{Ker}\, F\}$ and $\mathrm{Coker}_A := \cup_{F \in A} \{\mathrm{Coker}\, F\}$ were bundles, then we could define $i_A = [\mathrm{Ker}_A] - [\mathrm{Coker}_A]$. While Ker_A and Coker_A are not bundles, recall that $F \mapsto \dim \mathrm{Ker}\, F - \dim \mathrm{Coker}\, F$ is locally constant, which offers some hope that i_A can be defined. We cover compact A by a finite number of neighborhoods $U_1, \ldots, U_k$ as in Corollary 2.13 about $F_1, \ldots, F_k \in A$. Let V be a (closed) subspace of H of finite codimension such that for all $F \in A$, we have

(1) $V \cap \mathrm{Ker}\, F = \{0\}$

(2) $\mathrm{codim}\,(F(V)) = \dim\,(F(V))^\perp < \infty$ (and so $F(V)$ is closed as well)

(3) $\frac{H}{A(V)} := \cup_{F \in A} \left(F, \frac{H}{F(V)}\right) \to A$ is a vector bundle, where $\frac{H}{A(V)}$ gets the quotient topology from $\mathcal{F}(H) \times H$.

We say that V *is suitable for* A. An example of such V is $V_0 := (\mathrm{Ker}\, F_1)^\perp \cap \cdots \cap (\mathrm{Ker}\, F_k)^\perp$. That (1) and (2) are satisfied for $V = V_0$ is clear from Corollary 2.13. For (3), we use part (iii) of Corollary 2.13 with $W = V_0$ to get that $\frac{H}{F(V_0)} \cong F_i(V_0)^\perp$ for $F \in U_i$. Thus, over $U_i \cap A$, each $\frac{H}{F(V_0)}$ is identified with the vector space $F_i(V_0)^\perp$ independent of F, providing a suitable trivialization of $\frac{H}{A(V)}$ over $U_i \cap A$. Note that if V is suitable for A and $V' \subseteq V$ has finite codimension, then V' is also suitable for A. We set

$$i_A(V) = \left[A \times \frac{H}{V}\right] - \left[\frac{H}{A(V)}\right],$$

where $A \times \frac{H}{V}$ is the trivial bundle with fiber $\frac{H}{V}$. We need to show that $i_A(V)$ is independent of the choice of suitable V for A. If V_1 and V_2 are suitable for A, then we have exact sequences of vector bundles

$$0 \to A \times \frac{V_1}{V_1 \cap V_2} \to A \times \frac{H}{V_1 \cap V_2} \to A \times \frac{H}{V_1} \to 0 \text{ and}$$
$$0 \to A \times \frac{V_1}{V_1 \cap V_2} \to \frac{H}{A\,(V_1 \cap V_2)} \to \frac{H}{A\,(V_1)} \to 0.$$

Thus,

$$\left[A \times \frac{V_1}{V_1 \cap V_2}\right] = \left[A \times \frac{H}{V_1 \cap V_2}\right] - \left[A \times \frac{H}{V_1}\right],$$
$$\left[A \times \frac{V_1}{V_1 \cap V_2}\right] = \left[\frac{H}{A\,(V_1 \cap V_2)}\right] - \left[\frac{H}{A\,(V_1)}\right]$$

and

$$i_A(V_1) = \left[A \times \frac{H}{V_1}\right] - \left[\frac{H}{A\,(V_1)}\right] = \left[A \times \frac{H}{V_1 \cap V_2}\right] - \left[\frac{H}{A\,(V_1 \cap V_2)}\right] = i_A\,(V_1 \cap V_2)\,,$$

which is the same as $i_A(V_2)$ by symmetry, as required.

For continuous $f : X \to \mathcal{F}(H)$, we define

$$\mathrm{Index} : [X, \mathcal{F}(H)] \to K(X) \quad \text{by} \quad \mathrm{Index}\,([f]) = f^*(i_A),$$

where A is any compact set containing $f(X)$ and $f^* : K(A) \to K(X)$ is induced by pull-back of bundles. More directly we have

$$\mathrm{Index}\,([f]) = \left[X \times \frac{H}{V}\right] - \left[\frac{H}{f(V)}\right], \text{ where } \frac{H}{f(V)} := \cup_{x \in X} \left(x, \frac{H}{f(x)(V)}\right) \to X$$

which is a vector bundle. We show that $f^*(i_A)$ only depends on the homotopy class of f. If f_0 is homotopic to f_1, say there is $h : X \times [0,1] \to \mathcal{F}(H)$ with $h(x,0) = f_0(x)$ and $h(x,1) = f_1(x)$, then for $i_t : X \to X \times [0,1]$ given by $i_t(x) = (x,t)$ we have (where A is now any compact subset of $\mathcal{F}(H)$ containing $h(X \times [0,1])$)

$$f_0^*(i_A) = (h \circ i_0)^*\,(i_A) = i_0^* h^*(i_A) = i_1^* h^*(i_A) = (h \circ i_1)^*\,(i_A) = f_1^*(i_A).$$

To show that $\mathrm{Index} : [X, \mathcal{F}(H)] \to K(X)$ is a homomorphism, let $f, g : X \to \mathcal{F}(H)$ be continuous. Let V_f be suitable for $f(X)$. Let $\pi_{V_f}, \pi_{V_f^\perp} \in \mathcal{B}(H)$ denote orthogonal projections onto V_f and $(V_f)^\perp$. Note that $I - t\pi_{V_f^\perp} \in \mathcal{F}(H)$ for each $t \in [0,1]$, and hence we have a homotopy $h : X \times [0,1] \to \mathcal{F}(H)$ between g and $\pi_{V_f} \circ g$, given by $(I - t\pi_{V_f^\perp}) \circ g(x) : X \times [0,1] \to \mathcal{F}(H)$. Hence we may assume that $g(x)(H) \subseteq V_f$ for all $x \in X$. Let V_g be suitable for $g(X)$. Then V_g is also suitable for the set $(fg)\,(X) := \{f(x) \circ g(x) \mid x \in X\}$. Indeed, $\mathrm{Ker}\,(f(x) \circ g(x)) \subseteq \mathrm{Ker}\, g(x)$, since (using $g(x)(H) \subseteq V_f$)

$$\begin{aligned} &v \in \mathrm{Ker}\,(f(x) \circ g(x)) \Rightarrow 0 = (f(x) \circ g(x))\,(v) = f(x)(g(x)(v)) \\ &\Rightarrow g(x)(v) \in \mathrm{Ker}\, f(x) \cap g(x)(H) \subseteq \mathrm{Ker}\, f(x) \cap V_f = \{0\} \\ &\Rightarrow v \in \mathrm{Ker}\, g(x). \end{aligned}$$

Thus, $V_g \cap \mathrm{Ker}\,(f(x) \circ g(x)) \subseteq V_g \cap \mathrm{Ker}\, g(x) = \{0\}$. Note that $g(x)(V_g) \subseteq V_f$ has finite codimension and $f(x)\,(V_f)$ has finite codimension in H. Hence, $\mathrm{codim}\,((f(x) \circ g(x))\,(V_g)) < \infty$. Since $g(x)(V_g) \subseteq V_f$ and

$$\frac{H}{f(V_f)} := \cup_{x \in X} \left(x, \frac{H}{f(x)\,(V_f)}\right) \to X \text{ is a vector bundle,}$$

we have that

$$\frac{H}{(fg)\,(V)} := \cup_{x\in X}\left(x, \frac{H}{f(x)\,(g(x)(V_g))}\right) \to X \text{ is a vector bundle.}$$

Again using $g(V_g) \subseteq V_f$, we have exact sequences

$$0 \to \frac{V_f}{g(V_g)} \xrightarrow{f} \frac{H}{(fg)\,(V_g)} \to \frac{H}{f(V_f)} \to 0$$
$$\Rightarrow \left[\frac{H}{(fg)\,(V_g)}\right] = \left[\frac{V_f}{g(V_g)}\right] + \left[\frac{H}{f(V_f)}\right],$$

and

$$0 \to \frac{H}{V_f} \to \frac{H}{g(V_g)} \to \frac{V_f}{g(V_g)} \to 0$$
$$\Rightarrow \left[\frac{H}{g(V_g)}\right] = \left[\frac{H}{V_f}\right] + \left[\frac{V_f}{g(V_g)}\right] \Rightarrow \left[\frac{V_f}{g(V_g)}\right] = \left[\frac{H}{g(V_g)}\right] - \left[\frac{H}{V_f}\right].$$

Thus, recalling that V_g is suitable for $fg(X)$,

$$\mathrm{Index}\,([fg]) = \left[\frac{H}{V_g}\right] - \left[\frac{H}{(fg)\,(V_g)}\right] = \left[\frac{H}{V_g}\right] - \left(\left[\frac{V_f}{g(V_g)}\right] + \left[\frac{H}{f(V_f)}\right]\right)$$
$$= \left[\frac{H}{V_g}\right] - \left(\left(\left[\frac{H}{g(V_g)}\right] - \left[\frac{H}{V_f}\right]\right) + \left[\frac{H}{f(V_f)}\right]\right)$$
$$= \left[\frac{H}{V_f}\right] - \left[\frac{H}{f(V_f)}\right] + \left[\frac{H}{V_g}\right] - \left[\frac{H}{g(V_g)}\right] = \mathrm{Index}\,([f]) + \mathrm{Index}\,([g]).$$

The proof that $\mathrm{Index} : [X, \mathcal{F}(H)] \to K(X)$ is an isomorphism (i.e., The Atiyah-Jänich Theorem) consists of showing that it is injective and surjective. Injectivity rests on Kuiper's Theorem which is also of independent interest, and we cover that next.

3.3 Kuiper's Theorem

For finite n, the group $\mathrm{GL}(n, \mathbb{C})$ is not simply-connected. Indeed, $\mathrm{GL}(1)$ is $\mathbb{C} \setminus \{0\}$ and $\det : \mathrm{GL}(n, \mathbb{C}) \to \mathrm{GL}(1)$ induces an isomorphism $\pi_1(\mathrm{GL}(n, \mathbb{C})) \cong \pi_1(\mathrm{GL}(1)) \cong \mathbb{Z}$. However, in infinite dimensional spaces one can generally escape finite-dimensional topological constraints by moving aside into a new dimension. In particular, it was shown by N. I. Kuiper [33] that $\mathrm{Iso}(H)$, with the operator norm topology, is contractible. For our purposes, it suffices to prove Theorem 3.2 below, but first we establish

Proposition 3.1 *For $R, S : X \to \mathcal{B}(H)$, the two maps $SR \oplus I$, $R \oplus S : X \to \mathcal{B}(H \times H)$ are homotopic (i.e., $SR \oplus I \sim R \oplus S$).*

Proof Writing operators in $\mathcal{B}(H \times H)$ as 2×2 (block) matrices, define $F : X \times [0, \pi/2] \to \mathcal{B}(H \times H)$ by

$$F(x,t) := \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix} \begin{bmatrix} S(x) & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} \begin{bmatrix} R(x) & 0 \\ 0 & I \end{bmatrix},$$

where $\cos t$ and $\sin t$ stand for the operators $(\cos t)\, I$ and $(\sin t)\, I \in \mathcal{B}(H)$. Note that

$$F(x,0) = \begin{bmatrix} S(x)R(x) & 0 \\ 0 & \mathrm{Id} \end{bmatrix}, \text{ while } F(x,\pi/2) = \begin{bmatrix} R(x) & 0 \\ 0 & S(x) \end{bmatrix}.$$

Moreover, if S and R have values in $\mathrm{U}(H)$ (or $\mathrm{Iso}(H)$), then so does F. The continuity F in the norm topology of $\mathcal{B}(H \times H)$ is straightforward (but a bit tedious) to check. □

Theorem 3.2 (Kuiper) *Let X be compact and let $f : X \to \mathrm{Iso}(H)$ be continuous. Then f is homotopic in $\mathrm{Iso}(H)$ to the constant map $c : X \to \{I\}$.*

Proof Suppose that $H = H_1 \oplus H'$, where H' and H_1 are infinite-dimensional closed orthogonal complements. Let

$$Y = \left\{ A \in \mathrm{Iso}(H) : A_{|H'} = I_{H'} \text{ and } A\left(H_1\right) = H_1 \right\}.$$

We provide a homotopy $h : Y \times I \to Y$ with $h_0(A) = A$ and $h_1(A) = I$; i.e., Y is contractible in itself to I. Once this is done, it suffices to show that $f : X \to \mathrm{Iso}(H)$ is homotopic in $\mathrm{Iso}(H)$ to a map with values in Y relative some decomposition of the form $H = H_1 \oplus H'$. To show that Y is contractible, note that there is a decomposition $H' = H_2 \oplus H_3 \oplus \cdots$ of H' into an infinite number of infinite-dimensional closed, orthogonal subspaces, each of which can be identified with a copy of H_1. For $A \in Y$, let $B := A_{|H_1} \in \mathrm{Iso}(H_1)$. In block diagonal form, relative to the decomposition $H = H_1 \oplus H_2 \oplus H_3 \oplus \cdots = H_1 \oplus H_1 \oplus H_1 \oplus \cdots$, we have $A = B \oplus I \oplus I \oplus I \oplus \cdots$. By Proposition 3.1, we have

$$B \oplus B^{-1} \sim B^{-1}B \oplus I = I \oplus I = BB^{-1} \oplus I \sim B^{-1} \oplus B.$$

Using $I \oplus I \sim B^{-1} \oplus B$ and then using $B \oplus B^{-1} \sim I \oplus I$, we obtain

$$A = B \oplus I \oplus I \oplus I \oplus \cdots \sim B \oplus B^{-1} \oplus B \oplus B^{-1} \oplus \cdots) \sim I \oplus I \oplus I \oplus I \oplus \cdots = I_H.$$

To show that $f : X \to \mathrm{Iso}(H)$ is homotopic in $\mathrm{Iso}(H)$ to a map $X \to Y$, we first show that f is homotopic in $\mathrm{Iso}(H)$ to a map $f_1 : X \to \mathrm{Iso}(H)$ with values in the intersection of a finite-dimensional subspace V of $\mathcal{B}(H)$ with $\mathrm{Iso}(H)$. Note that the distance from the compact set $f(X)$ to the closed complement of $\mathrm{Iso}(H)$ in $\mathcal{B}(H)$ is positive, say 2ε. We cover the compact set $f(X)$ with a finite number of balls $B(f(x_i),\varepsilon) = \{A \in \mathcal{B}(H) : \|A - f(x_i)\| < \varepsilon\}$, $i = 1,\ldots,N$. For $A \in U := \bigcup_{i=1}^N B(f(x_i),\varepsilon)$ and $i = 1,\ldots,N$, define $\phi_i : U \to (0,1]$ by

$$\phi_i\,(A) := \frac{\psi_i(A)}{\sum_{i=1}^N \psi_i(A)}, \text{ where } \psi_i(A) := \begin{cases} \varepsilon - \|A - f(x_i)\| & \text{for } A \in B\left(f(x_i),\varepsilon\right) \\ 0 & \text{otherwise.} \end{cases}$$

For each $t \in [0,1]$ and $A \in U$, we define an operator

$$g_t(A) := (1-t)A + t\sum\nolimits_{i=1}^N \phi_i(A) f(x_i).$$

Note that $g_0 : U \to \mathcal{B}^\times(H)$ is the inclusion and $g_1 : U \to \mathrm{Iso}(H)$ maps U into the convex hull of the points $f(x_1), \ldots, f(x_N)$. (We thank the concerned reader in advance for verifying the joint continuity of $g_t(A)$ in (t, A), as well as for other homotopies defined later.) To show that $g_t(U) \subset \mathrm{Iso}(H)$, note that $A \in B(f(x_{i_0}), \varepsilon)$ for some i_0, and for $t \in [0,1]$

$$\|g_t(A) - f(x_{i_0})\| \leq \|g_t(A) - A\| + \|A - f(x_{i_0})\| < \varepsilon + \varepsilon, \text{ since}$$

$$\begin{aligned}\|g_t(A) - A\| &\leq \|t \sum\nolimits_{i=1}^{N} \phi_i(A) f(x_i) - tA\| = t\| \sum\nolimits_{i=1}^{N} \phi_i(A) \left(f(x_i) - A\right) \| \\ &\leq \sum\nolimits_{i=1}^{N} \phi_i(A) \left\|f(x_i) - A\right\| \leq \sum\nolimits_{\phi_i(A)>0} \phi_i(A)\varepsilon \leq \varepsilon.\end{aligned}$$

Thus, $g_t(A) \in B(f(x_{i_0}), 2\varepsilon) \subseteq \mathrm{Iso}(H)$. For $t \in [0,1]$, $f_t := g_t \circ f$ is a homotopy in $\mathrm{Iso}(H)$ from $f = f_0$ to $f_1 : X \to \mathcal{B}^{\times}(H)$ with the desired property $f_1(X) \subset V \cap \mathrm{Iso}(H)$, where $V := \mathrm{span}\,(f(x_1), ..., f(x_N), I)$ with $\dim V \leq N + 1 < \infty$. For any vector $w \in H$, let

$$Vw := \{Rw \in H \mid R \in V\} = \mathrm{span}\,(w, f(x_1)w, ..., f(x_N)w)$$

We construct a sequence of orthogonal unit vectors $a_1, a_2, ... \in H$ and a sequence of orthogonal $(N+2)$-dimensional subspaces $A_1, A_2, ... \subset H$, such that for all $i = 1, 2, ...$, $Va_i \subseteq A_i$ (in particular, $a_i \in A_i$ since $I \in V$). To begin, let a_1 be any unit vector, and let A_1 be an $(N+2)$-dimensional subspace such that $Va_1 \subseteq A_1$. For a subspace $W \subseteq H$, let

$$V^{-1}(W) := \{v \in H \mid Rv \in W \text{ for all } R \in V\} = W \cap \left(\cap_{i=1}^{N} f(x_i)^{-1}(W)\right).$$

Let a_2 be any unit vector in $V^{-1}\left(A_1^{\perp}\right)$ (so that $Va_2 \subseteq A_1^{\perp}$), and let A_2 be an $(N+2)$-dimensional subspace of $A_1^{\perp}$, such that $Va_2 \subseteq A_2$. Then A_1 and A_2 are orthogonal $(N+2)$-dimensional subspaces of H with $Va_i \subseteq A_i$, $i = 1, 2$. Inductively, given orthogonal $A_1, A_2, ..., A_k \subset H$ with $Va_i \subseteq A_i$ for $i = 1, \ldots, k$, let a_{k+1} be any unit vector in $V^{-1}\left(\cap_{i=1}^{k} A_i^{\perp}\right)$ (so that $Va_{k+1} \subseteq \cap_{i=1}^{k} A_i^{\perp}$). Choose A_{k+1} to be an $(N+2)$-dimensional subspace of $\cap_{i=1}^{k} A_i^{\perp}$ (of infinite codimension), such that $Va_{k+1} \subseteq A_{k+1}$.Then we have orthogonal $A_1, A_2, ..., A_{k+1} \subset H$ with $Va_i \subseteq A_i$ for $i = 1, \ldots, k+1$, and by induction the construction is complete.

Since $f_1(X)$ is compact in $V \cap \mathrm{Iso}(H)$, there is a constant $c \geq 1$ such that $c^{-1} \leq \|f_1(x)\| \leq c$ for all $x \in X$. Let $V_c := \left\{v \in V : c^{-1} < \|v\| < c\right\}$, so that $f_1(X) \subseteq V_c$. We now deform f_1 to f_2 with the property that so that $f_2(x)(a_i)$ is a unit vector for all $x \in X$. Define $g : V_c \times [0,1] \to \mathrm{Iso}(H)$, for $v \in V_c$ and $w \in H$, by

$$g_t(v)(w) = \begin{cases} v(w) & \text{for } w \in \left(\bigoplus_{i=1}^{\infty} A_i\right)^{\perp} \\ \left((1-t) + \frac{t}{\|v(a_i)\|}\right) v(w) & \text{for } w \in A_i. \end{cases}$$

Note that $\|v(a_i)\| > c^{-1}\|a_i\| = c^{-1} > 0$, and $g_1(v)(a_i) = \frac{v(a_i)}{\|v(a_i)\|}$ is a unit vector. Let $f_{t+1}(x) = g_t\left(f_1(x)\right)$. We have agreement at $t = 0$, since $g_0\left(f_1(x)\right)(w) = f_1(x)(w)$, and

$$f_2(x)(a_i) = g_1\left(f_1(x)\right)(a_i) = \frac{f_1(x)\left(a_i\right)}{\|f_1(x)(a_i)\|},$$

which is a unit vector. Since $\dim(Va_i) \leq N+1$ and $\dim(A_i) = N+2$, we can find a unit vector $b_i \in A_i$ with $b_i \perp Va_i$. We now deform f_2 to f_3 with $f_3(x)(a_i) = b_i$ for all $x \in X$.

Let $g' : V_c \times [0,1] \to \operatorname{Iso}(H)$ be defined for $v \in V_c, w \in H$ and $t \in [0,1]$ by

$$g'_t(v)(w) = \begin{cases} \cos(\frac{\pi}{2}t)v(a_i) + \sin(\frac{\pi}{2}t)b_i & \text{if } w = v(a_i) \text{ for some } i \\ -\sin(\frac{\pi}{2}t)v(a_i) + \cos(\frac{\pi}{2}t)b_i & \text{if } w = b_i \text{ for some } i \\ w & w \in \left(\bigoplus_{i=1}^{\infty} \operatorname{span}(b_i, v(a_i))\right)^{\perp}. \end{cases}$$

Thus, $g'_t(v)$ rotates $v(a_i)$ toward b_i through the angle $\frac{\pi}{2}t$ in each of the planes $\operatorname{span}(b_i, v(a_i))$. Let $f_{t+2}(x) = g'_t(f_2(x)) \circ f_2(x)$, $t \in [0,1]$. Then

$$f_3(x)(a_i) = g'_1(f_2(x))(f_2(x)a_i) = b_i.$$

Now we will deform f_3 to f_4 with $f_4(x)(a_i) = a_i$ for all $x \in X$. Define $g'' : V_c \times [0,1] \to \operatorname{Iso}(H)$ by

$$g''_t(w) = \begin{cases} \cos(\frac{\pi}{2}t)b_i + \sin(\frac{\pi}{2}t)a_i & \text{if } w = b_i \text{ for some } i \\ -\sin(\frac{\pi}{2}t)b_i + \cos(\frac{\pi}{2}t)a_i & \text{if } w = a_i \text{ for some } i \\ w & w \in \left(\bigoplus_{i=1}^{\infty} \operatorname{span}(b_i, a_i)\right)^{\perp}. \end{cases}$$

Thus, $g'_t(v)$ rotates b_i toward a_i through the angle $\frac{\pi}{2}t$ in each of the planes $\operatorname{span}(a_i, b_i)$. Let $f_{t+3}(x) = g''_t \circ f_3(x)$, $t \in [0,1]$. Then

$$f_4(x)(a_i) = g''_1(f_3(x)a_i) = g''_1(b_i) = a_i.$$

Let H' be the closed subspace of H with orthonormal basis $a_1, a_2, \ldots$, and let $H_1 = (H')^{\perp}$. Finally, let $\pi' : H \to H'$ and $\pi_1 : H \to H_1$ be the orthogonal projections. For $t \in [4,5]$, define

$$f_t(x) = (5-t)f_4(x) + (t-4)\left(\pi' + \pi_1 \circ f_4(x) \circ \pi_1\right)$$

Note that $f_5(x)|_{H'} = I_{H'}$ and $f_5(x)(H_1) \subset H_1$, as required. □

3.4 The Atiyah-Jänich Theorem

The following generalization of Theorem 2.15 appeared in [31], where a different proof due to Atiyah and Palais is also given; for other treatments, see [2, p.153-166] and [34, p.208-210].

Theorem 3.3 (Atiyah and Jänich) *The homomorphism* $\operatorname{Index} : [X, \mathcal{F}(H)] \to K(X)$ *is an isomorphism.*

Proof We use Kuiper's Theorem 3.2 to prove that $\operatorname{Index} : [X, \mathcal{F}(H)] \to K(X)$ is injective; i.e., for $f : X \to \mathcal{F}$ with $\operatorname{Index}([f]) = 0 \in K(X)$, f is homotopic to the constant map $X \to \{\mathrm{Id}\}$. Recall that

$$\operatorname{Index}([f]) = \left[X \times \frac{H}{V}\right] - \left[\frac{H}{f(V)}\right], \text{ where } \frac{H}{f(V)} := \cup_{x \in X}\left(x, \frac{H}{f(x)(V)}\right) \to X.$$

where V is suitable for $f(X)$. Now $\operatorname{Index}([f]) = 0$ implies that the bundles $X \times \frac{H}{V}$ and $\frac{H}{f(V)}$ are stably equivalent; i.e., for some integer $N \geq 0$,

$$\left(X \times \frac{H}{V}\right) \oplus \left(X \times \mathbb{C}^N\right) \cong \frac{H}{f(V)} \oplus \left(X \times \mathbb{C}^N\right).$$

If V' is a subspace of V with $\dim(V/V') = N$, then this is the same as saying

$$X \times \frac{H}{V'} \cong \frac{H}{f(V')},$$

by virtue of the exact bundle sequences

$$0 \to X \times \frac{V}{V'} \to X \times \frac{H}{V'} \to X \times \frac{H}{V} \to 0,\ 0 \to X \times \frac{V}{V'} \to \frac{H}{f(V')} \to \frac{H}{f(V)} \to 0.$$

Indeed, we may choose any space V' of V of codimension n, and then there is an isomorphism

$$k : X \times \frac{H}{V'} \cong \left(X \times \frac{H}{V}\right) \oplus (X \times \mathbb{C}^n) \cong \frac{H}{f(V)} \oplus (X \times \mathbb{C}^n) \cong \frac{H}{f(V')}.$$

We then have a continuous map $j : X \to B((V')^{\perp}, H)$ given by the composition

$$j(x) : (V')^{\perp} \cong \frac{H}{V'} \overset{k(x,\cdot)}{\cong} \frac{H}{f(x)(V')} \cong (f(x)(V'))^{\perp} \hookrightarrow H.$$

Since $j(x)$ is of finite rank, we have a homotopy $h : X \times [0,1] \to \mathcal{F}(H)$ given by

$$h(x,t) := f(x) + t\,(0 \oplus j(x)) : V' \oplus (V')^{\perp} \cong f(x)(V') \oplus f(x)(V')^{\perp}.$$

Note that $h(x,0) = f(x)$, while $h(x,1) = f(x) + (0 \oplus j(x))$ is an isomorphism for each $x \in X$. The surjectivity of $f(x) + (0 \oplus j(x))$ is shown via

$$\begin{aligned}(f(x) + (0 \oplus j(x)))\,(V' \oplus (V')^{\perp}) &\supseteq f(x)\,(V') + j(x)((V')^{\perp}) \\ &= f(x)\,(V') + (f(x)(V'))^{\perp} = H,\end{aligned}$$

and the injectivity of $f(x) + (0 \oplus j(x))$ is proved (for $(v,w) \in V' \oplus (V')^{\perp}$) via

$$\begin{aligned}&0 = (f(x) + (0 \oplus j(x)))\,(v,w) = f(x)(v) + f(x)(w) + j(x)(w) \\ &\quad \Rightarrow j(x)(w) = 0 \text{ and } f(x)(v) = -f(x)(w) \Rightarrow w = 0 \text{ and } f(x)(v) = 0 \\ &\quad \Rightarrow w = 0 \text{ and } v = 0, \text{ since } V' \cap \operatorname{Ker}(f(x)) = \{0\}.\end{aligned}$$

Hence, $\operatorname{Index}[f] = 0$ implies that f is homotopic to a map $g = f + (0 \oplus j(x)) : X \to \operatorname{Iso}(H)$. Theorem 3.2 (of Kuiper) yields that g (and hence f) is homotopic to the constant map $X \to \{I\}$; i.e., $[f]$ is the identity in $[X, \mathcal{F}]$.

We now show that $\operatorname{Index} : [X, \mathcal{F}(H)] \to K(X)$ is onto. Let $E \to X$ be any vector bundle of dimension m. We know that E can be regarded as a subbundle of a trivial bundle, say $X \times \mathbb{C}^n$, and we let $E^{\perp}$ be the orthogonal complement of E so that $E \oplus E^{\perp} = X \times \mathbb{C}^n$. Let $\pi : X \times \mathbb{C}^n \to E$ be orthogonal projection (with kernel $E^{\perp}$). Let S_L be a left shift operator on H (i.e., under the identification of H with $L^2(\mathbb{Z}_+)$ via an orthonormal basis of H, S_R is the right shift operator on $L^2(\mathbb{Z}_+)$ with index -1). Define $B : X \to \mathcal{F}(H \otimes \mathbb{C}^n)$ by

$$B(x) = S_R \otimes \pi_x + I_H \otimes (I_{\mathbb{C}^n} - \pi_x)$$

For $x \in X$, we let $e_1, \ldots, e_n$ be an orthonormal basis of $\mathbb{C}^n$ such that $e_1, \ldots, e_m$ is an orthonormal basis of E_x. For any $v_1,\ldots,v_n \in H$,

$$\begin{aligned}
&(S_R \otimes \pi_x + I_H \otimes (I_{\mathbb{C}^n} - \pi_x))(v_1 \otimes e_1 + \cdots + v_n \otimes e_n) \\
&= ((S_R \otimes \pi_x)(\oplus_{i=1}^n v_i \otimes e_i) + (I_H \otimes (I_{\mathbb{C}^n} - \pi_x)) \oplus_{i=1}^n v_i \otimes e_i) \\
&= S_R(v_1) \otimes e_1 + \cdots + S_R(v_m) \otimes e_m + v_{m+1} \otimes e_{m+1} + \cdots + v_n \otimes e_n,
\end{aligned}$$

which is $0 \Leftrightarrow v_1 = \cdots = v_n = 0$. If $0 \neq v \in S_R(H)^\perp$, then a complement to $B(H \otimes \mathbb{C}^n)$ in $H \otimes \mathbb{C}^n$ is $\operatorname{span}\{v \otimes e_1, \cdots, v \otimes e_m\} \cong S_R(H)^\perp \otimes E_x \cong E_x$. Thus, $B(x)$ is injective for all x, and

$$\frac{H \otimes \mathbb{C}^n}{B(H \otimes \mathbb{C}^n)} := \cup_{x \in X} \frac{H \otimes \mathbb{C}^n}{B(x)(H \otimes \mathbb{C}^n)} \cong \cup_{x \in X} E = E.$$

Following B with an isomorphism $H \otimes \mathbb{C}^n \cong H$, we obtain $\widetilde{B} : X \to H$ with $\operatorname{Index} \widetilde{B} = -[E]$, and $\operatorname{Index}(\widetilde{B} + (S_L)^k) = [k] - [E]$. Since any element of $K(X)$ is of the form $[k] - [E]$ (as its negative is of the form $[E] - [k]$) for some vector bundle $E \to X$, we have $\operatorname{Index} : [X, \mathcal{F}(H)] \to K(X)$ is onto. □

4 Elliptic operators and Sobolev spaces

Elliptic operators on sections of complex vector bundles over manifolds provide a primary source of Fredholm operators. In this section, we indicate how this happens.

4.1 Differential operators and symbols

Let X be a compact C^∞ n-manifold without boundary, and let $\pi_E : E \to X$ be a C^∞ complex vector bundle over X. We denote the linear space of C^∞ sections of E by $C^\infty(E) := \{s : X \to E \mid \pi_E \circ s = Id_X\}$. Unless otherwise stated, we remain in the C^∞ category. Let $\pi_F : F \to X$ be another complex vector bundle. A linear function $P : C^\infty(E) \to C^\infty(E)$ is called a differential operator of order k, if for all coordinate neighborhoods U and trivializations $\tau_E : E_{|U} \cong U \times \mathbb{C}^M$ and $\tau_F : F_{|U} \cong U \times \mathbb{C}^N$, P can be locally expressed in the form

$$P[s](x) = \tau_F^{-1}\left(\sum\nolimits_{|\alpha| \le k} a^\alpha(x) D_\alpha(\tau_E \circ s)\right), \ x \in U.$$

where α ranges over all multi-indices $(\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}_+ \times \overset{n}{\cdots} \times \mathbb{Z}_+$ with $|\alpha| := \alpha_1 + \cdots + \alpha_n \le k$, $a^\alpha(x) \in \operatorname{Hom}(\mathbb{C}^M, \mathbb{C}^N)$, and

$$D_\alpha := i^{-|\alpha|} \frac{\partial^{|\alpha|}}{\partial^{\alpha_1} x^1 \cdots \partial^{\alpha_n} x^n},$$

where $x^1, \ldots, x^n$ are the local coordinates on U. The reason for inclusion of the factor $i^{-|\alpha|}$ is related to the fact under Fourier transform $\frac{1}{i}\frac{d}{dx}$ is converted into a simple multiplication operator, and it's also convenient when integrating Hermitian inner products by parts. Let $\pi : T^*X \to X$ denote the cotangent bundle of X, and let $\pi^*E \to T^*X$ and $\pi^*F \to T^*X$

be the pull-backs of $\pi_E : E \to X$ and $\pi_E : F \to X$ via π. We then have a bundle $\mathrm{Hom}(\pi^*E, \pi^*F) \to T^*X$. We will define the (principal) **symbol** $\sigma(P)$ of P as a section $\sigma(P) : T^*X \to \mathrm{Hom}(\pi^*E, \pi^*F)$ of this bundle; i.e., $\sigma(P) \in C^\infty(\mathrm{Hom}(\pi^*E, \pi^*F))$. For $\xi_x \in T_x^*X$, note that the fiber $(\pi^*E)_{\xi_x}$ may be identified with E_x and we will do so; we write the identification as $(\pi^*E)_{\xi_x} \sim E_x$. In terms of local coordinates x^1, ..., x^n on U, we may write $\xi_x = \xi_1 dx^1 + \cdots + \xi_n dx^n$ and write $\xi_{\alpha_1} \cdots \xi_{\alpha_n}$ as ξ_α. For $e \in (\pi^*E)_{\xi_x} \sim E_x$, we then define

$$\sigma(P)(\xi_x)(e) := \tau_F^{-1}(x, \sum\nolimits_{|\alpha|=m} a^\alpha(x)\,(\tau_E(x,e))\,\xi_\alpha) \in F_x \sim (\pi^*F)_{\xi_x},$$
$$= \sum\nolimits_{|\alpha|=m} \tau_F^{-1}(x, a^{(\alpha_1,\ldots,\alpha_n)}(x)\,(\tau_E(x,e)))\xi_{\alpha_1}\cdots\xi_{\alpha_n}$$

Making identifications $(\pi^*E)_{\xi_x} \sim E_x \sim \mathbb{C}^M$ (via τ_E) and $(\pi^*F)_{\xi_x} \sim F_x \sim \mathbb{C}^N$ (via τ_F), we can write this more transparently as

$$\sigma(P)(\xi_x) = \sum\nolimits_{|\alpha|=m} a^{(\alpha_1,\ldots,\alpha_n)}(x)\xi_{\alpha_1}\cdots\xi_{\alpha_n} \in \mathrm{Hom}(\mathbb{C}^M, \mathbb{C}^N).$$

One can show that $\sigma(P) \in C^\infty(\mathrm{Hom}(\pi^*E, \pi^*F))$ is well-defined (i.e., independent of the choice of local coordinates and trivializations τ_E and τ_F), but this is not obvious. Indeed, if we had summed over α with $|\alpha| = m - 1$, the resulting so-called "subprincipal symbol" is not well-defined.

4.2 Elliptic differential operators

Definition 4.1 Let $P : C^\infty(E) \to C^\infty(F)$ be a differential operator with symbol $\sigma(P) : T^*X \to \mathrm{Hom}(\pi^*E, \pi^*F)$. If $\sigma(P)(\xi_x)$ is an isomorphism for all *nonzero* $\xi_x \in T_x^*X$, then P is called an **elliptic differential operator**.

We consider some standard examples. Even though $C^\infty(E)$ and $C^\infty(F)$ are not Hilbert spaces (and hence P is not Fredholm), here we take the index of P to be $\dim\mathrm{Ker}\,P - \dim\mathrm{Ker}\,P^*$, where P^* is the formal adjoint of P. This is the same as the usual index of a Fredholm extension of P to a suitable Sobolev space, as is explained in the next subsection.

Example 4.2 (The Laplace operator) Let ∂_{x^i} be a shorthand notation for $\frac{\partial}{\partial x^i}$. A simple example is the Laplace operator $\Delta = \partial_{x^1}^2 + \cdots + \partial_{x^n}^2$ on $C^\infty(\mathbb{R}^n, \mathbb{C})$, which we may regard as the space of sections of the trivial bundle $\mathbb{R}^n \times \mathbb{C} \to \mathbb{R}^n$. We have

$$\sigma(\Delta) : T^*\mathbb{R}^n \to \mathrm{Hom}(\pi^*(\mathbb{R}^n \times \mathbb{C}), \pi^*(\mathbb{R}^n \times \mathbb{C})), \text{ given by}$$
$$\sigma(\Delta)(\xi_1 dx^1 + \cdots + \xi_n dx^n) = -\left(\xi_1^2 + \cdots + \xi_2^2\right) \in \mathrm{End}(\pi^*(\mathbb{R}^n \times \mathbb{C})_\xi),$$

regarded as multiplication on $\pi^*(\mathbb{R}^n \times \mathbb{C})_\xi \cong \mathbb{C}$.

Example 4.3 (The Cauchy-Riemann operator) The Cauchy-Riemann operator

$$\partial_{\bar z} = \frac{\partial}{\partial \bar z} = \frac{1}{2}(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}) = \frac{1}{2}(\partial_x + i\partial_y) : C^\infty\left(\mathbb{R}^2, \mathbb{C}\right) \to C^\infty\left(\mathbb{R}^2, \mathbb{C}\right)$$

is a first-order elliptic operator on the same space of sections as in Example 4.2, but with a different symbol

$$\sigma(\frac{\partial}{\partial \bar z})(\xi_1 dx^1 + \xi_2 dx^2) = \frac{1}{2}i(\xi_1 + i\xi_2),$$

regarded as complex multiplication on $\pi^*\left(\mathbb{R}^2\times\mathbb{C}\right)_\xi\cong\mathbb{C}$. One similarly has the elliptic operator $\partial_z=\frac{1}{2}\left(\partial_x-i\partial_y\right)$. Note that $\Delta=4\frac{\partial^2}{\partial\bar{z}\partial z}$ and at $\xi_1dx^1+\xi_2dx^2$ we have

$$\sigma(\Delta)=4\sigma(\partial_z)\sigma(\partial_{\bar{z}})=4\tfrac{1}{2}i\left(\xi_1-i\xi_2\right)\tfrac{1}{2}i\left(\xi_1+i\xi_2\right)=-\left(\xi_1^2+\xi_2^2\right).$$

There is a type of exterior derivative on $\Omega^{0,0}\left(\mathbb{R}^2,\mathbb{C}\right):=C^\infty\left(\mathbb{R}^2,\mathbb{C}\right)$, namely the Dolbeault operator

$$\overline{\partial}:\Omega^{0,0}\left(\mathbb{R}^2,\mathbb{C}\right)\to\Omega^{0,1}\left(\mathbb{R}^2,\mathbb{C}\right):=\left\{gd\bar{z}\mid g\in C^\infty\left(\mathbb{R}^2,\mathbb{C}\right)\right\},$$
given by $\partial_{\bar{z}}(f):=\frac{\partial f}{\partial\bar{z}}d\bar{z}$.

Here $\Omega^{0,1}\left(\mathbb{R}^2,\mathbb{C}\right)$ is called the space of complex forms of type $(0,1)$. For a compact Riemann surface S, a strictly analogous operator $\overline{\partial}:\Omega^{0,0}(S,\mathbb{C})\to\Omega^{0,1}(S,\mathbb{C})$ can be defined. By the Riemann-Roch Theorem, we have $\operatorname{index}\overline{\partial}=1-g$, where g is the genus of S (i.e., the number of holes). For higher-dimensional compact, complex manifolds, there is a Dolbeault operator complex $\overline{\partial}:\Omega^{0,k}(X,\mathbb{C})\to\Omega^{0,k+1}(X,\mathbb{C})$, $k=0,1,\dots m:=\dim_{\mathbb{C}}X$ which can be "rolled up" to give an elliptic operator

$$\overline{\partial}:\oplus_{k\text{ even}}^n\Omega^{0,k}(X,\mathbb{C})\to\oplus_{k\text{ odd}}^n\Omega^{0,k}(X,\mathbb{C}).$$

The Hirzebruch-Riemann-Roch Theorem expresses $\operatorname{index}\overline{\partial}$, called the arithmetic genus of X or the holomorphic Euler characteristic $\chi(X)$ in terms of the *Todd genus of* X which in turn can be expressed in terms of Chern numbers, as in the table

$\dim_{\mathbb{C}}X$	1	2	3	4
$\chi(X)$	$\frac{1}{2}c_1$	$\frac{1}{12}\left(c_2+c_1^2\right)$	$\frac{1}{24}c_1c_2$	$\frac{1}{720}\left(-c_4+c_3c_1+3c_2^2+4c_2c_1^2-c_1^4\right).$

Example 4.4 (The Euler operator) Let $\Lambda^k(X)\to X$ denote the bundle of complex exterior k-covectors over the compact, orientable C^∞ Riemannian n-manifold X with metric tensor g. Let $\Omega^k(X)=C^\infty\left(\Lambda^k(X)\right)$ denote the space of C^∞ sections of $\Lambda^k(X)$, namely the space of $\mathbb{C}$-valued k-forms on X. We have the exterior derivative $d:\Omega^k(X)\to\Omega^{k+1}(X)$ and the codifferential $\delta:\Omega^{k+1}(X)\to\Omega^k(X)$ which is the formal adjoint of d; i.e.,

$$(d\alpha,\beta)=\int_X\langle d\alpha,\beta\rangle_g\nu_g=\int_X\langle\alpha,\delta\beta\rangle_g\nu_g=(\alpha,\delta\beta),$$

where ν_g is the volume element and $\langle\cdot,\cdot\rangle_g$ is the inner product on $\Lambda^*(X)$ induced by g. Let $\Omega^*(X)=\oplus_{k=0}^n\Omega^k(X)$. In terms of the Hodge star operator $*$,

$$\delta=-(-1)^{nk}*d*:\Omega^{k+1}(X)\to\Omega^k(X).$$

Then we have a first-order operator $d+\delta:\Omega^*(X)\to\Omega^*(X)$. That $d+\delta$ is elliptic is seen as follows. Let $\wedge$ and $\llcorner$ denote wedge product and interior product on $\Lambda^*(X)$. The symbol

$$\sigma(d+\delta):T^*X\to\operatorname{Hom}\left(\pi^*\Lambda^*(X),\pi^*\Lambda^*(X)\right)\text{ is given at }\xi\in T^*X\text{ by}$$
$$\sigma(d+\delta)_\xi(\alpha)=i\left(\xi\wedge\alpha-\xi\llcorner\alpha\right)\text{ where }\alpha\in\left(\pi^*\Lambda^*(X)\right)_\xi.$$

Note that $\sigma(d+\delta)_\xi$ is invertible for all $\xi \neq 0$ (and hence $d+\delta$ is elliptic), since $(\sigma(d+\delta)_\xi \circ \sigma(d+\delta)_\xi) = -|\xi|^2 \operatorname{Id}_{\pi^*\Lambda^*(X)_\xi}$:

$$\begin{aligned}
&(\sigma(d+\delta)_\xi \circ \sigma(d+\delta)_\xi)(\alpha) \\
&= i\,(\xi \wedge (\sigma(d+\delta)_\xi(\alpha)) - \xi \llcorner \sigma(d+\delta)_\xi(\alpha)) \\
&= -\,(\xi \wedge (\xi \wedge \alpha - \xi \llcorner \alpha) - \xi \llcorner (\xi \wedge \alpha - \xi \llcorner \alpha)) \\
&= \xi \wedge (\xi \llcorner \alpha) - \xi \llcorner (\xi \wedge \alpha) = -\,(\xi \llcorner \xi)\,\alpha = -\,|\xi|^2\,\alpha.
\end{aligned}$$

Since $d+\delta$ is formally self-adjoint, its index is zero. By definition, elements of $\operatorname{Ker}(d+\delta)$ are known as harmonic forms. Hodge theory tells us that the algebra of harmonic forms, say $\mathcal{H}^*(X)$, is isomorphic to the cohomology algebra $H^*(X;\mathbb{C})$ where wedge product of harmonic forms corresponds to cup product in $H^*(X;\mathbb{C})$. If we restrict $d+\delta$ to $\Omega^e(X) = \oplus_{k \text{ even}}^n \Omega^k(X)$, we obtain a differential operator

$$(d+\delta)^e : \Omega^e(X) \to \Omega^o(X) := \oplus_{k \text{ odd}}^n \Omega^k(X),$$

whose formal adjoint is $(d+\delta)^o := (d+\delta)|_{\Omega^o(X)}$. The restricted operators $(d+\delta)^e$ and $(d+\delta)^o$ are still elliptic with symbols that are inverses modulo a factor of $-|\xi|^2$. The index of $(d+\delta)^e$ is not necessarily zero. Indeed,

$$\begin{aligned}
\operatorname{index}(d+\delta)^e &= \dim \operatorname{Ker}(d+\delta)^e - \dim \operatorname{Ker}(d+\delta)^o \\
&= \sum_{k=0}^n (-1)^k \dim H^k(X;\mathbb{C}) = \chi(X),
\end{aligned}$$

the Euler characteristic of X. Consequently, $(d+\delta)^e$ is sometimes called the Euler operator.

Example 4.5 (The Hirzebruch signature operator) With the notation of Example 4.4, let $*_k : \Lambda^k(X) \to \Lambda^{n-k}(X)$ be the Hodge star operator, characterized by the property $(*_k\alpha)\wedge \beta = \langle \alpha, \beta \rangle\, v_g$. Assume that n is even, say $n = 2m$. Since $*_{2m-k} \circ *_k = (-1)^k$, $* := \bigoplus_{k=0}^n *_k$is not an involution. However, let $\tau_k = i^{m+k(k-1)} *_k$. Then

$$\tau_{2m-k} \circ \tau_k = i^{m+(2m-k)(2m-k-1)} i^{m+k(k-1)} (-1)^k = \cdots = 1,$$

and so $\tau := \bigoplus_{k=0}^{2m} \tau_k$ is an involution ($\tau^2 = 1 := \operatorname{Id}_{\Lambda^*(X)}$). Thus, $\Lambda^*(X) = \Lambda^+(X) \oplus \Lambda^-(X)$, where $\Lambda^+(X) = (1+\tau)\,\Lambda^*(X)$ and $\Lambda^-(X) = (1-\tau)\,\Lambda^*(X)$ are the ± 1 eigenbundles of τ; we set $\Omega^\pm(X) := C^\infty(\Lambda^\pm(X))$. Using $\delta = -*d*$ (for $n = 2m$ even), one can check that $(d+\delta)\,\tau = -\tau\,(d+\delta)$ so that $(d+\delta)\,(\Omega^\pm(X)) = \Omega^\mp(X)$. The Hirzebruch signature operator is

$$(d+\delta)^+ := (d+\delta)|_{\Omega^+(X)} : \Omega^+(X) \to \Omega^-(X).$$

with adjoint $(d+\delta)^- := (d+\delta)|_{\Omega^-(X)}$. We have

$$\operatorname{Ker}(d+\delta)^\pm = (1 \pm \tau_m)\mathcal{H}^m(X) \oplus \left(\oplus_{k<m}(1 \pm \tau_k)\mathcal{H}^k(X)\right)$$

For $k < m$, the maps

$$(1 \pm \tau_k) : \mathcal{H}^k(X) \to \mathcal{H}^k(X) \oplus \tau_k\mathcal{H}^k(X) = \mathcal{H}^k(X) \oplus \mathcal{H}^{2m-k}(X)$$

are injections, and so for any $k < m$, $(1 \pm \tau_k)\mathcal{H}^k(X)$ have equal dimensions. Consequently, only when $k = m$ do we have a contribution to the index:

$$\begin{aligned}\operatorname{index}(d+\delta)^+ &= \dim\operatorname{Ker}(d+\delta)^+ - \dim\operatorname{Ker}(d+\delta)^- \\ &= \dim\left((1+\tau_m)\mathcal{H}^m(X)\right) - \dim\left((1-\tau_m)\mathcal{H}^m(X)\right).\end{aligned}$$

Note that $\tau_m = i^{m+m(m-1)} *_m = i^{m^2} *_m = \begin{cases} \pm i *_m & \text{for } m \text{ odd} \\ *_m & \text{for } m \text{ even.}\end{cases}$

It follows that $\operatorname{index}(d+\delta)^+ = 0$ for m odd (i.e., $n \equiv 2 \operatorname{mod} 4$). Consider m even, so that $\tau_m = *_m$. Let $\mathcal{H}^m(X)_{\mathbb{R}}$ be the space of $\mathbb{R}$-valued m-forms α with $(d+\delta)\,\alpha = 0$. We have a quadratic form $Q : \mathcal{H}^m(X) \to \mathbb{R}$ given by

$$Q(\alpha) = \int_X \alpha \wedge \alpha = \int_X \langle \alpha, *_m \alpha\rangle_g \ \nu_g.$$

Note that Q is positive-definite on $(1+\tau_m)\mathcal{H}^m(X)_{\mathbb{R}}$ on which $*_m$ is Id and negative-definite on $(1-\tau_m)\mathcal{H}^m(X)_{\mathbb{R}}$. Thus, for $n \equiv 0 \operatorname{mod} 4$,

$$\operatorname{index}(d+\delta)^+ = \operatorname{Sign} Q := \text{the signature of } Q.$$

As a consequence of the Atiyah-Singer Index Formula, $\operatorname{Sign} Q$ is the L-genus of X, denoted by $L(X)$ expressible in terms of the Pontryagin numbers of X; e.g.,

$$\operatorname{index}(d+\delta)^+ = \operatorname{Sign}(Q) = L(X) = \begin{cases} \frac{1}{3}p_1 & \text{for } n = 4 \\ \frac{1}{45}\left(7p_2 - p_1^2\right) & \text{for } n = 8.\end{cases}$$

This result explained the once mysterious fact that p_1 is divisible by 3.

4.3 Sobolev spaces

To describe how an elliptic differential operator P determines a Fredholm operator (and hence an index), we need to introduce Sobolev spaces. We equip the bundle $E \to X$ with a Hermitian structure and compatible covariant differentiation operator $\nabla^E : C^\infty(E) \to C^\infty(T^*X \otimes E)$, and do the same for $F \to X$. By also employing a Riemannian metric g and Levi-Civita connection ∇^X on X, we obtain for any $k \geq 0$ a connection

$$\nabla = \nabla^{X,E} : C^\infty\left(\left(\otimes^k T^*X\right) \otimes E\right) \to C^\infty\left(\left(\otimes^{k+1} T^*X\right) \otimes E\right)$$

For $u, v \in C^\infty(E)$ and $k > 0$, we then set

$$(u,v)_k := \sum\nolimits_{j=0}^k \int_X \left\langle \nabla^j u, \nabla^j v\right\rangle \nu_g, \text{ and } \|u\|_k = \sqrt{(u,u)_k},$$

where ν_g is the volume element for g, and the inner product $\left\langle \nabla^j u, \nabla^j v\right\rangle$ is the natural one constructed from the one induced by g on $\otimes^j T^*X$ and the Hermitian structure on E.

Definition 4.6 The **Sobolev space** $W^k(E)$ is the completion of $C^\infty(E)$ with the norm $\|\cdot\|_k$.

The Hilbert space $W^k(E)$ coincides with the subspace of $L^2(E) := W^0(E)$ consisting of sections that have "weak derivatives" of orders $\leq k$ in L^2, a notion described as follows. The formal adjoint of

$$\nabla : C^\infty((\otimes^k T^*X) \otimes E) \to C^\infty((\otimes^{k+1} T^*X) \otimes E),$$

is the unique first order differential operator

$$\nabla^* : C^\infty((\otimes^{k+1} T^*X) \otimes E) \to C^\infty((\otimes^k T^*X) \otimes E),$$

such that for all $u \in C^\infty((\otimes^k T^*X) \otimes E)$ and $w \in C^\infty((\otimes^{k+1} T^*X) \otimes E)$, we have

$$\int_X \langle \nabla u, w \rangle \, \nu_h = \int_X \langle u, \nabla^* w \rangle \nu_h.$$

Now, $u \in W^k(E)$, if for each $j \leq k$, there is $v_j \in L^2(\otimes^j (T^*X) \otimes E)$, such that for all $w \in C^\infty(\otimes^j (T^*X) \otimes E)$, we have $\int_X \langle v_j, w \rangle \, \nu_h = \int_X \langle u, (\nabla^*)^j \, w \rangle \nu_h$. We say that "$\nabla^j u = v_j$ in the weak (or distributional) sense." Proofs of the following theorems in this subsection may be found in [27], [34], [40] and [41].

Theorem 4.7 *Let $P : C^\infty(E) \to C^\infty(F)$ be a linear differential operator of order m (not necessarily elliptic). For each $k \geq 0$, P has a unique continuous extension*

$$D_{k+m} : W^{k+m}(E) \to W^k(E) \text{ with } \|D_{k+m}(\alpha)\|_k \leq K \|\alpha\|_{k+m},$$

for some $K > 0$, independent of $\alpha \in W^{k+m}(E)$.

Let $C^m(E)$ be the Banach space of m-times (strongly) differentiable sections of E with norm

$$\|u\|_{C^m} := \sum_{j=0}^{m} \sup_{x \in X} \left| (\nabla^j u)_x \right|, \; u \in C^m(E).$$

Theorem 4.8 *For $n = \dim X$, $k \in \mathbb{Z}^+$ and $0 \leq m < k - \frac{n}{2}$, we have a compact (i.e., completely continuous) inclusion*

$$W^k(E) \subseteq C^m(E). \tag{1}$$

For $k > h$, we also have a compact inclusion

$$W^k(E) \subseteq W^h(E).$$

Theorem 4.9 (Fundamental elliptic estimate) *Assume that $P : C^\infty(E) \to C^\infty(F)$ is an elliptic differential operator of order m, with formal adjoint $P^* : C^\infty(F) \to C^\infty(E)$. Suppose that for some $u \in L^2(E)$, we have $v \in W^k(F)$ such that*

$$\int_X \langle v, w \rangle \; \nu_h = \int_X \langle u, P^* w \rangle \, \nu_h,$$

for all $w \in C^\infty(F)$ (i.e., $Pu := v$ exists weakly in $W^k(F)$). Then $u \in W^{k+m}(E)$. Moreover, for each $k \geq 0$, there is a constant $C_k > 0$ independent of u, such that

$$\|u\|_{k+m} \leq C_k \left(\|Pu\|_k + \|u\|_2 \right).$$

If $Pu \in C^\infty(F)$, then for all $k \geq 0$, $\|Pu\|_k < \infty$, and we have $\|u\|_{k+m} < \infty$, in which case $u \in C^\infty(E)$ by (1); e.g., weak solutions of $Pu = 0$ are C^∞.

Theorem 4.10 (Elliptic decomposition) *Let $P : C^\infty(E) \to C^\infty(F)$ be a differential operator of order m with a symbol which is injective or surjective, and let P^* be the formal adjoint of P. If $P^{k+m} : W^{k+m}(E) \to W^k(F)$ $(P^*)^{k+m} : W^{k+m}(F) \to W^k(E)$ are the Sobolev extensions of P and P^*, then we have the following direct sum decompositions into closed subspaces for $k \geq 0$,*

$$W^{k+m}(E) = \mathrm{Ker}(P^{k+m}) \oplus \mathrm{Im}\left((P^*)^{k+2m}\right) \quad \textit{and}$$
$$W^{k+m}(F) = \mathrm{Ker}((P^*)^{k+m}) \oplus \mathrm{Im}(P^{k+2m}).$$

If the symbol of P is injective, then $P^* \circ P$ is elliptic and

$$\mathrm{Ker}(P^{k+m}) = \mathrm{Ker}(P) = \mathrm{Ker}(P^* \circ P) \subseteq C^\infty(E)$$

is finite-dimensional. If the symbol of P is surjective, then $P \circ P^*$ is elliptic and

$$\mathrm{Ker}((P^*)^{k+m}) = \mathrm{Ker}(P^*) = \mathrm{Ker}(P \circ P^*) \subseteq C^\infty(F)$$

is finite-dimensional. In particular, if P is elliptic (i.e., with injective and surjective symbol), then both $\mathrm{Ker}(P)$ and $\mathrm{Ker}(P^*)$ are finite-dimensional and $P^{k+2m} : W^{k+2m}(E) \to W^{k+m}(F)$ is Fredholm with

$$\mathrm{index}\left(P^{k+2m}\right) = \dim\mathrm{Ker}(P) - \dim\mathrm{Ker}(P^*) =: \mathrm{index}\, P.$$

4.4 Pseudo-differential operators

We work within the C^∞ category unless stated otherwise. Let ρ be the injectivity radius of the compact manifold X with Riemannian metric g and Levi-Civita connection ∇; i.e., for all $x \in X$, the exponential map $\exp_x : T_xX \to X$ relative to g is injective on the disk of radius ρ about $0_x \in T_xX$. Let $\pi_E : E \to X$ and $\pi_F : F \to X$ be complex Hermitian vector bundles equipped with Hermitian connections $\nabla^E : C^\infty(E) \to C^\infty(T^*X \otimes E)$ and $\nabla^F : C^\infty(F) \to C^\infty(T^*X \otimes F)$, where $C^\infty(E)$ denotes the space of (smooth) sections of $\pi_E : E \to X$. For $x, y \in X$, with $d(x,y) < \rho$, let $\tau^E_{x,y} : E_y \to E_x$ denote parallel translation relative to ∇^E along the unique geodesic from y to x with minimal length $d(x,y)$. Let $\psi : [0,\infty) \to [0,1]$ be smooth, with $\psi(r) = 1$ for $r \in [0, \rho/3]$ and $\psi(r) = 0$ for $r \in [2\rho/3, \infty]$. For $\pi : T^*X \to X$ and $u \in C^\infty(E)$, let $u^\wedge \in C^\infty(\pi^*E)$ be defined (where $x = \pi(\xi)$ and $\xi \in T^*X$) by

$$u^\wedge(\xi) := \int_{T_xX} e^{-i\xi(v)}\psi(|v|)\tau^E_{x,\exp_x v}\left[u(\exp_x v)\right] d'v \in E_x \text{ for } \xi \in T^*_xX.$$

where $d'v = (2\pi)^{-n/2}dv$ and dv is the volume element on T_xX associated with g_x. For $x, y \in X$ with $d(x,y) < \rho$, we have $y = \exp_x v$ for a unique $v \in T_xX$ with $|v| = d(x,y)$, and we may define $\alpha \in C^\infty(X \times X, [0,1])$ by

$$\alpha(x,y) := \begin{cases} \psi\left(d(x,y)\right) = \psi(|v|) & \text{for } d(x,y) < \rho \\ 0 & \text{for } d(x,y) \geq \rho. \end{cases}$$

Note that we can think of the function (in $C^\infty(T_xX, E_x)$)

$$v \mapsto \psi(|v|)\tau^E_{x,\exp_x v}\left[u(\exp_x v)\right] = \tau^E_{x,\exp_x v}\left[\alpha(x,\exp_x v)u(\exp_x v)\right] \ (v \in T_xX)$$

as a "pull-back" (of sorts), using τ^E and $\exp_x : T_xX \to X$, of the bump function $\alpha(x,\cdot)$ times $u(\cdot)$ in a neighborhood of x, and $u^\wedge|_{T_x^*X}$ is the Fourier transform of this "pull-back" of $\alpha(x,\cdot)u(\cdot)$. The "inverse Fourier transform" $(u^\wedge)^\vee : TX \to E$ of $u^\wedge$ is given by

$$(u^\wedge)^\vee(v) := \int_{T_x^*X} e^{i\xi(v)} u^\wedge(\xi)\, d'\xi = \psi(|v|)\tau_{x,\exp_x v}[u(\exp_x v)],$$

where $d'\xi = (2\pi)^{-n/2}d\xi$. Since $(u^\wedge)^\vee(v) \in E_x$, $(u^\wedge)^\vee$ is a section of the pull-back of E to TX via $\pi : T^*X \to X$. Moreover, we can recover u locally about x from $u^\wedge{}_{|T_x^*X}$. In particular, for $v = 0_x \in T_xX$, we have $(u^\wedge)^\vee(0_x) = u(x)$. For $\pi : T^*X \to X$ and a section $p \in C^\infty(\mathrm{Hom}(\pi^*E, \pi^*F))$ (of $\mathrm{Hom}(\pi^*E,\pi^*F) \to T^*X$), we define an operator $\mathrm{Op}(p) : C^\infty(E) \to C^\infty(F)$ via

$$\begin{aligned}
\mathrm{Op}(p)(u)_x &:= \left.\int_{T_x^*X} e^{i\xi(v)} p(\xi)\left(u^\wedge(\xi)\right) d'\xi\right|_{v=0} = \int_{T_x^*X} p(\xi)\left(u^\wedge(\xi)\right) d'\xi \\
&= \int_{T_x^*X} p(\xi)\left(\int_{T_xX} e^{-i\xi(v)}\psi(|v|)\tau^E_{x,\exp_x v}\ [u(\exp_x v)]\, d'v\right) d'\xi \\
&= \int_{T_xX\times T_x^*X} p(\xi)\left(e^{-i\xi(v)}\psi(|v|)\tau^E_{x,\exp_x v}\ [u(\exp_x v)]\right) d'v d'\xi \\
&= \int_{T_xX\times T_x^*X} e^{-i\xi(v)} p(\xi)\left(\tau^E_{x,\exp_x v}\ [\alpha(x,\exp_x v)u(\exp_x v)]\right) d'v d'\xi. \qquad (2)
\end{aligned}$$

Recall roughly that quantization in quantum mechanics attempts to convert functions of position and momentum (i.e., functions on T_x^*X) into operators. One may think of $\mathrm{Op}(p)$ as a quantization of p, but $\mathrm{Op}(p)$ depends on many choices (e.g., the choice of metric, connections, and $\alpha : X \times X \to [0,1]$). Apart from these choices, there are other choices one can make, as is discussed in [46]. For example, if $s \in [0,1]$, let $T_{x,\exp_x sv} : T^*_{\exp_x sv}X \to T_x^*X$ denote parallel translation (with respect to the Levi-Civita connection) for T^*X along the geodesic $t \mapsto \exp_x tv$ in the reverse direction from $\exp_x sv$ to x. In [46] (but with notation that differs from ours), an operator $\mathrm{Op}(p;s)$ (depending on s) is associated to p via

$$\begin{aligned}
&\mathrm{Op}(p;s)(u)_x \\
&= \int_{T_xX\times T_x^*X} d'vd'\xi \\
&\quad e^{-i\xi(v)}\alpha(x,\exp_x v)\tau^F_{x,\exp_x sv} p(T_{x,\exp_x sv}(\xi))\tau^E_{\exp_x sv,\ \exp_x v} u(\exp_x v).
\end{aligned}$$

When $s = 0$, we get

$$\mathrm{Op}(p;0)(u)_x = \int_{T_xX\times T_x^*X} d'vd'\xi\, e^{-i\xi(v)}\alpha(x,\exp_x v)p(\xi)\tau^E_{x,\ \exp_x v}(u(\exp_x v))$$

which is precisely our $\mathrm{Op}(p)$. In cases of interest, the operators $\mathrm{Op}(p;s)$ for different s differ by "lower order" operators which do not affect the index (if defined). Hence, for simplicity, we only use $s = 0$. As stated in [46] the choice of s is related to the choice of operator ordering of monomials in position and momentum variables under quantization.

The connections ∇^E and ∇^F pull back via $\pi : T^*X \to X$ to connections on the bundles $\pi^*E \to T^*X$ and $\pi^*F \to T^*X$, which we continue to denote by ∇^E and ∇^F. The Levi-Civita connection for (X,g) determines a subbundle H of $T(T^*X)$ consisting of horizontal subspaces of $T(T^*X)$, which is complementary to the subbundle V of $T(T^*X)$ consisting of vectors which are tangent to the fibers of $T^*X \to X$. There is a natural Riemannian metric, say g^*, on T^*X such that V and H are orthogonal and g^* equals g on V and π^*g on H. Using ∇^E and ∇^F, along with the Levi-Civita connection for g^*, say ∇^*, we may construct a covariant derivative

$$\widetilde{\nabla} : C^\infty\left(\mathrm{Hom}(\pi^*E,\pi^*F)\right) \to C^\infty\left(T^*\left(T^*X\right) \otimes \mathrm{Hom}(\pi^*E,\pi^*F)\right).$$

Since ∇^* extends to $\otimes^k T^*\left(T^*X\right)$, we may "iterate" $\widetilde{\nabla}$ to obtain

$$\widetilde{\nabla}^k : C^\infty\left(\mathrm{Hom}(\pi^*E,\pi^*F)\right) \to C^\infty\left(\otimes^k T^*\left(T^*X\right) \otimes \mathrm{Hom}(\pi^*E,\pi^*F)\right).$$

Definition 4.11 We say that $p \in C^\infty(\mathrm{Hom}(\pi^*E,\pi^*F)$ is a **symbol of order** $m \in \mathbb{R}$ if for any $H_1,\ldots,H_I \in C^\infty(H)$ with $|H_1|,\ldots,|H_I| \le 1$ and $V_1,\ldots,V_J \in C^\infty(V)$, there are constants C_{IJ} (depending only on I, J and p), such that

$$\left|\left(\widetilde{\nabla}^{I+J} p\right)(H_1,\ldots,H_I,V_1,\ldots,V_J)\right| \le C_{IJ}\left(1+\sum\nolimits_{j=1}^{J}|V_j|\right)^{m-J}.$$

Moreover, we require that the m-th order asymptotic symbol of p, namely

$$\sigma_m(p)(\xi) := \lim_{t\to\infty}\frac{p(t\xi)}{t^m} \quad (\text{for } \xi \ne 0)$$

exist, where the convergence is uniform on $S(T^*X)$. Then we call $\mathrm{Op}(p)$ (see (2)) a **pseudo-differential operator of order** m. We denote the set of symbols of order m by $\mathrm{Symb}_m(E,F)$.

Clearly, for $m' > m$,

$$\mathrm{Symb}_{m'}(E,F) \supset \mathrm{Symb}_m(E,F) \supset \mathrm{Symb}_{-\infty}(E,F) := \bigcap\nolimits_{m=0}^{-\infty}\mathrm{Symb}_m(E,F).$$

For $p \in \mathrm{Symb}_m(E,F)$, we then have the operator, say $\mathrm{Op}(p) : C^\infty(E) \to C^\infty(F)$, given by (2), which extends to a bounded operator $\mathrm{Op}_s(p) : L^2_s(E) \to L^2_{s-m}(F)$, where for any $s \in \mathbb{R}$, $L^2_s(E)$ is the s-th Sobolev space of sections of E, namely the completion of $C^\infty(E)$ with respect to the norm $\|\cdot\|_s$ defined by

$$\|u\|_s^2 := \int_{T^*X}\left(1+|\xi|^2\right)^s \left|u^\wedge(\xi)\right|^2 d\xi.$$

Recall that for $k \in \mathbb{Z}^+$, and $s > n/2+k$, there is a compact inclusion $L^2_s(E) \subset C^k(E)$. For each s, the linear map

$$\mathrm{Op}_s : \mathrm{Symb}_m(E,F) \to \mathrm{B}(L^2_s(E), L^2_{s-m}(F))$$

into the Banach space $\mathrm{B}(L^2_s(E), L^2_{s-m}(F))$ of bounded linear transformations is continuous (see [34], p. 177f). Moreover, for $\varphi \in \mathrm{Symb}_{-\infty}(E,F)$, $\mathrm{Op}_s(\varphi)$ is a compact operator for any $s \in \mathbb{R}$, and $\mathrm{Op}_s(\varphi)\left(L^2_s(E)\right) \subset C^\infty(F)$; i.e., $\mathrm{Op}_s(\varphi)$ is a smoothing operator.

Definition 4.12 We say that $p \in \mathrm{Symb}_m(E,F)$, and the corresponding operator $\mathrm{Op}(p)$, are **elliptic** if for some constant $c > 0$, $p(\xi)^{-1}$ exists for $|\xi| > c$, and for some constant $K > 0$

$$\left|p(\xi)^{-1}\right| \le K\left(1+|\xi|\right)^{-m} \text{ for all } \xi \in T^*X \text{ with } |\xi| > c.$$

We set

$$\mathrm{Ell}_m(E,F) := \{p \in \mathrm{Symb}_m(E,F) : p \text{ is elliptic}\}.$$

For $p \in \mathrm{Ell}_m(E,F)$, there are $q \in \mathrm{Symb}_{-m}(E,F)$, $\varphi_E \in \mathrm{Symb}_{-\infty}(E,E)$ and $\varphi_F \in \mathrm{Symb}_{-\infty}(F,F)$, such that

$$\mathrm{Op}_{s-m}(q) \circ \mathrm{Op}_s(p) = \mathrm{Id}_{L^2_s(E)} + \mathrm{Op}_s(\varphi_E) \text{ and}$$
$$\mathrm{Op}_s(p) \circ \mathrm{Op}_{s-m}(q) = \mathrm{Id}_{L^2_{s-m}(F)} + \mathrm{Op}_{s-m}(\varphi_F).$$

Since $\mathrm{Op}_s(\varphi_E)$ and $\mathrm{Op}_{s-m}(\varphi_F)$ are compact operators, it follows that $\mathrm{Op}_s(p)$ is Fredholm, and hence we may define

$$\mathrm{index}(\mathrm{Op}_s(p)) := \dim\ker(\mathrm{Op}_s(p)) - \dim\mathrm{coker}(\mathrm{Op}_s(p)).$$

Note also that if $\mathrm{Op}_s(p)u \in C^\infty(F)$, then

$$u = \mathrm{Op}_{s-m}(q)\left(\mathrm{Op}_s(p)u\right) - \mathrm{Op}_s(\varphi_E)u \in C^\infty(E).$$

Thus, $\dim\ker\left(\mathrm{Op}_s(p)\right) < \infty$, $\ker\left(\mathrm{Op}_s(p)\right) \subset C^\infty(E)$, and $\ker\left(\mathrm{Op}_s(p)\right)$ is independent of s. As a consequence,

$$\mathrm{index}(\mathrm{Op}_s(p)) = \dim\ker(\mathrm{Op}(p)) - \dim\mathrm{coker}(\mathrm{Op}(p))$$

is independent of s.

5 The Atiyah-Singer Index Theorem

The standard, geometric, elliptic differential operators $P : C^\infty(E) \to C^\infty(F)$ (e.g., see subsection 4.2) typically depend on the choice of Riemannian metric on X, and Hermitian metrics and connections on E and F. Such choices might be called the geometric data defining the operator. As this geometric data varies smoothly, it can be shown that the operator $P^{k+2m} : W^{k+2m}(E) \to W^{k+m}(F)$ varies continuously in $\mathcal{F}\left(W^{k+2m}(E), W^{k+m}(F)\right)$ and hence $\mathrm{index}\left(P^{k+2m}\right)$ is constant. Thus, we expect $\mathrm{index}\, P$ to be a topological attribute of X, E and F, such as a combination of the Euler characteristic and/or other characteristic classes, which is invariant under a change of the geometric data. Since P is defined in terms of geometric data, one is tempted to try to compute $\mathrm{index}\, P$ in terms of this data. The Atiyah-Singer Index Theorem says that this is not only possible, but in fact $\mathrm{index}\, P$ is equal to a topological (integer) invariant known as the *topological index* of P determined (as described in the next subsection) by a suitable homotopy class of the principal symbol $\sigma(P)$. We denote the *topological index* of P by $\mathrm{index}_t\,[\sigma(P)]$. Succinctly, the Atiyah-Singer Index Theorem (or Formula) is $\mathrm{index}_a\, P = \mathrm{index}_t\,[\sigma(P)]$, where $\mathrm{index}_a\, P = \dim\mathrm{Ker}\, P - \dim\mathrm{Coker}\, P$ is the usual index of P as an elliptic operator (possibly pseudo-differential), with the subscript a standing "analytic". In the next subsection, there are several equivalent definitions of $\mathrm{index}_t\,[\sigma(P)]$. The utility and appreciation each definition depends on one's background.

5.1 Descriptions of the topological index

First we describe the so-called *difference construction*. Let (X, Y) be a compact pair where $Y \subseteq X$. Given bundles $\pi_0 : E_0 \to X$ and $\pi_1 : E_1 \to X$ and an isomorphism $\sigma : E_0|_Y \cong E_1|_Y$, we construct an element $\chi(E_0, E_1; \sigma) \in K(X, Y)$. Let $X_0 = X \times \{0\}$ and $X_1 = X \times \{1\}$, and let $Z = X_0 \cup_Y X_1 = X_0 \cup X_1/[(y, 0) \sim (y, 1)]$; i.e., the disjoint union of X_0 and X_1 but with $(y, 0)$ and $(y, 1)$ identified for all $y \in Y$. We define a vector bundle F over Z by $F = E_0 \cup E_1/[(e_0)_y \sim (\sigma(e_0))_y]$, which is the disjoint union of E_0 over X_0 and E_1 over X_1 but with the fibers over $(y, 0)$ and $(y, 1)$ identified via σ_y for all $y \in Y$. We have a retraction

$$\rho : Z \to X_1, \text{ given by } \rho(x, i) = \rho(x, 1) \text{ for } i \in \{0, 1\}.$$

From the sequence $(X_1, \phi) \xrightarrow{i} (Z, \phi) \xrightarrow{j} (Z, X_1)$, we obtain an exact sequence

$$0 \to K(Z, X_1) \xrightarrow{j^*} K(Z) \xrightarrow{i^*} K(X_1) \to 0,$$

and this sequence is split, with $\rho^* : K(X_1) \to K(Z)$ serving as a left inverse of i^*. Let $F_1 \to Z$ be $\rho^*(E_1)$, namely the pull-back of $E_1 \to X_1$ via ρ. Note that $i^*([F] - [F_1]) = 0$ since $F|_{X_1} = F_1|_{X_1}$. Thus, there is $\kappa \in K(Z, X_1)$ with $j^*(\kappa) = [F] - [F_1]$. Since $Z/X_1 \cong X/Y$, $K(Z, X_1) \cong K(X, Y)$, and hence κ corresponds to some element of $K(X, Y)$, which by definition is $\chi(E_0, E_1; \sigma) \in K(X, Y)$. If $Y = \phi$, then we claim that $\chi(E_0, E_1; \sigma) = [E_0] - [E_1]$. Indeed, for $Y = \phi$,

$$\begin{aligned}(F_1 \to Z) &= \rho^*(E_1 \to X_1) = (E_1 \to X_0) \cup (E_1 \to X_1), \text{ whereas} \\ (F \to Z) &= (E_0 \to X_0) \cup (E_1 \to X_1).\end{aligned}$$

Consequently,

$$\begin{aligned}j^*(\kappa) = [F] - [F_1] &= [(E_0 \to X_0) \cup (E_1 \to X_1)] - [(E_1 \to X_0) \cup (E_1 \to X_1)] \\ &= [(E_0 \to X_0) \cup (0 \to X_1)] - [(E_1 \to X_0) \cup (0 \to X_1)] \\ &= j^*([E_0 \to X_0] - [E_1 \to X_0]),\end{aligned}$$

and $[E_0 \to X_0] - [E_1 \to X_0] \in K(Z, X_1) = K(X_0 \cup X_1, X_1)$ corresponds to $[E_0] - [E_1] \in K(X) \cong K(X_0)$.

Definition 5.1 (difference element) Given a compact pair (X, Y), bundles $\pi_0 : E_0 \to X$ and $\pi_1 : E_1 \to X$, and an isomorphism $\sigma : E_0|_Y \cong E_1|_Y$, in the above notation the **difference element** $\chi(E_0, E_1; \sigma) \in K(X, Y)$ is defined by

$$\chi(E_0, E_1; \sigma) := (j^*)^{-1}([F] - [F_1]) \in K(X_0 \cup_Y X_1, X_1) \cong K(X, Y),$$

where we have identified $K(X_0 \cup_Y X_1, X_1)$ with $K(X, Y)$.

Let X be a compact C^∞ n-manifold. Relative to a Riemannian metric for X, we can consider a unit ball bundle $BX := \{\xi \in T^*X \mid |\xi| \leq 1\}$, with boundary the sphere bundle $S(X)$. If $\sigma(P) : T^*X \to \mathrm{Hom}(\pi^*E, \pi^*F)$ is the symbol of an elliptic operator $P : C^\infty(E) \to C^\infty(F)$, then $\sigma(P)$ restricts to an isomorphism $\sigma(P)|_{SX} : \pi^*E|_{SX} \to \pi^*F|_{SX}$. We can then apply the above difference construction to obtain $\chi(\pi^*E, \pi^*F; \sigma(P)|_{SX}) \in K(BX, SX)$. There is an isomorphism

$$K(T^*X) := \widetilde{K}\left((T^*X)^+\right) \cong \widetilde{K}(BX/SX) =: K(BX, SX),$$

and so we may regard $\chi\left(\pi^{*}E,\pi^{*}F;\sigma(P)|_{SX}\right)\in K\left(T^{*}X\right)$.

Definition 5.2 If $P : C^{\infty}(E) \to C^{\infty}(F)$ is an elliptic operator with symbol $\sigma(P) : T^{*}X \to \mathrm{Hom}(\pi^{*}E,\pi^{*}F)$, then the **symbol class** of P is denoted and defined by

$$[\sigma(P)] := \chi\left(\pi^{*}E,\pi^{*}F;\sigma(P)|_{SX}\right)\in K\left(T^{*}X\right).$$

After some preliminary work, we will eventually produce an integer from $[\sigma(P)]$ which will be the desired $\mathrm{index}_t[\sigma(P)]$, the topological index of P.

Let $\pi : V \to X$ be a complex vector bundle, where X is compact. Let $\Lambda^{i}(V)$ be the i-th exterior bundle of V over X. The pull-backs $\pi^{*}\Lambda^{i}(V)$ are then bundles over V, say $\pi^{i} : \pi^{*}\Lambda^{i}(V) \to V$. At each $v \in V$, we have a linear map $\alpha_v^i : \left(\pi^{*}\Lambda^{i}(V)\right)_v \to \left(\pi^{*}\Lambda^{i+1}(V)\right)_v$, given by $\alpha_v^i(w) = v\wedge w$. Since $\alpha_v^{i+1}\circ\alpha_v^i = 0$, we have a complex over V, namely

$$0 \to \pi^{*}\Lambda^{0}(V) \overset{\alpha^{0}}{\to} \pi^{*}\Lambda^{1}(V) \overset{\alpha^{1}}{\to} \cdots \overset{\alpha^{n-1}}{\to} \pi^{*}\Lambda^{n}(V) \to 0,$$

where n is the fiber dimension of V. If $v\neq 0$, $\mathrm{Im}\left(\alpha_v^i\right) = \mathrm{Ker}\left(\alpha_v^{i+1}\right)$, and so the complex is exact over V minus the zero section. Define bundles over V by

$$\pi^{*}\Lambda^{even}(V) := \oplus_{k\ \mathrm{even}}\pi^{*}\Lambda^{k}(V) \ \text{ and }\ \pi^{*}\Lambda^{odd}(V) := \oplus_{k\ \mathrm{odd}}\pi^{*}\Lambda^{k}(V).$$

Relative to a Hermitian structure for V, let $BV := \{v\in V \mid |v|\le 1\}$ and $SV := \{v\in V \mid |v| = 1\}$. There is an isomorphism over SV, namely

$$\left(\alpha^{e}-(\alpha^{e})^{*}\right)|_{SV} := \oplus_{i\ \mathrm{even}}\alpha^{i}|_{SV} : \pi^{*}\Lambda^{even}(V)|_{SV} \cong \pi^{*}\Lambda^{odd}(V)|_{SV}$$

We mention that the adjoint of α_v^i, $\left(\alpha^{i}\right)_v^{*} : \left(\pi^{*}\Lambda^{i+1}(V)\right)_v \to \left(\pi^{*}\Lambda^{i}(V)\right)_v$, is given by interior multiplication by the Hermitian dual v^{*}. Applying the difference construction relative to the compact pair (BV, SV), we obtain the difference element

$$\lambda_V := \chi\left(\pi^{*}\Lambda^{even}(V)|_{BV},\pi^{*}\Lambda^{odd}(V)|_{BV};\left(\alpha^{e}-(\alpha^{e})^{*}\right)|_{SV}\right)\in K(BV,SV)\cong K(V).$$

For the proof of the following result, see e.g., [34], Appendix C.

Theorem 5.3 (Thom Isomorphism Theorem in K-Theory) *For a complex vector bundle* $\pi : V \to X$ *, where* X *is compact, the homomorphism*

$$\varphi : K(X)\to K(V),\ \textit{given by}\ \varphi(a) = (\pi^{*}a)\,\lambda_V,$$

is an isomorphism.

Remark 5.4 To indicate the dependence of φ on $\pi : V \to X$, we use the notation

$$\varphi_{V\to X} : K(X)\to K(V) \text{ or simply } \varphi_V.$$

The analogous result for noncompact X is proven in [32].

Of concern to us, is a special case of this isomorphism which arises as follows. Let X and Y be manifolds and $f : X \to Y$ a smooth, proper embedding. We have $f_{*} : TX \to TY$. While the normal bundle N of X in Y does not have a complex structure, the normal bundle of TX in TY does.

$$TY|_X = TX\oplus N \text{ and } T(TY)|_{TX} = T(TX)\oplus TN$$

Thus, the normal bundle of TX in TY is TN. Roughly, the fiber of $TN \to TX$ over $v \in T_xX$ consists of pairs (u, v) of vectors in N_x, where $u \in N_x$ is a normal vector to $f(X)$ in Y at $f(x) \in Y$, and v is a normal vector to $f_*(T_xX)$ in $T_{f(x)}Y$ at $0_{f(x)} \in T_{f(x)}Y$; thus v can also be regarded as in N_x under the identification of $T_{f(x)}Y$ with $T_{0_{f(x)}}\left(T_{f(x)}Y\right)$. The complex structure maps (u, v) to $(v, -u)$. Thus, we have $\varphi_{TN\to TX} : K(TX) \to K(TN)$. Note that TN can be embedded into TY as an open subset, and this embedding induces an extension homomorphism $h : K(TN) \to K(TY)$. The composition $h \circ \varphi_{TN\to TX}$ gives us a homomorphism $f_! := h \circ \varphi_{TN\to TX} : K(TX) \to K(TY)$. In the case where $Y = \mathbb{R}^{n+m}$, we have $TY = \mathbb{R}^{2(n+m)}$. If $i : \{\mathbf{0}\} \to \mathbb{R}^{n+m}$ is the inclusion of the origin, then $i_! : K(T\{\mathbf{0}\}) \cong K(\mathbb{R}^{2(n+m)})$, and plainly $K(T\{\mathbf{0}\}) \cong \mathbb{Z}$, since $T\{\mathbf{0}\}$ is just a point. Then $i_!^{-1} \circ f_!$ is a homomorphism,

$$\operatorname{index}_t : K(TX) \xrightarrow{f_!} K(\mathbb{R}^{2(n+m)}) \overset{i_!^{-1}}{\cong} K(T\{\mathbf{0}\}) \cong \mathbb{Z}.$$

Of course, some work is needed to show that this is well defined (e.g., independent of the choice of f); see [34], p.244.

Definition 5.5 For an elliptic operator $P : C^\infty(E) \to C^\infty(F)$ with symbol class $[\sigma(P)] \in K(T^*X)$, the **topological index** of P is defined by

$$\operatorname{index}_t[\sigma(P)] := i_!^{-1} f_![\sigma(P)].$$

For those more familiar with cohomology and characteristic classes than with K-theory, the following considerations may provide a better way of computing $\operatorname{index}_t[\sigma(P)]$. First we review definitions of relevant characteristic classes. Let $E \to X$ be a complex vector bundle of dimension m over a C^∞ n-manifold X. The *total Chern class*

$$c(E) = \oplus_{k\geq 0} c_k(E), \text{ where } c_k(E) \in H^{2k}(X, \mathbb{Z}),$$

can be written formally as $\prod_{i=1}^m (1 + x_i)$ so that $c_k(E)$ is the k-th elementary symmetric polynomial in the x_i. The *Chern character* of E is given by

$$ch(E) := \sum\nolimits_{i=1}^m e^{x_i} \in H^*(X, \mathbb{Q}),$$

with the understanding that when $c_k(E)$ is substituted for the k-th elementary symmetric polynomial in the x_i, when $\sum_{i=1}^m e^{x_i}$ is expressed in terms of the elementary symmetric polynomials. Explicitly one computes (where we write $c_k(E)$ simply as c_k)

$$ch_0 = \dim E,\ ch_1 = c_1,\ ch_2 = \tfrac{1}{2}c_1^2 - c_2,\ ch_3 = \tfrac{1}{6}\left(3c_3 - 3c_2c_1 + c_1^3\right),$$
$$ch_4 = \tfrac{1}{24}\left(-4c_4 + 4c_3c_1 + 2c_2^2 - 4c_2c_1^2 + c_1^4\right), \ldots.$$

For two complex bundles E_1 and E_2 over X, we have

$$ch(E_1 \oplus E_2) = ch(E_1) + ch(E_2) \text{ and } ch(E_1 \otimes E_2) = ch(E_1)ch(E_2).$$

Then $E \mapsto ch(E)$ induces a ring homomorphism $ch : K(X) \to H^{even}(X, \mathbb{Q})$. The Todd class of $E \to X$ is defined in the analogous way by

$$\mathbf{Td}(E) = \oplus_{k\geq 0} Td_k(E) = \prod\nolimits_{i=1}^m \frac{x_i}{1 - e^{-x_i}} \in H^{even}(X, \mathbb{Q}).$$

One finds $Td_0 = 1$, $Td_1 = \frac{1}{2}c_1$, $Td_2 = \frac{1}{12}\left(c_2 + c_1^2\right)$, $Td_3 = \frac{1}{24}c_1c_2$,
$$Td_4 = \frac{1}{720}\left(-c_4 + c_3c_1 + 3c_2^2 + 4c_2c_1^2 - c_1^4\right), \ldots .$$

Let $H_c^*(\cdot)$ denote cohomology with compact supports, and let $ch : K^*(TX) \to H_c^{even}(TX)$ be the Chern character ring homomorphism. We can now state

Theorem 5.6 (Cohomological formula for topological index) *For an elliptic operator $P : C^\infty(E) \to C^\infty(F)$ with symbol class $[\sigma(P)] \in K(T^*X)$, the topological index of P is given by*

$$\operatorname{index}_t[\sigma(P)] = (-1)^n \{ch([\sigma(P)])\mathbf{Td}(TX \otimes \mathbb{C})\}[TX], \tag{3}$$

where $ch([\sigma(P)])\mathbf{Td}(TX \otimes \mathbb{C}) \in H_c^*(TX)$, and the right side of (3) is evaluation of this element on the fundamental cycle $[TX]$; i.e., integration over TX.

Proof For a complex vector bundle $\pi : E \to X$ of complex dimension k, we have Thom isomorphisms

$$\varphi_E : K(X) \to K(E) \text{ and } \psi_E : H^*(X) \to H_c^{*+2k}(E;\mathbb{Z}) \cong H^{*+2k}(E, E \setminus \{0\};\mathbb{Z}),$$

in K-theory (see Theorem 5.3) and in cohomology (see [37], §10). By Theorem 5.3, $\varphi_E : K(X) \to K(E)$ is given by $\varphi_E(u) = \pi^*(u)\lambda_E = \pi^*(u)\varphi_E(1)$. If $i_E : X \to E$ is the zero section, with induced map $i_E^* : K(E) \to K(X)$, then

$$\begin{aligned}(i_E^* \circ \varphi_E)(u) &= i_E^*(\pi^*(u)\varphi_E(1)) = i_E^*(\varphi_E(1))\,u \\ &= i_E^*(\lambda_E)\,u = \left([\Lambda^{even}(E)] - [\Lambda^{odd}(E)]\right)u.\end{aligned}$$

For $1 \in H^0(X)$, the Thom class is $\psi_E(1) \in H^{2k}(E)$, and for $\pi^* : H^*(X) \to H^*(E)$ induced by $\pi : E \to X$, we have $\psi_E(u) = \pi^*(u)\psi_E(1)$. For the zero section $i_E : X \to E$, we have $i_E^* : H^*(E) \to H^*(X)$, and for $u \in H^j(X)$,

$$(i_E^*\psi_E)(u) = (i_E^*\psi_E)(u) = i_E^*(\pi^*(u)\psi_E(1)) = (i_E^*\psi_E(1)) \cdot u = \chi(E) \cdot u, \tag{4}$$

where the pull-back $i_E^*\psi_E(1)$ of the Thom class is the Euler class of E, namely $\chi(E) := i_E^*\psi_E(1) \in H^{2k}(X;\mathbb{Z})$. The diagram

$$\begin{array}{ccc} K(X) & \xrightarrow{\varphi_E} & K(E) \\ \downarrow ch_X & & \downarrow ch_E \\ H^{\text{even}}(X) & \xrightarrow{\psi_E} & H^{\text{even}}(E) \end{array}$$

does not commute in general, and this leads to the introduction of the *Chern character defect*, $\mathcal{I}(E) := \psi_E^{-1}(ch_E(\varphi_E(1)))$. We may write $ch_E(\varphi_E(1)) = \psi_E(1) \cdot \pi^*x$ for some $x \in H^*(X)$. We then have

$$\begin{aligned}\chi(E)\mathcal{I}(E) &= (i_E^*\psi_E(1)) \cdot \mathcal{I}(E) = (i_E^*\psi_E(1)) \cdot \psi_E^{-1}ch_E(\varphi_E(1)) \\ &= (i_E^*\psi_E(1)) \cdot \psi_E^{-1}(\psi_E(1) \cdot \pi^*x) = (i_E^*\psi_E(1)) \cdot x \\ &= (i_E^*\psi_E(1)) \cdot i_E^*\pi^*x = i_E^*(\psi_E(1) \cdot \pi^*x) \\ &= i_E^*ch_E(\varphi_E(1)) = ch_X(i_E^*\varphi_E(1)) = ch_X\left([\Lambda^{\text{even}}(E)] - [\Lambda^{\text{odd}}(E)]\right).\end{aligned}$$

Thus, formally

$$\psi_E^{-1}\left(ch_E(\varphi_E(1))\right) = \mathcal{I}\left(E\right) = \frac{ch_X\left([\Lambda^{\text{even}}(E)] - [\Lambda^{\text{odd}}(E)]\right)}{\chi(E)}.$$

Writing $c(E) = \prod_{j=1}^{k}(1+x_j)$, we then have

$$\begin{aligned}\mathcal{I}\left(E\right) &= \frac{ch_X\left([\Lambda_{\mathbb{C}}^{\text{even}}(E)] - [\Lambda_{\mathbb{C}}^{\text{odd}}(E)]\right)}{\chi(E)} \\ &= \frac{\exp\left(\sum_{2p\le k}\left(\sum_{j_1<j_2\cdots<j_{2p}} x_{j_i}\right) - \sum_{2q+1\le k}\left(\sum_{j_1<j_2\cdots<j_{2q+1}} x_{j_i}\right)\right)}{\prod_{j=1}^{k} x_j} \\ &= \prod_{j=1}^{k}\frac{1-e^{x_j}}{x_j}.\end{aligned}$$

Recall that

$$\begin{aligned}\mathbf{Td}\left(E\right) &= \prod_{j=1}^{k}\frac{x_j}{1-e^{-x_j}}, \text{ and so} \\ \mathbf{Td}(\overline{E}) &= \prod_{j=1}^{k}\frac{-x_j}{1-e^{x_j}} = (-1)^k\prod_{j=1}^{k}\frac{x_j}{1-e^{x_j}} = (-1)^k\,\mathcal{I}\left(E\right)^{-1}, \text{ or} \\ \mathcal{I}\left(E\right)^{-1} &= (-1)^k\,\mathbf{Td}(\overline{E}).\end{aligned}$$

For $\xi \in K(X)$, we have

$$\left(\psi_E^{-1}\circ ch_E \circ \varphi_E\right)(\xi)) = \psi_E^{-1}(ch_E(\varphi_E(1)))ch_X\left(\xi\right) = \mathcal{I}\left(E\right)ch_X\left(\xi\right).$$

Indeed,

$$\begin{aligned}&\psi_E^{-1}ch_E(\varphi_E\left(\xi\right)) = \psi_E^{-1}ch_E(\varphi_E(1)\cdot\pi^*\xi) = \psi_E^{-1}\left(ch_E(\varphi_E(1))\,ch_E\left(\pi^*\xi\right)\right) \\ &= \psi_E^{-1}(ch_E(\varphi_E(1))\pi^*\left(ch_X\left(\xi\right)\right)) = \psi_E^{-1}(ch_E(\varphi_E(1))ch_X\left(\xi\right) = \mathcal{I}\left(E\right)ch_X\left(\xi\right).\end{aligned}$$

For $i:\{p_0\}\to\mathbb{R}^{n+m}$, we have the normal bundle $N_0 = \mathbb{R}^{n+m}_{p_0}\to\{p_0\}$ and $TN_0 = T\mathbb{R}^{n+m} = \mathbb{C}^{n+m}\to T\left\{p_0\right\}$. Moreover, there are Thom isomorphisms

$$i_! = \varphi_0 : K(T\left\{p_0\right\}) \to K(T\mathbb{R}^{n+m}) \text{ and } \psi_0 : H_c^*(T\left\{p_0\right\}) \to H_c^*(T\mathbb{R}^{n+m}).$$

Let $u \in K(\mathbb{C}^{n+m}) \cong \mathbb{Z}$, say $u = \varphi_0(\xi)$, for $\xi \in K\left(T\left\{p_0\right\}\right) \cong \mathbb{Z}$. We have

$$\begin{aligned}ch(u)[\mathbb{C}^{n+m}] &= {\psi_0}^{-1}\left(ch\left(u\right)\right) = {\psi_0}^{-1}\left(ch\left(\varphi_0(\xi)\right)\right) \\ &= \mathcal{I}\left(\mathbb{C}^N_{p_0}\right)ch\left(\xi\right) = 1\xi = \xi = {\varphi_0}^{-1}\left(u\right).\end{aligned}$$

Hence, the right-most square in the following diagram commutes:

$$\begin{array}{ccccccc} K(TX) & \xrightarrow{\varphi} & K(TN) & \xrightarrow{h} & K(T\mathbb{R}^{n+m} = \mathbb{C}^{n+m}) \cong \mathbb{Z} & \xrightarrow{\varphi_0^{-1}} & K(T\left\{p_0\right\}) \cong \mathbb{Z} \\ \downarrow ch & & \downarrow ch & & \downarrow ch & & \downarrow ch \\ H_c^*(TX) & \xrightarrow{\psi} & H_c^*(TN) & \xrightarrow{k} & H_c^*(T\mathbb{R}^{n+m} = \mathbb{C}^{n+m}) \cong \mathbb{Z} & \xleftarrow{\psi_0^{-1}} & H_c^*(T\left\{p_0\right\}) \cong \mathbb{Z}, \end{array}$$

and under the identifications with $\mathbb{Z}$ of rings in the right-most square the homomorphisms are all just identities. The middle square also commutes, where h and k are the extension homomorphisms. The left-most square does not commute in general, since there is generally a nontrivial Chern character defect $\mathcal{I}(TN)$ in the relation

$$\psi^{-1} ch(\varphi(\xi)) = \mathcal{I}(TN)\, ch(\xi),\ \xi \in K(TX)$$

In the midst of the computation of $\mathrm{index}_t(P)$ below, we use we use the result $\mathcal{I}(TN) = \mathcal{I}(TX \otimes \mathbb{C})^{-1}$, which is seen as follows. From the relation $\mathcal{I}(E) = \prod_{j=1}^{k} \frac{1-e^{x_j}}{x_j}$, we have $\mathcal{I}(E_1 \oplus E_2) = \mathcal{I}(E_1)\mathcal{I}(E_2)$. Hence, $\mathcal{I}(TN) = \mathcal{I}(TX \otimes \mathbb{C})^{-1}$ follows from the fact that as complex bundles over TX, we have

$$(TX \otimes \mathbb{C}) \oplus TN \cong (\pi^*(TX) \oplus \pi^*(TX)) \oplus TN \cong T(TX) \oplus TN = T\left(T\mathbb{R}^{n+m}\right)|_{TX},$$

which is trivial. Without further interruption, we obtain

$$\begin{aligned}
\mathrm{index}_t(P) &= (i_!)^{-1}\left(f_!\left([\sigma(P)]\right)\right) = (\varphi_0)^{-1}\left((h \circ \varphi)\left([\sigma(P)]\right)\right) \\
&= ch((h \circ \varphi)([\sigma(P)]))[\mathbb{C}^{n+m}] = k\left(ch(\varphi([\sigma(P)]))\right)[\mathbb{C}^{n+m}] \\
&= ch(\varphi([\sigma(P)])\, k_*[\mathbb{C}^{n+m}] = ch(\varphi([\sigma(P)])\,[TN] \\
&= \psi^{-1} ch(\varphi([\sigma(P)])\,[TX] = (\mathcal{I}(TN) ch([\sigma(P)]))\,[TX] \\
&= \left(\mathcal{I}(TX \otimes \mathbb{C})^{-1} ch([\sigma(P)])\right)[TX] \\
&= (-1)^n \left(\mathbf{Td}(\overline{T_{\mathbb{C}} X \otimes \mathbb{C}}) ch_{TX}([\sigma(P)])\right)[TX] \\
&= (-1)^n \left(ch_{TX}([\sigma(P)])\, \mathbf{Td}(T_{\mathbb{C}} X \otimes \mathbb{C})\right)[TX].
\end{aligned}$$

□

Corollary 5.7 *The index of any elliptic differential operator P over an odd dimensional compact n-manifold X is 0.*

Proof Let $f : TX \to TX$ be defined by $f(v) = -v$. Since n is odd, $f_*[TX] = -[TX]$. Also, since P is a differential operator $f^*(\sigma(P))_\xi = \sigma(P)_{-\xi} = (-1)^m \sigma(P)_\xi$, where m is the order of P. If m is even $f^*(\sigma(P)) = \sigma(P)$, while if m is odd, we still have $f^*(\sigma(P)) = -\sigma(P)$ is homotopic to $\sigma(P)$ via $e^{i\pi t}\sigma(P)$, $0 \le t \le 1$. In either case $f^*(\sigma(P))$ is homotopic to $\sigma(P)$. Thus,

$$\begin{aligned}
\mathrm{index}_t[\sigma(P)] &= (-1)^n \{ch([\sigma(P)])\mathbf{Td}(TX \otimes \mathbb{C})\}[TX] \\
&= -(-1)^n \{ch([f^*\sigma(P)])\mathbf{Td}(TX \otimes \mathbb{C})\} f_*[TX] = -\mathrm{index}_t[\sigma(P)],
\end{aligned}$$

where the last equality is due to the invariance of evaluation under diffeomorphism. □

Corollary 5.8 *Let X be a compact oriented manifold. For an elliptic operator $P : C^\infty(E) \to C^\infty(F)$ with symbol class $[\sigma(P)] \in K(T^*X)$, the topological index of P is given by*

$$\mathrm{index}_t(P) = (-1)^{n(n+1)/2}\left(\psi_{TX}^{-1}\left(ch_{TX}([\sigma(P)])\right)\mathbf{Td}(TX \otimes \mathbb{C})\right)[X]. \tag{5}$$

If we assume $\chi(X) \neq 0$, then there is the simpler formula

$$\mathrm{index}_t(P) = (-1)^{n(n+1)/2}\left(\frac{ch([E]-[F])}{\chi(X)}\mathbf{Td}(TX \otimes \mathbb{C})\right)[X]. \tag{6}$$

Proof For a compact oriented manifold X, we have the Thom isomorphism ψ_{TX} : $H^*(X) \to H^*(TX)$, which has the property

$$u[TX] = (-1)^{n(n-1)/2} \left(\psi_{TX}^{-1}(u)\right)[X] \tag{7}$$

for any $u \in H^*(TX)$. The $(-1)^{\frac{1}{2}n(n-1)}$ is explained as follows. Let $x_1, \ldots, x_n$ be positively oriented local coordinates for X. Then the induced oriented local coordinates for $v_1\partial x_1 + \cdots + v_n\partial x_n \in TX$ are $x_1, \ldots, x_n, v_1, \ldots, v_n$. However, for the orientation that we chose for TX (even if X is not orientable), the coordinates $x_1, v_1, \ldots, x_n, v_n$ are positively oriented. Transforming $x_1, \ldots, x_n, v_1, \ldots, v_n$ to $x_1, v_1, \ldots, x_n, v_n$ requires $(n-1) + (n-2) + \cdots + 1 = n(n-1)/2$ transpositions. We use (7) with $u = (-1)^n \{ch([\sigma(P)])\mathbf{Td}(TX \otimes \mathbb{C})\}$, noting that $n + n(n-1)/2 = n(n+1)/2$, to obtain (5). To obtain (6), recall from (4) that if $i : X \to TX$ is the zero section, then

$$\chi(TX)\cdot\psi_{TX}^{-1}(ch_{TX}([\sigma(P)]) = i^* ch_{TX}([\sigma(P)]) = ch_X(i^*[\sigma(P)]) = ch_X([E]-[F]).$$

□

5.2 On the K-theoretic proof of Atiyah and Singer

Here we provide an outline of the K-theoretic embedding proof of the Atiyah-Singer Index Formula (Theorem 5.9 below). This proof appeared in [12], after announcement of the main result in [11] and the exposition [41] of the unpublished proof involving cobordism.

Theorem 5.9 (Atiyah-Singer Index Theorem) *Let $P : C^\infty(E) \to C^\infty(F)$ be an elliptic operator (possibly pseudo-differential) where $E \to X$ and $F \to X$ are complex vector bundles over the compact n-manifold X. Let $[\sigma(P)] \in K(T^*X) \cong K(TX)$ be the symbol class of P, as in Definition 5.2. Then*

$$\mathrm{index}_a\, P = \mathrm{index}_t\, [\sigma(P)].$$

Here $\mathrm{index}_a\, P = \dim \mathrm{Ker}\, P - \dim \mathrm{Coker}\, P$ *is the analytic index of P, and in the notation preceding Definition 5.5,*

$$\mathrm{index}_t\, [\sigma(P)] := i_!^{-1} f_! [\sigma(P)]$$

is the topological index of P (or its symbol class) which is also given cohomologically by Theorem 5.6 or Corollary 5.8 when applicable.

Outline of proof. It is not difficult to prove that any $u \in K(TX)$ is of the form $[\sigma_m(p)]$ for some $p \in \mathrm{Ell}_m(E,F) \in C^\infty(\mathrm{Hom}(\pi^*E,\pi^*F))$, where $\pi : TX \to X$ and $\sigma_m(p)$ is the m-th order asymptotic symbol of p; indeed, m can be chosen freely; e.g., see [12, p.492], [34, p.245]). With some work, the analytical index of $u \in K(TX)$ is shown to be well-defined, independent of the choice of p, by

$$\mathrm{index}_a\, u := \mathrm{index}_a\, \mathrm{Op}(p);$$

e.g., see [12, p.518], [34, p.246]. We define $\mathrm{index}_t\, u$ in the same way that $\mathrm{index}_t[\sigma(P)]$ was defined, namely $\mathrm{index}_t\, u = i_!^{-1} f_! u$ where $f : X \to Y := \mathbb{R}^{n+m}$ is a proper embedding. The goal is to show $\mathrm{index}_a\, u = \mathrm{index}_t\, u$ for all $u \in K(TX)$. To prove this, one "only" needs to show that $\mathrm{index}_a(u) = \mathrm{index}_a(f_! u)$.

Then $\mathrm{index}_a(u) = \mathrm{index}_a(f_! u) = \mathrm{index}_a\left(\left(i_! i_!^{-1}\right)(f_! u)\right) = \mathrm{index}_a\left(i_!\left(i_!^{-1} f_! u\right)\right)$

$$= \operatorname{index}_a(i_!^{-1} f_! u) = i_!^{-1} f_! u = \operatorname{index}_t(u).$$

For last two equalities, note that for a singleton $\{x_0\}$, $T\{x_0\}$ is also a point, and so $K(T\{x_0\})$ is identified with $\mathbb{Z}$. The elliptic operator over $\{x_0\}$ (as well as its symbol) associated with $i_!^{-1} f_! u \in K(T\{0\})$ is just a linear map $P : E \to F$ between vector spaces over $\{x_0\}$ (or $T\{x_0\}$), and

$$\operatorname{index}_a P = \dim \operatorname{Ker} P - \dim \operatorname{Coker} P = \dim E - \dim F = \operatorname{index}_t P.$$

Since $f_! = h \circ \varphi_{TN \to TX}$ is a composition of two maps, the proof that $\operatorname{index}_a(u) = \operatorname{index}_a(f_! u)$ has two parts, namely

1. $\operatorname{index}_a(\varphi_{TN \to TX}(u)) = \operatorname{index}_a(u)$ and
2. $\operatorname{index}_a(\varphi_{TN \to TX}(u)) = \operatorname{index}_a(h(\varphi_{TN \to TX}(u)))$.

Part 2 follows from the Excision Property and its proof is easier than part 1 (e.g., see [34, p.248 and 254], [12, p.522]). Part 1 is a consequence of the multiplicative property of the analytic index with respect to embeddings. We will elucidate this property in the next subsection. Once this is done, the proof of Part 1 and the is completed by the following

$$\begin{aligned} \operatorname{index}_a(\varphi_{TN \to TX}(u)) &= \operatorname{index}_a((\pi_{TN}^* u)\lambda_{TN}) = \operatorname{index}_a(u \cdot i_! 1) \\ &= \operatorname{index}_a\left(\left(\operatorname{index}_{\mathrm{O}(m)} i_! 1\right) \cdot u\right) = \operatorname{index}_a(u), \end{aligned}$$

in which the notation is yet to be introduced. The multiplicative property is used for the third equation and $\operatorname{index}_{\mathrm{O}(m)}(i_! 1)$ is the $\mathrm{O}(m)$-equivariant analytic index of the Bott element $i_! 1 \in K(T^* S^m)$ where i is the inclusion of a point in S^m. Since it can be shown that $\operatorname{index}_{\mathrm{O}(m)}(i_! 1) = 1$ in the representation ring $R(\mathrm{O}(m))$ (see [12, p.505 and 524]), the last equality follows.

5.3 The multiplicative property

Consider an embedding $f : X \to Y$ of the compact manifold X into some manifold Y (say $\mathbb{R}^{n'}$ or $S^{n'}$). From an elliptic pseudo-differential operator on X, we will construct an appropriate elliptic pseudo-differential operator, with the same index, on a suitably compactified tubular neighborhood, say S, of $f(X)$ in Y. In other words, from a symbol

$$a \in \operatorname{Ell}_m(E, F) \subset C^\infty(T^* X, \operatorname{Hom}(\pi_X^* E, \pi_X^* F)), \text{ where } \pi_X : T^* X \to X$$

with associated operator $\operatorname{Op}(a) : C^\infty(E) \to C^\infty(F)$, we construct suitable complex vector bundles $\widetilde{E} \to S$ and $\widetilde{F} \to S$ and a symbol

$$c \in \operatorname{Ell}_m(\widetilde{E}, \widetilde{F}) \subset C^\infty(T^* S, \operatorname{Hom}(\pi_S^* \widetilde{E}, \pi_S^* \widetilde{F})), \tag{8}$$

with associated operator $\operatorname{Op}(c) : C^\infty(\widetilde{E}) \to C^\infty(\widetilde{F})$; here $\pi_S : T^* S \to S$. The essential ingredient which is needed to produce c is an equivariant K- theory element $b \in K_{\mathrm{O}(m)}(T^* S^m)$, where $m = n' - n$ and S^m is the unit m-sphere. The choice of $b \in K_{\mathrm{O}(m)}(T^* S^m)$ which yields $\operatorname{index} \operatorname{Op}(c) = \operatorname{index} \operatorname{Op}(a)$ is essentially the famous generating Bott element, but b will be arbitrary here. We begin with a short review of relevant equivariant K-theory for those who desire it. The work of Graeme Segal [43] is an excellent, authoritative exposition of the foundations of equivariant K-theory.

Let G be a group which acts to the left on X, via a $L : G \times X \to X$. We write $g \cdot x = L_g(x) = L(g, x)$. Let $\pi : E \to X$ be a complex vector bundle over X and suppose that there is a left action of G on E such that $\pi(g \cdot e) = g \cdot \pi(e)$ and $e \mapsto g \cdot e$ is linear on each fiber E_x. Then $\pi : E \to X$ is called a G-vector bundle. As an example, if X is a manifold and G acts on X smoothly, then the action on $T_{\mathbb{C}}X := \mathbb{C} \otimes TX$ given by $v \mapsto d(L_g)(v)$ for $v \in T_{\mathbb{C}}X$ makes $T_{\mathbb{C}}X \to X$ a G-vector bundle. More generally, $\Lambda^k(T_{\mathbb{C}}X) \to X$ is a G-vector bundle. A morphism from G-vector bundle $\pi_1 : E_1 \to X$ to G-vector bundle $\pi_2 : E_2 \to X$ is a vector bundle morphism (linear on fibers) $\varphi : E_1 \to E_2$ such that $\varphi(g \cdot e) = g \cdot \varphi(e)$. An isomorphism of G-vector bundles is a morphism which is bijective. The direct sum of G-vector bundles is clearly a G-vector bundle and this operation induces an abelian semi-group structure on the set of isomorphism classes of G-vector bundles. We can then form the associated abelian group $K_G(X)$ via the Grothendieck construction. Moreover, the tensor product of G-vector bundles yields a G-vector bundle, and this induces a ring structure on $K_G(X)$. For a homogeneous space G/H where H is a closed subgroup of G, there is a ring isomorphism $K_G(G/H) \cong R(H) :=$ the representation ring of H. Recall that $R(H)$ is the Grothendieck ring obtained from the abelian semi-group of equivalence classes of representations of H with addition induced by the direct sum. Tensor product of representations induces a multiplication on $R(H)$ making it a ring. More concisely, $R(H) = K_H(\{\text{point}\})$. As with ordinary K-theory, an element of $K_G(X)$ can also be described as equivalence classes of G-equivariant morphisms $E \to F$ of G-bundles which are isomorphisms outside of a compact support (i.e., morphisms with compact support).

We proceed with the construction of $c \in \mathrm{Ell}_m(\widetilde{E}, \widetilde{F})$ in (8). Let $\pi_P : P \to X$ be the principal $\mathrm{O}(m)$-bundle of orthonormal frames of the normal bundle $N \to X$ for the embedding $f : X \to Y$, where $\dim X = n$ and $\dim Y = n'$. We regard a frame $p \in P_x$ as a linear isometry $p : \mathbb{R}^m \to N_x$, where $m = n' - n$ and N_x is the fiber of the normal bundle at $x \in X$. In terms of associated bundles, we have $N = P \times_{\mathrm{O}(m)} \mathbb{R}^m = (P \times \mathbb{R}^m) / \mathrm{O}(m)$, where $\mathrm{O}(m)$ acts on $P \times \mathbb{R}^m$ via $(p, v) \cdot A := (p \circ A, A^{-1}v)$. Note that $\mathrm{O}(m)$ also acts on $\mathbb{R}^{m+1} = \mathbb{R}^m \times \mathbb{R}$ via $A \cdot (v, a) = (A(v), a)$, and the m-sphere $S^m \subset \mathbb{R}^{m+1}$ is invariant under this action with two fixed points, the poles $(\mathbf{0}, \pm 1) \in S^m$. Let

$$S := P \times_{\mathrm{O}(m)} S^m \text{ and let } Q : P \times S^m \to P \times_{\mathrm{O}(m)} S^m = (P \times S^m) / \mathrm{O}(m)$$

be the quotient map. We may regard $\pi_S : S \to X$ as the m-sphere bundle over X obtained by compactification of the normal bundle N via adjoining the section at infinity. Choose a $\mathrm{so}(m)$-valued connection 1-form ω on P; there is actually a natural ω induced by $f : X \to Y$ and a given Riemannian metric on Y. Then we have an $\mathrm{O}(m)$-invariant distribution H of horizontal subspaces (i.e., $H_p = \mathrm{Ker}\,\omega_p$) on P and hence on $P \times S^m$. By the $\mathrm{O}(m)$-invariance of H, $Q_*(H)$ is a well defined distribution on S. Moreover, since $\pi_{S*}Q_*(H_p) = \pi_{P*}(H_p) = T_{\pi(p)}X$, $Q_*(H)$ is complementary to the vertical distribution V_S of tangent spaces of the fibers of $\pi_S : S \to X$. We denote $Q_*(H)$ by H_S. Thus, we have a splitting

$$TS = V_S \oplus H_S = V_S \oplus Q_*(H). \tag{9}$$

We also have $T^*S = \widetilde{V}_S^* \oplus \widetilde{H}_S^*$, where

$$\widetilde{H}_S^* := \{\alpha \in T^*S : \alpha(V_S) = 0\} \text{ and}$$

$$\widetilde{V}_S^* := \{\beta \in T^*S : \beta(H_S) = 0\} \cong P \times_{\mathrm{O}(m)} T^*S^m.$$

In view of the splitting (9), there are identifications $\widetilde{V}_S^* \cong V_S^* := (V_S)^*$ and $\widetilde{H}_S^* \cong H_S^* := (H_S)^*$. Note that $\mathrm{O}(m)$ acts on the sphere S^m, and hence on T^*S^m via pull-back of covectors. Thus, we may consider $K_{\mathrm{O}(m)}(T^*S^m)$. The projection $P \times T^*S^m \to T^*S^m$ induces a map $K_{\mathrm{O}(m)}(T^*S^m) \to K_{\mathrm{O}(m)}(P \times T^*S^m)$. Moreover, there is the general fact that if G acts freely on X, then the projection $Q : X \to X/G$ induces an isomorphism $Q^* : K(X/G) \cong K_G(X)$ (see [43], p. 133). Thus, we have

$$K_{\mathrm{O}(m)}(T^*S^m) \to K_{\mathrm{O}(m)}(P \times T^*S^m) \overset{(Q^*)^{-1}}{\underset{\to}{\cong}} K(P \times_{\mathrm{O}(m)} T^*S^m) = K(V_S^*). \quad (10)$$

We define

$$K(T^*X) \otimes K(V_S^*) \to K(T^*S), \quad (11)$$

as follows. If $E \to T^*X$ and $F \to V_S^*$ are complex vector bundles, then for $\alpha' \in T^*X$ and $\beta' \in V_S^*$, we have unique $\alpha \in \widetilde{H}_S^*$ and $\beta \in \widetilde{V}_S^*$ such that $\alpha(v) = \alpha'\left((\pi_S)_*(v)\right)$ for v in TS, and $\beta|_{V_S} = \beta'$ and $\beta(H_S) = 0$. Then $E_{\alpha'} \otimes F_{\beta'}$ is the fiber of a bundle over T^*S at the point $\alpha + \beta$. Thus, we have $K(T^*X) \otimes K(V_S^*) \to K(T^*S)$ induced by $[E] \otimes [F] \mapsto [E \otimes F]$. Using the homomorphisms (10) and (11), we then have

$$K(T^*X) \otimes K_{\mathrm{O}(m)}(T^*S^m) \to K(T^*X) \otimes K(V_S^*) \to K(T^*S). \quad (12)$$

For any representation $\rho : \mathrm{O}(m) \to \mathrm{GL}(\mathbb{C}^q)$, we have the associated vector bundle $P \times_\rho \mathbb{C}^q \to X$. Let $R(\mathrm{O}(m))$ be the representation ring of $\mathrm{O}(m)$. The assignment $\rho \mapsto P \times_\rho \mathbb{C}^q$ extends to a ring homomorphism $R(\mathrm{O}(m)) \to K(X)$, which is to say that $K(X)$ is a $R(\mathrm{O}(m))$-module. Moreover, recall that $K(T^*X)$ is a $K(X)$-module via $u \cdot v = (\pi^*u)\, v$. Thus, ultimately $K(T^*X)$ is an $R(\mathrm{O}(m))$-module. We are now in a position to state

Theorem 5.10 (**The Multiplicative Property**) *For* $v \in K_{\mathrm{O}(m)}(T^*S^m)$ *and* $u \in K(T^*X)$*, we have* $u \cdot v \in K(T^*S)$*, via* (12)*. Moreover,*

$$\mathrm{index}_a(u \cdot v) = \mathrm{index}_a\left(\left(\mathrm{index}_{\mathrm{O}(m)}\, v\right) \cdot u\right),$$

where $\left(\mathrm{index}_{\mathrm{O}(m)}\, v\right) \cdot u \in K(T^*X)$ *makes sense since* $\mathrm{index}_{\mathrm{O}(m)}\, v \in R(\mathrm{O}(m))$*, and as we have just noted,* $K(T^*X)$ *is an* $R(\mathrm{O}(m))$*-module. In particular, if* $\mathrm{index}_{\mathrm{O}(m)}\, v = 1 \in R(\mathrm{O}(m))$*, then* $\mathrm{index}_a(u \cdot v) = \mathrm{index}_a\, u$*.*

Proofs can be found in [34, p.252] and [12, p.526-9].

5.4 Heat kernel methods for twisted Dirac operators

The classical geometric operators such as the Hirzebruch signature operator, the de Rham operator, the Dolbeaut operator and even the Yang-Mills operator can all be locally expressed in terms of chiral halves of twisted Dirac operators. Thus, we will focus on index theory for such operators. As we will show in the next subsection, the index of any of these classical operators (and *their* twists) can then be obtained from the Local Index Theorem for twisted Dirac operators. This theorem supplies a well-defined n-form on X, whose integral is the index of the twisted Dirac operator. This n-form (or "index density") is expressed in terms of forms for characteristic classes which are polynomials in curvature forms. The Index Theorem thus obtained then becomes a formula that relates a global invariant quantity, namely the index of an operator, to the integral of a local quantity involving curvature. This is in the spirit of the Gauss-Bonnet Theorem which can be considered a special case. We begin with an introduction and/or review of Clifford algebras and spinors.

Let V be a real vector space with positive-definite inner product $\langle \cdot, \cdot \rangle$.

Definition 5.11 The **Clifford algebra** $Cl\,(V)$ is the real algebra generated by V and $\mathbb{R}$ with the relation

$$vw + wv = -2\,\langle v, w\rangle\,, \text{ for all } v, w \in V.$$

Note that the product of v and w in $Cl\,(V)$ is denoted by the plain juxtaposition vw. Also, $v^2 := vv = -\langle v, v\rangle = -\|v\|^2$, and $vw = -wv$ if $\langle v, w\rangle = 0$.

Let $\Lambda^*(V) = \oplus_{k=1}^{n}\Lambda^k(V)$ be the exterior algebra of V. While $\Lambda^*(V)$ is not isomorphic to $Cl\,(V)$ as an algebra, there is a linear isomorphism of vector spaces

$$\mathcal{L} : \Lambda^*(V) \cong Cl\,(V) \text{ with } \mathcal{L}\,(e_{i_1} \wedge \cdots \wedge e_{i_k}) := e_{i_1} \cdots e_{i_k},\ (i_1 < \cdots < i_k)\,,$$

where we let $\{e_1, \ldots, e_n\}$ be an orthonormal basis of V. It can be shown that $\mathcal{L}$ is O(n)-equivariant and independent of the choice of orthonormal basis. Moreover via $\mathcal{L}$, the natural inner product on $\Lambda^*(V)$ gives us an inner product and norm on $Cl\,(V)$. There is an exponential map $\exp : Cl\,(V) \to Cl\,(V)$ given by $\exp(x) = \sum_{k=0}^{\infty}\frac{1}{k!}x^k,\ x \in Cl\,(V)\,,$ which converges, since $\left\|x^k\right\| \le C^k\,\|x\|^k$ for some constant C depending on n but not on x.

We define the *bracket* (or *commutator*) of any $x,\, y \in Cl\,(V)\,,$ by $[x, y] = xy - yx$. The linear subspace $\mathcal{L}\left(\Lambda^2(V)\right)$ is closed under bracket. Indeed, for $A = [a_{ij}] \in \mathfrak{so}\,(n)$ (i.e., A, B anti-symmetric), we have that

$$c' : \mathcal{L}\left(\Lambda^2(V)\right) \cong \mathfrak{so}\,(n)\,, \text{ given by } c'(-\tfrac{1}{4}\textstyle\sum_{i,j} a_{ij}e_ie_j) := A \tag{13}$$

is an isomorphism of Lie algebras. We define $\mathrm{Spin}\,(n) := \exp\left(\mathcal{L}\left(\Lambda^2(V)\right)\right)$. The following result is not hard to verify:

Proposition 5.12 *For $g \in \mathrm{Spin}\,(n)$ and $v \in V = \mathcal{L}\left(\Lambda^1(V)\right)$, let*

$$c\,(g)\,(v) := gvg^{-1} \in Cl\,(V)\,.$$

Then $c\,(g)\,(v) \in \mathcal{L}\left(\Lambda^1(V)\right) = V$. Also, $c\,(g) \in \mathrm{SO}\,(n) := \mathrm{SO}\,(V)$ and

$$c : \mathrm{Spin}\,(n) \to \mathrm{SO}\,(n)\,.$$

is a double covering homomorphism (universal for $n \ge 3$). Moreover, for c' defined as in (13), we have

$$c\,(\exp a) = \exp\left(c'\,(a)\right),$$

and so c' is the Lie algebra homomorphism for c.

Besides the vector representation $c : \mathrm{Spin}\,(n) \to \mathrm{SO}\,(n)$, there are fundamental spinor representations, which we will describe. Since the index of an elliptic differential operator on a compact, odd-dimensional manifold is always 0, for simplicity we assume that n is even, say $n = 2m$. Then there is a unique (up to equivalence) irreducible representation (homomorphism of algebras over $\mathbb{R}$)

$$\rho : Cl_{2m} \to \mathrm{End}\,(\Sigma_{2m})\,,$$

where $\mathrm{End}\,(\Sigma_{2m})$ is the algebra of $\mathbb{C}$-linear endomorphisms of some complex vector space Σ_{2m}, the elements of which are called *spinors*. Here "irreducible" means that Σ_{2m} has no proper subspace which is invariant under all operators in $\rho\,(Cl_{2m})$. In the following we give an explicit construction of Σ_{2m} and ρ.

Let $\langle \cdot, \cdot \rangle$ be the standard Hermitian inner product on $\mathbb{C}^m$ given by $\langle z, w \rangle = \sum_{k=1}^{m} z_k \overline{w_k}$. We identify $\mathbb{C}^m$ with $\mathbb{R}^n = \mathbb{R}^{2m}$, and for $w \in \mathbb{C}^m$, we have $\mathbb{C}$-linear function

$$w\wedge : \Lambda^k\,(\mathbb{C}^m) \to \Lambda^{k+1}\,(\mathbb{C}^m)$$

given by $\alpha \mapsto w \wedge \alpha$ for $\alpha \in \Lambda^k\,(\mathbb{C}^m)$. Moreover, there is a $\mathbb{C}$-linear function

$$w \llcorner : \Lambda^k\,(\mathbb{C}^m) \to \Lambda^{k-1}\,(\mathbb{C}^m)\,, \text{ defined via}$$

$$w \llcorner (v_1 \wedge \cdots \wedge v_k) := \sum_{j=1}^{k} (-1)^{j+1} \langle v_j, w \rangle\, v_1 \wedge \cdots \wedge \widehat{v_j} \wedge \cdots \wedge v_k,$$

where $\widehat{v_j}$ means that the factor v_j is omitted. While $w\llcorner$ is $\mathbb{C}$-linear, the function $\mathbb{C}^m \to \mathrm{End}\,(\Lambda^*\,(\mathbb{C}^m))$ given by $w \mapsto w\llcorner$ is $\mathbb{R}$-linear (but $\mathbb{C}$-conjugate linear). There is a Hermitian inner product $\langle \cdot, \cdot \rangle$ on $\Lambda^k\,(\mathbb{C}^m)$ induced by that on $\mathbb{C}^m$, such that $\{e_{i_1} \wedge \cdots \wedge e_{i_k} : i_1 < \cdots < i_k\}$ is an orthonormal basis for $\Lambda^k\,(\mathbb{C}^m)$ if $e_1, \ldots, e_m$ is an orthonormal basis for $\mathbb{C}^m$. Relative to this inner product, it is easy to check that $w\llcorner$ and $w\wedge$ are adjoints, and to verify

Proposition 5.13 *Let $\rho_1 : \mathbb{C}^m \to \mathrm{End}\,(\Lambda^*\,(\mathbb{C}^m))$ be given by*

$$\rho_1\,(w)\,(\alpha) := (w \wedge\; - w\llcorner)\,(\alpha) = w \wedge \alpha - w\llcorner\alpha.$$

Then ρ_1 uniquely extends to an $\mathbb{R}$-linear homomorphism

$$\rho : Cl_{2m} \to \mathrm{End}\,(\Lambda^*\,(\mathbb{C}^m))\,,$$

of algebras over $\mathbb{R}$.

Let $\mathbb{C}l_{2m} := \mathbb{C} \otimes_{\mathbb{R}} Cl_{2m}$ be the complex Clifford algebra. To see that the complex linear extension of ρ, say $\rho_{\mathbb{C}} : \mathbb{C}l_{2m} \to \mathrm{End}\,(\Lambda^*\,(\mathbb{C}^m))$, is an isomorphism of algebras over $\mathbb{C}$, it is convenient to introduce more notation. Let $(f_1, \cdots, f_m)$ be an orthonormal basis of $\mathbb{C}^m$, then

$$(e_1, \cdots, e_{2m}) := (f_1, if_1, \cdots, f_m, if_m)$$

is an oriented, orthonormal basis of $\mathbb{R}^{2m}$, and the *complex volume element* is

$$\omega_{\mathbb{C}} := i^m e_1 \cdots e_{2m} \in \mathbb{C}l_{2m};$$

this is independent of the choice of oriented orthonormal basis of $\mathbb{R}^{2m}$, and it is easy to check that $\omega_{\mathbb{C}}^2 = 1$ and hence $\rho_{\mathbb{C}}(\omega_{\mathbb{C}})^2 = \rho_{\mathbb{C}}(\omega_{\mathbb{C}}^2) = \rho_{\mathbb{C}}(1) = \mathrm{Id}$. Since $\rho_1(w) = w\wedge - w\llcorner$ is the difference between an operator and its adjoint, $\rho_1(w)$ is skew-adjoint. We define skew-adjoint operators $\gamma_j := \rho_{\mathbb{C}}(e_j) = \rho(e_j)$, for $j \in \{1,\ldots,2m\}$. In harmony with the physics literature, we set $\gamma_{n+1} = \gamma_{2m+1} := \gamma_1\cdots\gamma_n$, so that $\rho_{\mathbb{C}}(\omega_{\mathbb{C}}) = i^m\gamma_{2m+1}$. One can show that $\rho_{\mathbb{C}}(\omega_{\mathbb{C}})$ is self-adjoint. Since $\rho_{\mathbb{C}}(\omega_{\mathbb{C}})^2 = \mathrm{Id}$, the eigenvalues of $\rho_{\mathbb{C}}(\omega_{\mathbb{C}})$ are ± 1. As $\rho_{\mathbb{C}}(\omega_{\mathbb{C}})$ is self-adjoint, the eigenspaces of $\rho_{\mathbb{C}}(\omega_{\mathbb{C}})$, say

$$\Sigma_{2m}^{+} := (\rho_{\mathbb{C}}(\omega_{\mathbb{C}}) + \mathrm{Id})\,\Lambda^*(\mathbb{C}^m) \;\text{ and }\; \Sigma_{2m}^{-} := (\rho_{\mathbb{C}}(\omega_{\mathbb{C}}) - \mathrm{Id})\,\Lambda^*(\mathbb{C}^m),$$

are orthogonal. One finds $\rho_{\mathbb{C}}(\omega_{\mathbb{C}})\,|_{\Lambda^l(\mathbb{C}^m)} = (-1)^l\,\mathrm{Id}$, and so

$$\Sigma_{2m}^{+} = \Lambda^{\text{even}}(\mathbb{C}^m) := \bigoplus\nolimits_{l\text{ even}} \Lambda^l(\mathbb{C}^m), \text{ while}$$
$$\Sigma_{2m}^{-} = \Lambda^{\text{odd}}(\mathbb{C}^m) := \bigoplus\nolimits_{l\text{ odd}} \Lambda^l(\mathbb{C}^m).$$

Using this, we have $\gamma_j\left(\Sigma_{2m}^{\pm}\right) = \Sigma_{2m}^{\mp}$, and indeed $\gamma_j : \Sigma_{2m}^{\pm} \cong \Sigma_{2m}^{\mp}$ with inverse $-\gamma_j$, and so $\rho(\mathfrak{spin}(2m))\left(\Sigma_{2m}^{\pm}\right) \subset \Sigma_{2m}^{\pm}$. Moreover, for $j \neq k$,

$$\left(\gamma_j\circ\gamma_k\right)^* = \gamma_k^*\circ\gamma_j^* = -\gamma_k\circ-\gamma_j = -\gamma_j\circ\gamma_k \text{ and}$$
$$\mathrm{Tr}\left(\gamma_j\circ\gamma_k\right) = -\mathrm{Tr}\left(\gamma_k\circ\gamma_j\right) = -\mathrm{Tr}\left(\gamma_j\circ\gamma_k\right).$$

Thus, $\rho(\mathrm{Spin}(2m)) \subset \mathrm{SU}(\Lambda^*(\mathbb{C}^m))$, since the elements of $\rho(\mathfrak{spin}(2m))$ are skew-adjoint and traceless. In summary, $\rho : \mathrm{Spin}(2m) \to \mathrm{SU}(\Lambda^*(\mathbb{C}^m))$ is the orthogonal direct sum of two special unitary "half-spinor" or "chiral" representations

$$\rho^{\pm} : \mathrm{Spin}(2m) \to \mathrm{SU}\left(\Sigma_{2m}^{\pm}\right).$$

Definition 5.14 Let

$$\Sigma_{2m} := \Lambda^*(\mathbb{C}^m) \text{ and } \pi^{\pm} := \tfrac{1}{2}(\rho_{\mathbb{C}}(\omega_{\mathbb{C}}) \pm \mathrm{Id}) : \Sigma_{2m} \to \Sigma_{2m}^{\pm}.$$

The **supertrace** of an endomorphism $A \in \mathrm{End}(\Sigma_{2m})$ is

$$\begin{aligned}\mathrm{Str}(A) &:= Tr\left(\pi^{+}\circ\left(A|_{\Sigma_{2m}^{+}}\right)\right) - Tr\left(\pi^{-}\circ\left(A|_{\Sigma_{2m}^{-}}\right)\right)\\ &= Tr\left(A\circ\rho_{\mathbb{C}}(\omega_{\mathbb{C}})\right) = i^m Tr\left(A\circ\gamma_{2m+1}\right).\end{aligned}$$

The following result is crucial for evaluating the local index density of the twisted Dirac operator; an easy proof is in [16, p.2056].

Proposition 5.15 *For $k \in \{1,\ldots,2m\}$ with $j_1, j_2, \ldots, j_k$ distinct, we have*

$$\begin{aligned}&\mathrm{Tr}\left(\gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_k}\right) = 0, \textit{ and}\\ &\mathrm{Str}\left(\gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_k}\right) = i^m\,\mathrm{Tr}\left(\gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_k}\gamma_{2m+1}\right)\\ &\qquad = \begin{cases} 0 & \textit{if } k < 2m\\ (-2i)^m\,\varepsilon_{j_1\cdots j_{2m}} & \textit{if } k = 2m.\end{cases}\end{aligned}$$

The 2^n endomorphisms consisting of Id *and those $\gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_k}$ with $j_1 < j_2 < \cdots < j_k$, $k \in \{1,\ldots,2m\}$, form a basis of* $\mathrm{End}(\Sigma_{2m})$.

Corollary 5.16 *The representation $\rho_{\mathbb{C}} : \mathbb{C}l_{2m} \to \mathrm{End}\,(\Sigma_{2m})$ is irreducible and an isomorphism of complex algebras. Moreover, $\rho^{\pm} : \mathrm{Spin}\,(2m) \to \mathrm{End}\left(\Sigma^{\pm}_{2m}\right)$ are irreducible, and since $\rho\,(e_1 e_2 \cdots e_{2m-1} e_{2m}) = \gamma_{2m+1} = \pm i^{-m}$ on $\Sigma^{\pm}_{2m}$ they are inequivalent.*

Proposition 5.17 *Let $R : \mathbb{C}l_{2m} \to \mathrm{End}\,(V)$ be a finite-dimensional representation. Then $V = \bigoplus_{k=1}^{N} W_k$ where $W_1, \ldots, W_N$ of V are invariant subspaces, such that $R_k : \mathbb{C}l_{2m} \to \mathrm{End}\,(W_k)$ defined by $R_k\,(\alpha) = R\,(\alpha)\,|_{W_k}$ is equivalent to $\rho_{\mathbb{C}} : \mathbb{C}l_{2m} \to \mathrm{End}\,(\Lambda^*\,(\mathbb{C}^m))$ for all $k = \{1, \ldots, N\}$. In particular, all irreducible representations of $\mathbb{C}l_{2m}$ are equivalent to $\rho_{\mathbb{C}}$. Moreover, let*

$$\mathrm{Hom}_0\,(\Sigma_{2m}, V) := \{F \in \mathrm{Hom}\,(\Sigma_{2m}, V) : F\,(\rho_{\mathbb{C}}\,(\alpha)\,(w)) = R\,(\alpha)\,(F\,(w))\}$$

be the subspace of $\mathrm{Hom}\,(\Sigma_{2m}, V)$ of $\mathbb{C}l_{2m}$-equivariant linear maps; note that $\mathbb{C}l_{2m}$ acts trivially on $\mathrm{Hom}_0\,(\Sigma_{2m}, V)$. There is then an isomorphism of $\mathbb{C}l_{2m}$-modules

$$\Phi : \mathrm{Hom}_0\,(\Sigma_{2m}, V) \otimes \Sigma_{2m} \cong V \text{ given by } \Phi\,(\phi \otimes \psi) := \phi\,(\psi)\,.$$

Given what we have done, a direct proof of this is not difficult. Alternatively, it is known (see [47]) that, up to equivalence, the only irreducible representation of the algebra $\mathrm{End}\,(W)$ for any complex or real vector space W is the defining representation, namely $\mathrm{Id} : \mathrm{End}\,(W) \to \mathrm{End}\,(W)$. Since $\mathbb{C}l_{2m} \cong \mathrm{End}\,(\Sigma_{2m})$, it follows that, up to equivalence, $\rho_{\mathbb{C}} : \mathbb{C}l_{2m} \to \mathrm{End}\,(\Sigma_{2m})$ is the only irreducible representation of $\mathbb{C}l_{2m}$.

Definition 5.18 Let X be an oriented Riemannian n-manifold ($n = 2m$ even) with metric h, and oriented orthonormal frame bundle FX. Assume that X has a spin structure $P \to FX$, where P is a principal $\mathrm{Spin}\,(n)$-bundle and the projection $P \to FX$ is a two-fold cover, equivariant with respect to $\mathrm{Spin}\,(n) \to \mathrm{SO}\,(n)$. Furthermore, let $E \to X$ be a Hermitian vector bundle with unitary connection ε. The *twisted Dirac operator* $\mathcal{D}$ associated with the above data is

$$\mathcal{D} := (1 \otimes \mathbf{c}) \circ \nabla : C^\infty\,(E \otimes \Sigma\,(X)) \to C^\infty\,(E \otimes \Sigma\,(X))\,.$$

Here, $\Sigma\,(X)$ is the spin bundle over X associated to $P \to FX \to X$ via the spinor representation $\mathrm{Spin}\,(n) \to \mathrm{End}\,(\Sigma_n)$,

$$\mathbf{c} : C^\infty\,(\Sigma\,(X) \otimes TX^*) \to C^\infty\,(\Sigma\,(X))$$

is Clifford multiplication, and

$$\nabla : C^\infty\,(E \otimes \Sigma\,(X)) \to C^\infty\,(E \otimes \Sigma\,(X) \otimes TX^*)$$

is the covariant derivative determined by the connection ε and the spinorial lift to P of the Levi-Civita connection form, say θ, on FX.

Let $\Sigma^{\pm}\,(X)$ denote the ± 1 eigenbundles of the complex Clifford volume element in $C^\infty\,(\mathfrak{Cl}(X))$, given at a point $x \in X$ by $i^m e_1 \cdots e_n$, where $e_1, \ldots, e_n$ is an oriented, orthonormal basis of $T_x X$. The $\Sigma^{\pm}\,(X)$ are the so-called chiral halves of $\Sigma\,(X) = \Sigma^+\,(X) \oplus \Sigma^+\,(X)$. Since

$$\nabla\left(C^\infty\left(E \otimes \Sigma^{\pm}\,(X)\right)\right) \subseteq C^\infty\left(E \otimes \Sigma^{\pm}\,(X) \otimes TX^*\right) \text{ and}$$
$$(1 \otimes \mathbf{c})\left(C^\infty\left(E \otimes \Sigma^{\pm}\,(X) \otimes TX^*\right)\right) \subseteq C^\infty\left(E \otimes \Sigma^{\mp}\,(X)\right), \text{ we have}$$

$$\mathcal{D}=\mathcal{D}^{+}\oplus\mathcal{D}^{-}, \text{ where } \mathcal{D}^{\pm}:C^{\infty}\left(E\otimes\Sigma^{\pm}\left(X\right)\right)\rightarrow C^{\infty}\left(E\otimes\Sigma^{\mp}\left(X\right)\right).$$

The symbol of the first-order differential operator $\mathcal{D}$ is computed as follows. For $\phi\in C^{\infty}\left(X\right)$ with $\phi\left(x\right)=0$ and $\psi\in C^{\infty}\left(E\otimes\Sigma\left(X\right)\right)$, we have at x

$$\begin{aligned}\left(1\otimes\mathbf{c}\right)\circ\nabla\left(\phi\psi\right)&=\left(1\otimes\mathbf{c}\right)\circ\left(\left(d\phi\right)\psi+\phi\nabla\psi\right)=\left(1\otimes\mathbf{c}\right)\circ\left(d\phi\right)\psi\\&=\left(1\otimes\mathbf{c}\left(d\phi\right)\right)\psi.\end{aligned}$$

Thus, the symbol $\sigma\left(\mathcal{D}\right):T_{x}^{*}X\rightarrow\operatorname{End}\left(\Sigma\left(X\right)\right)$ at the covector $\xi_{x}\in T_{x}^{*}X$ is given by

$$\sigma\left(\mathcal{D}\right)\left(\xi_{x}\right)=1\otimes\mathbf{c}\left(\xi_{x}\right)\in\operatorname{End}\left(E_{x}\otimes\Sigma_{x}\right).$$

For $\xi_{x}\neq0$, $\sigma\left(\mathcal{D}\right)\left(\xi_{x}\right)$ is an isomorphism, since

$$\sigma\left(\mathcal{D}\right)\left(\xi_{x}\right)\circ\sigma\left(\mathcal{D}\right)\left(\xi_{x}\right)=1\otimes\mathbf{c}\left(\xi_{x}\right)^{2}=-\left|\xi_{x}\right|^{2}\mathrm{I}.$$

Thus, $\mathcal{D}$ is an elliptic operator. Moreover, since $\sigma\left(\mathcal{D}^{+}\right)$ and $\sigma\left(\mathcal{D}^{-}\right)$ are restrictions of $\sigma\left(\mathcal{D}\right)$, it follows that $\mathcal{D}^{+}$ and $\mathcal{D}^{-}$ are elliptic. It can be shown that $\mathcal{D}$ is formally self-adjoint, and $\mathcal{D}^{+}$ and $\mathcal{D}^{-}$ are formal adjoints of each other (see [34]). We also have a pair of self-adjoint elliptic operators

$$\begin{aligned}\mathcal{D}_{+}^{2}&:=\mathcal{D}^{2}|C^{\infty}\left(E\otimes\Sigma^{+}\left(X\right)\right)=\mathcal{D}^{-}\circ\mathcal{D}^{+}\\\mathcal{D}_{-}^{2}&:=\mathcal{D}^{2}|C^{\infty}\left(E\otimes\Sigma^{-}\left(X\right)\right)=\mathcal{D}^{+}\circ\mathcal{D}^{-}.\end{aligned}$$

For $\lambda\in\mathbb{C}$, let $V_{\lambda}\left(\mathcal{D}_{\pm}^{2}\right):=\left\{\psi\in C^{\infty}\left(E\otimes\Sigma^{\pm}\left(X\right)\right)\mid\mathcal{D}_{\pm}^{2}\psi=\lambda\psi\right\}$. From the general theory of formally self-adjoint, elliptic operators on compact manifolds, we know that $\operatorname{Spec}\left(\mathcal{D}_{\pm}^{2}\right)=\left\{\lambda\in\mathbb{C}\mid V_{\lambda}\left(\mathcal{D}_{\pm}^{2}\right)\neq\{0\}\right\}$ consists of the eigenvalues of $\mathcal{D}_{\pm}^{2}$ and is a discrete subset of $[0,\infty)$, the eigenspaces $V_{\lambda}\left(\mathcal{D}_{\pm}^{2}\right)$ are finite-dimensional, and an $L^{2}\left(E\otimes\Sigma^{\pm}\left(X\right)\right)$-complete orthonormal set of vectors can be selected from the $V_{\lambda}\left(\mathcal{D}_{\pm}^{2}\right)$. Note that $\mathcal{D}^{+}\left(V_{\lambda}\left(\mathcal{D}_{+}^{2}\right)\right)\subseteq V_{\lambda}\left(\mathcal{D}_{-}^{2}\right)$, since for $\psi\in V_{\lambda}\left(\mathcal{D}_{+}^{2}\right)$,

$$\begin{aligned}\mathcal{D}_{-}^{2}\left(\mathcal{D}^{+}\psi\right)&=\left(\mathcal{D}^{+}\circ\mathcal{D}^{-}\right)\left(\mathcal{D}^{+}\psi\right)=\mathcal{D}^{+}\left(\left(\mathcal{D}^{-}\circ\mathcal{D}^{+}\right)\left(\psi\right)\right)\\&=\mathcal{D}^{+}\left(\mathcal{D}_{+}^{2}\left(\psi\right)\right)=\mathcal{D}^{+}\left(\lambda\psi\right)=\lambda\mathcal{D}^{+}\left(\psi\right),\end{aligned}$$

and similarly $\mathcal{D}^{-}\left(V_{\lambda}\left(\mathcal{D}_{-}^{2}\right)\right)\subseteq V_{\lambda}\left(\mathcal{D}_{+}^{2}\right)$. For $\lambda\neq0$,

$$\mathcal{D}^{\pm}|V_{\lambda}\left(\mathcal{D}_{\pm}^{2}\right):V_{\lambda}\left(\mathcal{D}_{\pm}^{2}\right)\rightarrow V_{\lambda}\left(\mathcal{D}_{\mp}^{2}\right)$$

is an isomorphism, since it has inverse $\frac{1}{\lambda}\mathcal{D}^{\mp}$. Thus the set of nonzero eigenvalues (and their multiplicities) of $\mathcal{D}_{+}^{2}$ coincides with that of $\mathcal{D}_{-}^{2}$. However, in general

$$\dim V_{0}\left(\mathcal{D}_{+}^{2}\right)-\dim V_{0}\left(\mathcal{D}_{-}^{2}\right)=\dim\operatorname{Ker}\left(\mathcal{D}_{+}^{2}\right)-\dim\operatorname{Ker}\left(\mathcal{D}_{-}^{2}\right)=\operatorname{index}\left(\mathcal{D}^{+}\right)\neq0.$$

Since $\dim V_{\lambda}\left(\mathcal{D}_{+}^{2}\right)-\dim V_{\lambda}\left(\mathcal{D}_{-}^{2}\right)=0$ for $\lambda\neq0$, obviously

$$\begin{aligned}\operatorname{index}\left(\mathcal{D}^{+}\right)&=\dim V_{0}\left(\mathcal{D}_{+}^{2}\right)-\dim V_{0}\left(\mathcal{D}_{-}^{2}\right)\\&=\sum\nolimits_{\lambda\in\operatorname{Spec}\left(\mathcal{D}_{+}^{2}\right)}e^{-t\lambda}\left(\dim V_{\lambda}\left(\mathcal{D}_{+}^{2}\right)-\dim V_{\lambda}\left(\mathcal{D}_{-}^{2}\right)\right).\end{aligned}$$

This may seem like a very inefficient way to write $\operatorname{index}(\mathcal{D}^+)$, but the point is that the sum can be expressed as the integral of the supertrace of the heat kernel for the spinorial heat equation $\frac{\partial \psi}{\partial t} = -\mathcal{D}^2 \psi$, from which the Local Index Theorem (Theorem 5.21 below) for $\mathcal{D}^+$ will eventually follow. However, first the existence of the heat kernel needs to be established.

Let the *positive* eigenvalues of $\mathcal{D}^2_\pm$ be placed in a sequence $0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \ldots$ where each eigenvalue is repeated according to its multiplicity. Let $u_1^\pm, u_2^\pm, \ldots$ be an L^2-orthonormal sequence in $C^\infty\left(E \otimes \Sigma^+(X)\right)$ with $\mathcal{D}^2_\pm\left(u_j^\pm\right) = \lambda_j u_j^\pm$ (i.e., $u_j^\pm \in V_{\lambda_j}\left(\mathcal{D}^2_\pm\right)$). We let $u^+_{0_1}, \ldots, u^+_{0_{n^+}}$ be an L^2-orthonormal basis of $\operatorname{Ker} \mathcal{D}^2_+ = \operatorname{Ker} \mathcal{D}_+$, and $u^-_{0_1}, \ldots, u^-_{0_{n^-}}$ be an L^2-orthonormal basis of $\operatorname{Ker} \mathcal{D}^2_- = \operatorname{Ker} \mathcal{D}_-$. We can pull back the bundle $E \otimes \Sigma^\pm(X)$ via either of the projections $X \times X \times (0,\infty) \to X$ given by $\pi_1(x,y,t) := x$ and $\pi_2(x,y,t) := y$ and take the tensor product of the results to form a bundle

$$\mathcal{K}^\pm := \pi_1^*\left(E \otimes \Sigma^\pm(X)\right) \otimes \pi_2^*\left(E \otimes \Sigma^\pm(X)\right) \to X \times X \times (0,\infty).$$

Note that for $x \in X$, the Hermitian inner product $\langle\ ,\ \rangle_x$ on $\left(E \otimes \Sigma^\pm(X)\right)_x$ gives us a conjugate-linear map $\psi \mapsto \psi^*(\cdot) := \langle \cdot, \psi\rangle_x$ from $\left(E \otimes \Sigma^\pm(X)\right)_x$ to its dual $\left(E \otimes \Sigma^\pm(X)\right)^*_x$. Thus, we can (and do) make the identifications

$$\begin{aligned}
&\pi_1^*\left(E \otimes \Sigma^\pm(X)\right) \otimes \pi_2^*\left(E \otimes \Sigma^\pm(X)\right) \\
&\cong \left(\pi_1^*\left(E \otimes \Sigma^\pm(X)\right)\right)^* \otimes \pi_2^*\left(E \otimes \Sigma^\pm(X)\right) \\
&\cong \operatorname{Hom}\left(\pi_1^*\left(E \otimes \Sigma^\pm(X)\right), \pi_2^*\left(E \otimes \Sigma^\pm(X)\right)\right).
\end{aligned}$$

The full proof of the following Proposition is contained in [27] for readers of sufficient background.

Proposition 5.19 *For $t > t_0 > 0$, the series $k'^\pm$, defined by*

$$k'^\pm(x,y,t) := \sum_{j=1}^\infty e^{-\lambda_j t} u_j^\pm(x) \otimes u_j^\pm(y),$$

converges uniformly in $C^q(\mathcal{K}^\pm | X \times X \times (t_0,\infty))$ for all $q \geq 0$. Hence $k'^\pm \in C^\infty(\mathcal{K}^\pm)$, and (for $t > 0$)

$$\frac{\partial}{\partial t} k'^\pm(x,y,t) = -\sum_{j=1}^\infty \lambda_j e^{-\lambda_j t} u_j^\pm(x) \otimes u_j^\pm(y) = -\mathcal{D}^2_\pm k'^\pm(x,y,t).$$

Definition 5.20 The *positive and negative twisted spinorial heat kernels* (or the *heat kernels for $\mathcal{D}^2_\pm$*) $k^\pm \in C^\infty(\mathcal{K}^\pm)$ are given by

$$\begin{aligned}
k^\pm(x,y,t) &:= \sum\nolimits_{i=1}^{n^\pm} u^\pm_{0_i}(x) \otimes u^\pm_{0_i}(y) + k'^\pm(x,y,t) \\
&= \sum\nolimits_{i=1}^{n^\pm} u^\pm_{0_i}(x) \otimes u^\pm_{0_i}(y) + \sum\nolimits_{j=1}^{\infty} e^{-\lambda_j t} u_j^\pm(x) \otimes u_j^\pm(y) \ \text{ for } t > 0.
\end{aligned}$$

The *total twisted spinorial heat kernel* (or the *heat kernel for $\mathcal{D}^2$*) is

$$k = \left(k^+, k^-\right) \in C^\infty\left(\mathcal{K}^+\right) \oplus C^\infty\left(\mathcal{K}^-\right) \cong C^\infty\left(\mathcal{K}^+ \oplus \mathcal{K}^-\right) \subseteq C^\infty(\mathcal{K}),$$
$$\text{where } \mathcal{K} := \mathcal{K}^+ \oplus \mathcal{K}^- = \operatorname{Hom}\left(\pi_1^*(E \otimes \Sigma(X)), \pi_2^*(E \otimes \Sigma(X))\right).$$

For any finite dimensional Hermitian vector space $(V, \langle\cdot,\cdot\rangle)$ with orthonormal basis $e_1, \ldots, e_N$, we have (for $v \in V$)

$$\begin{aligned}\operatorname{Tr}(v^* \otimes v) &= \sum\nolimits_{i=1}^{N} \langle (v^* \otimes v)(e_i), e_i\rangle = \sum\nolimits_{i=1}^{N} \langle v^*(e_i)\, v, e_i\rangle \\ &= \sum\nolimits_{i=1}^{N} \langle \langle e_i, v\rangle v, e_i\rangle = \sum\nolimits_{i=1}^{N} \langle e_i, v\rangle \langle v, e_i\rangle = \sum\nolimits_{i=1}^{N} |\langle e_i, v\rangle|^2 = |v|^2 .\end{aligned}$$

In particular, $k^{\pm}(x,x,t) \in \operatorname{End}\left((E \otimes \Sigma^{\pm}(X))_x\right)$ and

$$\operatorname{Tr}\left(k^{\pm}(x,x,t)\right) = \sum\nolimits_{i=1}^{n^{\pm}} \left|u_{0_i}^{\pm}(x)\right|^2 + \sum\nolimits_{j=1}^{\infty} e^{-\lambda_j t} \left|u_j^{\pm}(x)\right|^2 .$$

Since this series converges uniformly and $\left\|u_{0_i}^{\pm}\right\|_{2,0} = \left\|u_j^{\pm}\right\|_{2,0} = 1$, we have

$$\int_X \operatorname{Tr}\left(k^{\pm}(x,x,t)\right)\, \nu_x = n^{\pm} + \sum\nolimits_{j=1}^{\infty} e^{-\lambda_j t} < \infty.$$

For $t > 0$, we define the bounded operator $e^{-t\mathcal{D}_{\pm}^2} \in \operatorname{End}\left(L^2\left(E \otimes \Sigma^{\pm}(X)\right)\right)$ by

$$e^{-t\mathcal{D}_{\pm}^2}\left(\psi^{\pm}\right) = \sum\nolimits_{i=1}^{n^{\pm}} \left(u_{0_i}^{\pm}, \psi_0^{\pm}\right) u_{0_i}^{\pm} + \sum\nolimits_{j=1}^{\infty} e^{-\lambda_j t} \left(u_j^{\pm}, \psi_0^{\pm}\right) u_j^{\pm}.$$

Note that $e^{-t\mathcal{D}_{\pm}^2}$ is of trace class, since

$$\operatorname{Tr}\left(e^{-t\mathcal{D}_{\pm}^2}\right) = n^{\pm} + \sum\nolimits_{j=1}^{\infty} e^{-\lambda_j t} = \int_X \operatorname{Tr}\left(k^{\pm}(x,x,t)\right)\, \nu_x < \infty.$$

Now, we have

$$\begin{aligned}\operatorname{index}\left(\mathcal{D}^+\right) &= \dim V_0\left(\mathcal{D}_+^2\right) - \dim V_0\left(\mathcal{D}_-^2\right) \\ &= n^+ - n^- + \sum\nolimits_{j=1}^{\infty} \left(e^{-\lambda_j t} - e^{-\lambda_j t}\right) \\ &= n^+ + \sum\nolimits_{j=1}^{\infty} e^{-\lambda_j t} - \left(n^- + \sum\nolimits_{j=1}^{\infty} e^{-\lambda_j t}\right) \\ &= \int_X \left(\operatorname{Tr}\left(k^+(x,x,t)\right) - \operatorname{Tr}\left(k^-(x,x,t)\right)\right)\, \nu_x. \end{aligned} \tag{14}$$

Since $\mathcal{D}^2 = \mathcal{D}_+^2 \oplus \mathcal{D}_-^2$, we also have the operator $e^{-t\mathcal{D}^2} \in \operatorname{End}\left(L^2\left(E \otimes \Sigma(X)\right)\right)$, whose trace is given by

$$\operatorname{Tr}\left(e^{-t\mathcal{D}^2}\right) = \int_X \operatorname{Tr}\left(k(x,x,t)\right)\, \nu_x = \int_X \left(\operatorname{Tr}\left(k^+(x,x,t)\right) + \operatorname{Tr}\left(k^-(x,x,t)\right)\right)\, \nu_x.$$

The *supertrace* of $k(x,x,t)$ is defined by

$$\operatorname{Str}\left(k(x,x,t)\right) := \operatorname{Tr}\left(k^+(x,x,t)\right) - \operatorname{Tr}\left(k^-(x,x,t)\right),$$

and in view of (14), we have

$$\operatorname{index}\left(\mathcal{D}^+\right) = \int_X \operatorname{Str}\left(k(x,x,t)\right)\, \nu_x. \tag{15}$$

The left side is independent of t and so the right side is also independent of t. The main task now is to determine the behavior of $\mathrm{Str}\,(k\,(x,x,t))$ as $t \to 0^+$. We suspect that for each $x \in X$, as $t \to 0^+$, $k\,(x,x,t)$ and $\mathrm{Str}\,(k\,(x,x,t))$ are influenced primarily by the geometry (e.g., curvature form Ω^θ of X with metric h and Levi-Civita connection θ, and the curvature Ω^ε of the unitary connection for E) near x, since the heat sources of points far from x are not felt very strongly at x for small t. Indeed, we have the following Local Index Formula and we will give an outline a proof after explaining the notation.

Theorem 5.21 (The Local Index Theorem) *In the notation of Definitions 5.18 and 5.20, let $\mathcal{D} : C^\infty\,(E \otimes \Sigma\,(X)) \to C^\infty\,(E \otimes \Sigma\,(X))$ be a twisted Dirac operator and let $k \in C^\infty\,(\mathcal{K})$ be the heat kernel for $\mathcal{D}^2$. If Ω^ε is the curvature form of the unitary connection ε for E and Ω^θ is the curvature form of the Levi-Civita connection θ for (X,h) with volume element ν, then*

$$\lim_{t\to 0^+} \mathrm{Str}\,(k\,(x,x,t)) = \left\langle \mathrm{Tr}\left(e^{i\Omega^\varepsilon/2\pi}\right) \wedge \det\left(\frac{i\Omega^\theta/4\pi}{\sinh\,(i\Omega^\theta/4\pi)}\right)^{\frac{1}{2}}, \nu_x \right\rangle. \tag{16}$$

Remark 5.22 As will be explained below, the right side is really the inner product, with volume form ν at x, of the canonical form $\mathbf{ch}\,(E,\varepsilon) \smile \widehat{\mathbf{A}}\,(X,\theta)$ (depending on the connections ε for E and the Levi-Civita connection θ for the metric h) which represents $\mathbf{ch}\,(E) \smile \widehat{\mathbf{A}}\,(X)$. As a consequence, we obtain the Index Theorem for twisted Dirac operators from the Local Index Formula in Corollary 5.23 below. Thus, the Local Index Formula is stronger than the Index Theorem for twisted Dirac operators. Indeed, the Local Index Formula yields the Index Theorem for elliptic operators which are locally expressible as twisted Dirac operators or direct sums of such.

Corollary 5.23 (Index formula for twisted Dirac operators) *For an oriented Riemannian n-manifold X (n even) with spin structure, and a Hermitian vector bundle $E \to X$ with unitary connection, let $\mathcal{D} = \mathcal{D}^+ \oplus \mathcal{D}^-$ be the twisted Dirac operator, with $\mathcal{D}^+ : C^\infty\,(E \otimes \Sigma^+\,(X)) \to C^\infty\,(E \otimes \Sigma^-\,(X))$. We have*

$$\mathrm{index}\left(\mathcal{D}^+\right) = \left(\mathbf{ch}\,(E) \smile \widehat{\mathbf{A}}\,(X)\right)[X],$$

where $\mathbf{ch}\,(E)$ is the total Chern character class of E and $\widehat{\mathbf{A}}\,(X)$ is the total $\widehat{\mathbf{A}}$ class of X, both defined below. In particular, we obtain:

$$n = 2 \Rightarrow \ \mathrm{index}\,(\mathcal{D}^+) = ch_1\,(E)\,[X] = c_1\,(E)\,[X] \ \textit{and}$$

$$n = 4 \Rightarrow \begin{cases} \mathrm{index}\,(\mathcal{D}^+) = \left(\mathbf{ch}\,(E) \smile \widehat{\mathbf{A}}\,(X)\right)[X] \\ = \left(-\dim E \cdot \frac{1}{24} p_1\,(TX) + ch_2\,(E)\right)[X] \\ = \left(-\frac{\dim E}{24} p_1\,(TX) + \frac{1}{2} c_1\,(E)^2 - c_2\,(E)\right)[X]. \end{cases}$$

Proof By (15), (16) and Remark 5.22, we have

$$\begin{aligned} &\mathrm{index}\left(\mathcal{D}^+\right) = \int_X \mathrm{Str}\,(k\,(x,x,t))\,\nu_x = \int_X \lim_{t\to 0^+} \mathrm{Str}\,(k\,(x,x,t))\,\nu_x \\ &= \int_X \langle \mathbf{ch}\,(E,\varepsilon)_x \smile \widehat{\mathbf{A}}\,(X,\theta)_x\,, \nu_x \rangle \nu_x = (\mathbf{ch}\,(E) \smile \widehat{\mathbf{A}}\,(X))\,[X]. \end{aligned}$$

□

We now explain the meaning of the form

$$\mathrm{Tr}\left(e^{i\Omega^{\varepsilon}/2\pi}\right)\wedge\det\left(\frac{i\Omega^{\theta}/4\pi}{\sinh\left(i\Omega^{\theta}/4\pi\right)}\right)^{\frac{1}{2}}.$$

The first part $\mathrm{Tr}\left(e^{i\Omega^{\varepsilon}/2\pi}\right)$ is relatively easy. We have (recall $2m=\dim X$)

$$e^{i\Omega^{\varepsilon}/2\pi}:=\sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{i}{2\pi}\right)^{k}\Omega^{\varepsilon}\wedge\overset{k}{\cdots}\wedge\Omega^{\varepsilon}=\sum_{k=0}^{m}\frac{1}{k!}\left(\frac{i}{2\pi}\right)^{k}\Omega^{\varepsilon}\wedge\overset{k}{\cdots}\wedge\Omega^{\varepsilon},\qquad(17)$$

where $\Omega^{\varepsilon}\wedge\overset{k}{\cdots}\wedge\Omega^{\varepsilon}\in\Omega^{2k}\left(\mathrm{End}\left(E\right)\right)$. Also $\mathrm{Tr}\left(i^{k}\Omega^{\varepsilon}\wedge\overset{k}{\cdots}\wedge\Omega^{\varepsilon}\right)\in\Omega^{2k}\left(X\right)$ and

$$\mathrm{Tr}\left(e^{i\Omega^{\varepsilon}/2\pi}\right)\in\bigoplus\nolimits_{k=1}^{m}\Omega^{2k}\left(X\right).$$

This (by one of many equivalent definitions) is a representative of the total Chern character $\mathbf{ch}\left(E\right)\in\bigoplus_{k=1}^{m}H^{2k}\left(X,\mathbb{Q}\right)$. The curvature Ω^{θ} of the Levi-Civita connection θ for the metric h has values in the skew-symmetric endomorphisms of TX; i.e., $\Omega^{\theta}\in\Omega^{2}\left(\mathrm{End}\left(TX\right)\right)$. A skew-symmetric endomorphism of $\mathbb{R}^{2m}$, say $B\in\mathfrak{so}\left(n\right)$, has pure imaginary eigenvalues $\pm ir_k$, where $r_k\in\mathbb{R}$ $(1\le k\le m)$. Thus, iB has real eigenvalues $\pm r_k$. Now $\frac{z/2}{\sinh(z/2)}$ is a power series in z with radius of convergence 2π. Thus, $\frac{isB/2}{\sinh(isB/2)}$ is defined for s sufficiently small and has eigenvalues $\frac{r_k s/2}{\sinh(r_k s/2)}$ each repeated twice. Hence

$$\det\left(\frac{isB/2}{\sinh\left(isB/2\right)}\right)=\prod\nolimits_{k=1}^{m}\left(\frac{r_k s/2}{\sinh\left(r_k s/2\right)}\right)^{2}\text{ and}$$
$$\det\left(\frac{isB/2}{\sinh\left(isB/2\right)}\right)^{\frac{1}{2}}=\prod\nolimits_{k=1}^{m}\frac{r_k s/2}{\sinh\left(r_k s/2\right)}.$$

The last product is a power series in s of the form

$$\prod_{k=1}^{m}\frac{r_k s/2}{\sinh\left(r_k s/2\right)}=\sum_{k=0}^{\infty}a_k\left(r_1^2,\ldots,r_m^2\right)s^{2k},$$

where the coefficient $a_k\left(r_1^2,\ldots,r_m^2\right)$ is a homogeneous, symmetric polynomial in $r_1^2,\ldots,r_m^2$ of degree k. One can always express any such a symmetric polynomial as a polynomial in the elementary symmetric polynomials $\sigma_1,\ldots,\sigma_m$ in $r_1^2,\ldots,r_m^2$, where

$$\sigma_1=\sum\nolimits_{i=1}^{m}r_i^2,\ \sigma_2=\sum\nolimits_{i<j}^{m}r_i^2r_j^2,\ \sigma_2=\sum\nolimits_{i<j<k}^{m}r_i^2r_j^2r_k^2,\ldots.$$

These in turn may be expressed in terms of $\mathrm{SO}\left(n\right)$-invariant polynomials in the entries of $B\in\mathfrak{so}\left(n\right)$ via

$$\det\left(\lambda I-B\right)=\prod_{j=1}^{m}\left(\lambda+ir_j\right)\left(\lambda-ir_j\right)=\prod_{j=1}^{m}\left(\lambda^2+r_j^2\right)=\sum_{k=1}^{m}\sigma_k\left(r_1^2,\ldots,r_m^2\right)\lambda^{2(m-k)}.$$

On the other hand,

$$\det(\lambda I - B) = \sum_{k=1}^{m} \left(\frac{1}{(2k)!} \sum_{(i),(j)} \delta^{j_1 \cdots j_{2k}}_{i_1 \cdots i_{2k}} B^{i_1}_{\ j_1} \cdots B^{i_{2k}}_{\ j_{2k}} \right) \lambda^{2(m-k)}, \text{ and so}$$

$$\sigma_k\left(r_1^2, \ldots, r_m^2\right) = \frac{1}{(2k)!} \sum_{(i),(j)} \delta^{j_1 \cdots j_{2k}}_{i_1 \cdots i_{2k}} B^{i_1}_{\ j_1} \cdots B^{i_{2k}}_{\ j_{2k}},$$

where $(i) = (i_1, \cdots, i_{2k})$ is an ordered $2k$-tuple of distinct elements of $\{1, \ldots, 2m\}$ and (j) is a permutation of (i) with sign $\delta^{j_1 \cdots j_{2k}}_{i_1 \cdots i_{2k}}$. If we replace $B^i_{\ j}$ with the 2-form $\frac{1}{2\pi}\left(\Omega^\theta\right)^i_{\ j}$ relative to an orthonormal frame field, we obtain the Pontryagin forms

$$p_k\left(\Omega^\theta\right) := \frac{1}{(2\pi)^{2k}(2k)!} \sum_{(i),(j)} \delta^{j_1 \cdots j_{2k}}_{i_1 \cdots i_{2k}} \Omega^\theta_{i_1 j_1} \wedge \cdots \wedge \Omega^\theta_{i_{2k} j_{2k}},$$

which represent the Pontryagin classes of the $\mathrm{SO}(n)$ bundle FX. Note that $p_k\left(\Omega^\theta\right)$ is independent of the choice of framing by the ad-invariance of the polynomials σ_k. If we express the $a_k\left(r_1^2, \ldots, r_m^2\right)$ as polynomials, say $\widehat{A}_k(\sigma_1, \ldots, \sigma_k)$, in the σ_j ($j \leq k$), we can ultimately write

$$\det\left(\frac{isB/2}{\sinh(isB/2)}\right)^{\frac{1}{2}} = \sum_{k=0}^{\infty} \widehat{A}_k(\sigma_1, \ldots, \sigma_k)\, s^{2k}.$$

Formally replacing B by $\frac{1}{2\pi}\Omega^\theta$, we finally have the reasonable definition

$$\det\left(\frac{i\Omega^\theta/4\pi}{\sinh\left(i\Omega^\theta/4\pi\right)}\right)^{\frac{1}{2}} := \sum_{k=0}^{\infty} \widehat{A}_k\left(p_1\left(\Omega^\theta\right), \ldots, p_k\left(\Omega^\theta\right)\right),$$

where the $p_j\left(\Omega^\theta\right)$ are multiplied via wedge product when evaluating the terms in the sum; the order of multiplication does not matter since $p_j\left(\Omega^\theta\right)$ is of even degree $4j$. Also, since $\widehat{A}_k\left(p_1\left(\Omega^\theta\right), \ldots, p_k\left(\Omega^\theta\right)\right)$ is a $4k$-form, there are only a finite number of nonzero terms in the infinite sum. Abbreviating $p_j(\Omega^\theta)$ simply by p_j, one finds

$$\begin{aligned} \det\left(\frac{i\Omega^\theta/4\pi}{\sinh\left(i\Omega^\theta/4\pi\right)}\right)^{\frac{1}{2}} &= 1 - \frac{1}{24}p_1 + \frac{1}{5760}\left(7p_1^2 - 4p_2\right) \\ &\quad - \frac{1}{967\,680}\left(31p_1^3 - 44p_1p_2 + 16p_3\right) + \cdots . \end{aligned} \tag{18}$$

This (by one definition) represents the total $\widehat{A}$-class of X, denoted by

$$\widehat{\mathbf{A}}(X) \in \bigoplus_{k=1}^{m} H^{2k}(X, \mathbb{Q}),$$

where actually $\widehat{\mathbf{A}}(X)$ has only nonzero components in $H^{2k}(X, \mathbb{Q})$ when k is even (or $2k \equiv 0 \bmod 4$). In (16) the multi-degree forms (17) and (18) have been wedged, and the top ($2m$-degree) component (relative to the volume form) has been harvested.

We now turn to our outline of the proof of Theorem 5.21. The well-known heat kernel (or fundamental solution) for the ordinary heat equation $u_t = \Delta u$ in Euclidean space $\mathbb{R}^n$, is given by

$$e(x,y,t) = (4\pi t)^{-n/2} \exp\left(-|x-y|^2/4t\right).$$

Since $H(x,y,t)$ only depends on $r = |x-y|$ and t, it is convenient to write

$$e(x,y,t) = \mathcal{E}(r,t) := (4\pi t)^{-n/2} \exp\left(-r^2/4t\right).$$

We do not expect such a simple expression for the heat kernel $k = (k^+, k^-)$ of Definition 5.20. However, it can be shown that for $x, y \in X$ (of even dimension $n = 2m$) with $r = d(x,y) :=$ Riemannian distance from x to y sufficiently small, we have an asymptotic expansion as $t \to 0^+$ for $k(x,y,t)$ of the form

$$k(x,y,t) \sim H_Q(x,y,t) := \mathcal{E}(d(x,y),t) \sum\nolimits_{j=0}^{Q} h_j(x,y)\, t^j, \tag{19}$$

for any fixed integer $Q > m+4$, where

$$h_j(x,y) \in \operatorname{Hom}\left((E \otimes \Sigma(X))_x, (E \otimes \Sigma(X))_y\right), \; j \in \{0,1,\ldots,Q\}.$$

The meaning of $k(x,y,t) \sim H_Q(x,y,t)$ is that for $d(x,y)$ and t sufficiently small,

$$|k(x,y,t) - H_Q(x,y,t)| \le C_Q \mathcal{E}(d(x,y),t)\, t^{Q+1} \le C_Q t^{Q-m+1},$$

where C_Q is a constant, independent of (x,y,t). We then have

$$k(x,x,t) \sim (4\pi t)^{-m} \sum_{j=0}^{Q} h_j(x,x)\, t^j = (4\pi)^{-m} \sum_{j=0}^{Q} h_j(x,x)\, t^{j-m}. \tag{20}$$

Using (15), i.e., $\int_X \operatorname{Str}(k(x,x,t))\, \nu_x = \operatorname{index}(\mathcal{D}^+)$ and (20), we deduce that

$$\int_X \operatorname{Str}(h_j(x,x))\, \nu_x = 0 \;\text{ for }\; j \in \{0,1,\ldots,m-1\}, \text{ while}$$
$$(4\pi)^{-m} \int_X \operatorname{Str}(h_m(x,x))\, \nu_x = \int_X \operatorname{Str}(k(x,x,t))\, \nu_x = \operatorname{index}\left(\mathcal{D}^+\right).$$

Thus, to prove the Local Index Formula, it suffices to show that

$$(4\pi)^{-m} \operatorname{Str}(h_m(x,x)) = \left\langle \operatorname{Tr}\left(e^{i\Omega^{\varepsilon}/2\pi}\right) \wedge \det\left(\frac{i\Omega^{\theta}/4\pi}{\sinh(i\Omega^{\theta}/4\pi)}\right)^{\frac{1}{2}}, \nu_x \right\rangle.$$

While this may not be the intellectual equivalent of climbing Mount Everest, it is not for the faint of heart.

We choose a normal coordinate system $\left(y^1,\ldots,y^n\right)$ in a coordinate ball $\mathcal{B}$ centered at the fixed point $x \in X$, so that $\left(y^1,\ldots,y^n\right) = 0$ at x. The coordinate fields $\partial_1 := \partial/\partial y^1, \ldots, \partial_n := \partial/\partial y^n$ are orthonormal at x, and for any fixed $y_0 \in \mathcal{B}$ with coordinates

$\left(y_0^1,\ldots,y_0^n\right)$, the curve $t\mapsto t\left(y_0^1,\ldots,y_0^n\right)$ is a geodesic through x. By parallel translating the frame $(\partial_1,\ldots,\partial_n)$ at x along these radial geodesics, we obtain an orthonormal frame field $(E_1,\ldots,E_n)$ on $\mathcal{B}$ which generally does not coincide with $(\partial_1,\ldots,\partial_n)$ at points $y\in\mathcal{B}$ other than at x. The framing $(E_1,\ldots,E_n)$ defines a particularly nice section $\mathcal{B}\to FX|B$ and we may lift this to a section $\mathcal{B}\to P|B$ of the spin structure, which enables us to view the space $C^\infty\left(\Sigma\left(X\right)|\mathcal{B}\right)$ of spinor fields on $\mathcal{B}$ as $C^\infty\left(\mathcal{B},\Sigma_n\right)$, i.e., functions on $\mathcal{B}$ with values in the fixed spinor representation vector space $\Sigma_n=\Sigma_n^+\oplus\Sigma_n^-$. By similar radial parallel translation (with respect to the connection ε) of an orthonormal basis of the twisting bundle fiber E_x, we can identify $C^\infty\left(E|\mathcal{B}\right)$ with $C^\infty\left(\mathcal{B},\mathbb{C}^N\right)$, where $N=\dim_{\mathbb{C}}E$. The coordinate expressions for the curvatures Ω^θ, Ω^ε and $\mathcal{D}^2$ are as simple as possible in this so-called radial gauge.

With the above identifications, we proceed as follows. For $0\le Q\in\mathbb{Z}$, let $\Psi_Q\in C^\infty\left(\mathcal{B}\times(0,\infty),\mathbb{C}^N\otimes\Sigma_{2m}\right)$ be of the form

$$\Psi_Q\left(y,t\right):=\mathcal{E}\left(r,t\right)\sum\nolimits_{k=0}^{Q}U_k\left(y\right)t^k,$$

where $U_k\in C^\infty\left(\mathcal{B},\mathbb{C}^N\otimes\Sigma_{2m}\right)$. If $U_0\left(0\right)\in\mathbb{C}^N\otimes\Sigma_{2m}$ is arbitrarily specified, we seek a formula for $U_k\left(y\right)$, $k=0,\ldots,Q$, such that

$$\left(\mathcal{D}^2+\partial_t\right)\Psi_Q\left(y,t\right)=\mathcal{E}\left(r,t\right)t^Q\mathcal{D}^2\left(U_Q\right)\left(y\right), \tag{21}$$

where the square $\mathcal{D}^2$ of the Dirac operator $\mathcal{D}$ can be written (where "$\cdot$" is Clifford multiplication) as

$$\mathcal{D}^2\psi=-\Delta\psi+\tfrac{1}{2}\sum\nolimits_{j,k}\Omega^\varepsilon_{jk}E_j\cdot E_k\cdot\psi+\tfrac{1}{4}S\psi,$$

by virtue of the generalized Lichnerowicz formula (see [34, p. 164]). It is convenient to define the 0-th order operator $\mathcal{F}$ on $C^\infty\left(\mathcal{B},\mathbb{C}^N\otimes\Sigma_{2m}\right)$ via

$$\mathcal{F}\left[\psi\right]:=\tfrac{1}{2}\sum\nolimits_{j,k}\Omega^\varepsilon_{jk}E_j\cdot E_k\cdot\psi,\ \text{so that}\ \mathcal{D}^2=-\Delta\psi+\left(\mathcal{F}+\tfrac{1}{4}S\right)\left[\psi\right].$$

The desired formula for the $U_k\left(y\right)$ involves the operator A on $C^\infty\left(\mathcal{B},\mathbb{C}^N\otimes\Sigma_{2m}\right)$ given by

$$A\left[\psi\right]:=-h^{1/4}\mathcal{D}^2[h^{-1/4}\psi]=h^{1/4}\Delta[h^{-1/4}\psi]-\left(\mathcal{F}+\tfrac{1}{4}S\right)\left[\psi\right],$$

where $h^{1/4}:=(\sqrt{\det h})^{1/2}$. For $s\in[0,1]$, let

$$A_s\left[\psi\right]\left(y\right):=A\left[\psi\right]\left(sy\right).$$

As is proved in [16], we have

Proposition 5.24 *Let $U_0\left(0\right)\in\mathbb{C}^N\otimes\Sigma_{2m}$, and let $V_0\in C^\infty\left(\mathcal{B},\mathbb{C}^N\otimes\Sigma_{2m}\right)$ be the constant function $V_0\left(y\right)\equiv U_0\left(0\right)$. Then the $U_k\left(y\right)$ which satisfy (21) are given by*

$$U_k\left(y\right)=h\left(y\right)^{-1/4}V_k\left(y\right),\ \textit{where}$$

$$V_k\left(y\right)=\int_{I^k}\prod\nolimits_{i=0}^{k-1}\left(s_i\right)^i\left(A_{s_{k-1}}\circ\cdots\circ A_{s_0}\left[V_0\right]\right)\left(y\right)ds_0\ldots ds_{k-1}, \tag{22}$$

and where $I^k=\left\{(s_0,\ldots s_k):s_i\in\left[0,1\right],i\in\left\{0,\ldots,k-1\right\}\ \right\}$.

Note that $U_0(0) \in \mathbb{C}^N \otimes \Sigma_{2m}$ may be arbitrarily specified, and once $U_0(0)$ is chosen, the $U_m(y)$ are uniquely determined via (22). Let $h_k(y) \in \mathrm{End}\left(\mathbb{C}^N \otimes \Sigma_{2m}\right)$ be given by

$$h_k(y)(U_0(0)) := U_k(y) \tag{23}$$

(in particular, $h_0(0) = \mathrm{I} \in \mathrm{End}\left(\mathbb{C}^N \otimes \Sigma_{2m}\right)$), and

$$H_Q(0,y,t) := \mathcal{E}(r,t)\sum_{k=0}^{Q} h_k(y)\, t^k \in C^\infty\left(\mathcal{B}, \mathrm{End}\left(\mathbb{C}^N \otimes \Sigma_{2m}\right)\right).$$

We may regard $H_Q(0,y,t)$ as

$$H_Q(x,y,t) \in \mathrm{Hom}\left(E_x \otimes \Sigma(X)_x, E_y \otimes \Sigma(X)_y\right),$$

where we recall that $x \in X$ is the point about which we have chosen normal coordinates. For y sufficiently close to x, we set

$$H_Q(x,y,t) := \mathcal{E}(d(x,y),t)\sum_{k=0}^{Q} h_k(x,y)\, t^k.$$

Of course, one expects that $H_Q(x,y,t)$ provides the desired asymptotic expansion (19). We will only state here without proof that

$$k(x,y,t) \sim H_Q(x,y,t) := \mathcal{E}(d(x,y),t)\sum\nolimits_{j=0}^{Q} h_j(x,y)\, t^j$$

for $d(x,y)$ sufficiently small, where the h_j are given in (23).

Using normal coordinates $\left(y^1,\ldots,y^{2m}\right) \in B(r_0,0)$ about $x \in X$ and the radial gauge, and selecting $V_0 \in \mathbb{C}^N \otimes \Sigma_{2m}$, by Proposition 5.24, we have

$$h_m(x,x)(V_0) = \int_{I^m} \prod\nolimits_{i=0}^{m-1} (s_i)^i \left(\left(A_{s_{m-1}} \circ \cdots \circ A_{s_0}\right)\left[\widetilde{V_0}\right]\right)(0)\, ds_0 \ldots ds_{m-1}, \tag{24}$$

where $\widetilde{V_0} \in C^\infty\left(B(r_0,0), \mathbb{C}^N \otimes \Sigma_{2m}\right)$ is the constant extension of V_0. Recall that for $\psi \in C^\infty\left(B(r_0,0), \mathbb{C}^N \otimes \Sigma_{2m}\right)$, we have

$$A_s[\psi](y) := A[\psi](sy), \text{ where } A[\psi] := h^{1/4}\Delta[h^{-1/4}\psi] - \left(\mathcal{F} + \tfrac{1}{4}S\right)[\psi].$$

While the right side of (24) may seem unwieldy, there is substantial simplification due to facts that $\left(A_{s_{m-1}} \circ \cdots \circ A_{s_0}\right)[\widetilde{V_0}](y)$ is evaluated at $y = 0$ in (24). Also, if $\gamma^1,\ldots,\gamma^n$ denote the so-called gamma matrices for Clifford multiplication by $\partial_1,\ldots,\partial_n$, only those terms of $A_{s_{m-1}} \circ \cdots \circ A_{s_0}[\widetilde{V_0}](0)$ which involve the product $\gamma_{n+1} := \gamma^1 \cdots \gamma^n$ will survive when the supertrace $\mathrm{Str}(h_m(x,x))$ is taken. As a consequence, we have the following simplification contained in [16].

Proposition 5.25 *Let $R_{klji}(0) = h\left(\Omega^\theta(\partial_i,\partial_j)\partial_l, \partial_k\right)$ denote the components of the Riemann curvature tensor of h at x, and let $\Omega^\varepsilon_{ij}(0) := \Omega^\varepsilon(\partial_i,\partial_j)$ at x. Set*

$$\widetilde{\theta}^1(\partial_j) := \tfrac{1}{8}\sum\nolimits_{k,l,i} R_{klji}(0)\, \gamma^k\gamma^l y^i,$$

$$\mathcal{F}^0 := \frac{1}{2}\sum\nolimits_{i,j} F_{ij} \otimes \gamma^i\gamma^j = \frac{1}{2}\sum\nolimits_{i,j} \Omega^{\varepsilon}_{ij}(0) \otimes \gamma^i\gamma^j, \text{ and}$$
$$A^0 := \sum\nolimits_i \left(\partial_i^2 + \widetilde{\theta}^1 (\partial_i)^2\right) - \mathcal{F}_0.$$

For $V_0 \in \mathbb{C}^N \otimes \Sigma_{2m}$*, define*

$$h^0_m(0,0)(V_0) := \int_{I^k} \prod\nolimits_{i=0}^{m-1} (s_i)^i \left(\left(A^0_{s_{m-1}} \circ \cdots \circ A^0_{s_0}\right)\left[\widetilde{V}_0\right]\right)(0)\, ds_0 \ldots ds_{m-1}, \quad (25)$$

where $\widetilde{V}_0 \in C^\infty\left(B(r_0,0), \mathbb{C}^N \otimes \Sigma_{2m}\right)$ *is the constant extension of* $V_0 \in \mathbb{C}^N \otimes \Sigma_{2m}$*. Then*

$$\operatorname{Str}(h_m(0,0)) = \operatorname{Str}\left(h^0_m(0,0)\right).$$

In other words, in the computation of $\operatorname{Str}(h_m(0,0))$ *given by (24), we may replace* A *by* A^0*.*

This is a substantial simplification, not only in that A^0 is a second-order differential operator with coefficients which are at most quadratic in y, but it also shows that $\operatorname{Str}(h_m(x,x))$ only depends on the curvatures Ω^θ and Ω^ε at the point x. One might regard the gist of the Index Formula for twisted Dirac operator as exhibiting the global quantity $\operatorname{index}(\mathcal{D}^+)$ as the integral of a form which may be locally computed. From this perspective, Proposition 5.25 does the job. Also, knowing in advance that $\operatorname{index}(\mathcal{D}^+)$ is insensitive to perturbations in h and ε, one suspects that $\operatorname{Str}(h_m(x,x))\,\nu_x$ can be expressed in terms of the standard forms which represent characteristic classes for TX and E. The Local Index Formula confirms this. Moreover, for low values of m, say $m = 1$ or 2 (i.e., for 2 and 4-manifolds), one can directly compute $\operatorname{Str}\left(h^0_m(0,0)\right)$ using (25), and thereby verify Theorem 5.21 and hence obtain Corollary 5.23 rather easily. For readers who have no use for the Local Index Theorem beyond dimension 4, this is sufficient. It requires more effort to prove Theorem 5.21 for general m. For lack of space, we cannot go into the details of this here, but they can be found in [16]. It is well worth mentioning that the appearance of the $\sinh$ function in the Local Index Formula has its roots in Mehler's formula for the heat kernel

$$e_a(x,y,t) = \frac{1}{\sqrt{4\pi\frac{\sinh(2at)}{2a}}} \exp\left(-\frac{1}{4\frac{\sinh(2at)}{2a}}\left(\cosh(2at)\left(x^2+y^2\right) - 2xy\right)\right).$$

of the generalized 1-dimensional heat problem

$$u_t = u_{xx} - a^2x^2u, \quad u(x,t) \in \mathbb{R},\ (y,t) \in \mathbb{R}\times(0,\infty),$$
$$u(x,0) = f(x),$$

where $0 \neq a \in \mathbb{R}$ is a given constant. A solution of this problem is given by

$$u(x,t) = \int_{-\infty}^{\infty} e_a(x,y,t)\, f(y)\, dy,$$

and this reduces to the usual formula as $a \to 0$. The nice idea of using Mehler's formula in a rigorous derivation of the Local Index Theorem appears to be due to Getzler in [24] and

[25], although it was at least implicitly involved in earlier heuristic supersymmetric path integral arguments for the Index Theorem. In the same vein, further simplifications and details can be found in [15] and [16].

While the Local Index Theorem (Theorem 5.21) is stated for twisted Dirac operators, the same proof may be applied to obtain the index formulas for an elliptic operator, possibly on a *nonspin* manifold, which is only locally of the form of twisted Dirac operator $\mathcal{D}^+$. Indeed, if $\mathcal{A}$ is such an operator and k is the heat kernel for $\mathcal{A}^*\mathcal{A} \oplus \mathcal{A}\mathcal{A}^*$, then from the spectral resolution of $\mathcal{A}$, we can still deduce from the asymptotic expansion of k that

$$\operatorname{index}(\mathcal{A}) = (4\pi)^{-m} \int_X \operatorname{Str}\left(h_m(x,x)\right) \nu_x,$$

where the supertrace Str is defined in the natural way. The crucial observation is that since $\mathcal{A}$ is locally in the form of a twisted Dirac operator, we can compute $\operatorname{Str}\left(h_m(x,x)\right)$ in exactly the same way (i.e., locally) as we have done. Since it is not easy to find first-order elliptic operators of geometrical significance which are not expressible in terms of locally twisted Dirac operators (or 0-th order perturbations thereof), the Local Index Formula for twisted Dirac operators is much more comprehensive than it would appear at first glance.

5.5 Standard geometric operators and applications

We state some special cases of the index formula for standard elliptic operators of natural geometric significance (and their twists), other than the twisted Dirac operator which was covered in Corollary 5.23. These geometric operators include

the signature operator $d + \delta : (1 + *)\,\Omega^*(M) \to (1 - *)\,\Omega^*(M)$,

the Euler-Dirac operator $d + \delta : \Omega^{even}(M) \to \Omega^{odd}(M)$, and

the Dolbeult-Dirac operator $\sqrt{2}\left(\bar\partial + \bar\partial^*\right) : \Omega^{0,even}(M) \to \Omega^{0,odd}(M)$.

Here, d is exterior derivative, δ is the exterior coderivative (i.e., the formal adjoint d^* of d), $*$ is the Hodge star operator, and $\bar\partial$ and its adjoint $\bar\partial^*$ are complex analogs of d and d^* on complex manifolds, which, along with $\Omega^{0,*}(M)$, will be defined. The index formula obtained for the above operators yields the Hirzebruch Signature Theorem, the Chern-Gauss-Bonnet Theorem, and the Hirzebruch-Riemann-Roch Theorem, respectively. Each of these is locally a twisted Dirac operator (or a direct sum of such), although space limitations do not permit us to demonstrate this in each case. Indeed, perhaps the best way to do this is uniformly is to introduce the notion of a Clifford module bundles (see [34], Chapter II, §5 and §6). We will just state the final results and introduce the definitions needed to understand them.

The Hirzebruch Signature Formula.

Let $\Lambda^k(X) \to X$ denote the bundle of complex exterior k-covectors over the compact, orientable C^∞ Riemannian manifold X of even dimension $n = 2m$ with metric tensor g. Let $\Omega^k(X) = C^\infty\left(\Lambda^k(X)\right)$ denote the space of C^∞ sections of $\Lambda^k(X)$, namely the space of $\mathbb{C}$-valued k-forms on X. We have the exterior derivative $d : \Omega^k(X) \to \Omega^{k+1}(X)$ and the codifferential $\delta : \Omega^{k+1}(X) \to \Omega^k(X)$ which is the formal adjoint of d. Let $*_k : \Lambda^k(X) \to \Lambda^{n-k}(X)$ be the Hodge star operator and define

$$\tau := \bigoplus_{k=0}^n \tau_k := \bigoplus_{k=0}^n i^{m+k(k-1)} *_k \text{ and } \Lambda^\pm(X) := (1 \pm \tau)\,\Lambda^*\left((T_{\mathbb{C}}X)^*\right).$$

For $\Omega^{\pm}(X) = C^{\infty}(\Lambda^{\pm}(X))$, we have the **Hirzebruch signature operator**

$$(d+\delta)^{\pm} : \Omega^{\pm}(X) \to \Omega^{\mp}(X).$$

Let $\mathcal{H}^m(X)_{\mathbb{R}}$ be the space of $\mathbb{R}$-valued m-forms α with $(d+\delta)\alpha = 0$. Recall from Example 4.5 that there is a quadratic form $Q : \mathcal{H}^m(X) \to \mathbb{R}$ given by

$$Q(\alpha) = \int_X \alpha \wedge \alpha = \int_X \langle \alpha, *_m \alpha \rangle_g \ \nu_g,$$

and for $n \equiv 0 \bmod 4$,

$$\operatorname{sig}(X) := \text{the signature of } Q = \operatorname{index}(d+\delta)^{+}.$$

(i.e., the analytical index of $(d+\delta)^{+}$). We describe the topological index of $(d+\delta)^{+}$ as follows. The total L-class of X is defined by

$$\mathbf{L}(X) = \prod\nolimits_{j=1}^{m} \frac{r_j}{\tanh r_j} = \sum\nolimits_{k=0}^{\infty} L_k\left(\sigma_1\left(r_1^2,\ldots,r_m^2\right),\ldots,\sigma_k\left(r_1^2,\ldots,r_m^2\right)\right)$$

in $H^*(X;\mathbb{Q})$, where the elementary symmetric polynomials $\sigma_i\left(r_1^2,\ldots,r_m^2\right)$ are identified with the Pontryagin classes of TX. In particular,

$$\begin{aligned}
&L_0 = 1,\ L_1(\sigma_1) = \frac{1}{3}\sigma_1,\ L_2(\sigma_1,\sigma_2) = \frac{7\sigma_2 - \sigma_1^2}{45}\\
&L_3(\sigma_1,\sigma_2,\sigma_3) = \frac{62\sigma_3 - 13\sigma_2\sigma_1 - 2\sigma_1^3}{945}\\
&L_4(\sigma_1,\ldots,\sigma_3) = \frac{381\sigma_4 - 71\sigma_3\sigma_1 - 19\sigma_2^2 + 22\sigma_2\sigma_1^2 - 3\sigma_1^4}{14175},\ldots.
\end{aligned} \tag{26}$$

As is proved in [34] (Theorem 13.9, p.256), the topological index of $(d+\delta)^{+}$ is $\mathbf{L}(X)[X]$, which is known as the L-genus of X, denoted by $L(X)$. The index formula then yields

Theorem 5.26 (Hirzebruch Signature Theorem) *Let X be a compact, oriented Riemannian $2m$-manifold, where m is even. Then*

$$\operatorname{sig}(X) = \mathbf{L}(X)[X] =: L(X) = L\text{-genus of } X.$$

In terms of the Pontryagin classes $p_k = p_k(TX)$, we have

$$\begin{aligned}
2m = 4 \quad &\Rightarrow \quad \operatorname{sig}(X) = \tfrac{1}{3} p_1 [X]\\
2m = 8 \quad &\Rightarrow \quad \operatorname{sig}(X) = \tfrac{1}{45}\left(7p_2 - p_1^2\right)[X],
\end{aligned}$$

in particular, and one may extend this using (26).

There is a twisted version of this theorem which we now describe. Let $E \to X$ be a Hermitian vector bundle with a covariant derivative ∇^E arising from a unitary connection 1-form ε on the frame bundle $\mathrm{U}(E)$. There is an exterior covariant derivative operator $D^{\varepsilon} : \Omega^k(X,E) \to \Omega^{k+1}(X,E)$ which generalizes $d : \Omega^k(X) \to \Omega^{k+1}(X)$, but $D^{\varepsilon} \circ D^{\varepsilon} \neq 0$ if ε is not flat. Moreover, D^{ε} has a formal adjoint $\delta^{\varepsilon} : \Omega^{k+1}(X,E) \to \Omega^k(X,E)$. For $W = \Lambda^*\left(\left(\mathbb{C}^{2m}\right)^*\right)$, the **twisted Hirzebruch signature operator** is

$$(d+\delta)^{E,+} := D^{\varepsilon} + \delta^{\varepsilon} : \Omega^{+}(X,E) \to \Omega^{-}(X,E), \tag{27}$$

As is proved in [34] (Theorem 13.9, p.256), we have

Theorem 5.27 (Twisted Hirzebruch Signature Theorem) *Let X be a compact, oriented Riemannian $2m$-manifold, where m is even, and let $E \to X$ be a Hermitian vector bundle with a covariant derivative ∇^E arising from a unitary connection 1-form ε on $\mathrm{U}(E)$. Then the index of the twisted Hirzebruch signature operator*

$$(d+\delta)^{E,+} : \Omega^+(X,E) \to \Omega^-(X,E)$$

(where $\Omega^{\pm}(X,E) := C^\infty(E \otimes \Lambda^{\pm}(X))$) is given by

$$\text{index}\left((d+\delta)^{E,+}\right) = (\mathbf{ch}_2(E) \smile \mathbf{L}(X))[X],$$

where $\mathbf{ch}_2(E) := \bigoplus_{j=0}^{m} 2^j ch_j(E)$.

The Gauss-Bonnet-Chern Formula.

Let $\Lambda^{even}(\mathbb{C}^{2m}) = \oplus_{k \text{ even}} \Lambda^k(\mathbb{C}^{2m})$ and $\Lambda^{odd}(\mathbb{C}^{2m}) = \oplus_{k \text{ odd}} \Lambda^k(\mathbb{C}^{2m})$. The **Euler operator** is

$$(d+\delta)^{\chi} := d+\delta : \Omega^{even}(X) \to \Omega^{odd}(X),$$

whose index (according to Hodge Theory) is the Euler characteristic $\chi(X)$. Indeed,

$$\begin{aligned}
\chi(X) &:= \dim\left(\bigoplus\nolimits_{k=0}^{m} H^{2k}(X)\right) - \dim\left(\bigoplus\nolimits_{k=1}^{m} H^{2k-1}(X)\right) \\
&= \dim\left(\bigoplus\nolimits_{k=0}^{m} \mathcal{H}^{2k}(X)\right) - \dim\left(\bigoplus\nolimits_{k=1}^{m} \mathcal{H}^{2k-1}(X)\right) \\
&= \dim(\text{Ker}\,(d+\delta)^{\chi}) - \dim\left(\text{Ker}\,((d+\delta)^{\chi})^*\right) \\
&= \text{index}\,((d+\delta)^{\chi}).
\end{aligned}$$

The topological index of $(d+\delta)^{\chi}$ is the evaluation, on the fundamental cycle $[X]$, of the Euler class which is represented by the **Gauss-Bonnet form**

$$\text{GB}\left(\Omega^\theta\right) = (-1)^m \,\text{Pf}\left(\tfrac{1}{2\pi}\Omega^\theta\right) := \frac{1}{2^{2m}\pi^m m!}\sum\nolimits_{(i)} \varepsilon_{i_1 \cdots i_{2m}} \Omega^\theta_{i_1 i_2} \wedge \cdots \wedge \Omega^\theta_{i_{2m-1} i_{2m}},$$

where $\Omega^\theta_{ij} = \frac{1}{2}\sum_{k,l} R_{ijkl}\varphi^k \wedge \varphi^l$ and the R_{ijkl} denote the components of the Riemann curvature tensor (for the Levi-Civita connection θ) relative to a local orthonormal frame field, with conventions such that R_{ijij} is the sectional curvature of the plane dual to $\varphi^i \wedge \varphi^j$. $\text{GB}\left(\Omega^\theta\right)$ is independent of the choice of local orthonormal framing. For $\dim X = 2$, $\text{GB}\left(\Omega^\theta\right)$ is $K\,dv$ where K is the Gaussian curvature and dv is the area element. The index formula for $(d+\delta)^{\chi}$ is

Theorem 5.28 (Gauss-Bonnet-Chern Theorem) *Let X be a compact, orientable, Riemannian manifold of even dimension $n = 2m$, and let $\mathcal{H}^k(X)$ be the space of harmonic k-forms on X. Then*

$$\text{index}\,((d+\delta)^{\chi}) = \sum\nolimits_{k=0}^{n} (-1)^k \dim\left(\mathcal{H}^k(X)\right) = \text{GB}(TX)[X] = \int_X \text{GB}\left(\Omega^\theta\right).$$

There is also a twisted version of the Gauss-Bonnet-Chern Theorem. We define the twisted Euler operator

$$(d+\delta)^{E,\chi} : \Omega^{even}(X,E) \to \Omega^{odd}(X,E) := (d+\delta)^{E}|_{\Omega^{even}(X,E)}$$

as the restriction $(d+\delta)^{E}|_{\Omega^{even}(X,E)}$ of the twisted DeRham-Dirac operator $(d+\delta)^{E}$ in (27). It happens that the index of the twisted Euler operator is only affected by the dimension of the twisting bundle E, rather than any actual twisting (i.e., nontriviality) of E.

Theorem 5.29 (Twisted Gauss-Bonnet-Chern Theorem) *Let X be a compact, oriented Riemannian $2m$-manifold, and let $E \to X$ be a Hermitian vector bundle with a covariant derivative ∇^{E} arising from a unitary connection 1-form ε on $\mathrm{U}(E)$. Then the index of the twisted Euler operator is given by*

$$\operatorname{index}\left((d+\delta)^{E,\chi}\right) = \dim E \cdot \mathrm{GB}[TX][X] = \dim E \cdot \chi(X).$$

The Generalized Yang-Mills Index Theorem

There are two interesting operators of Dirac type from which the signature operator and the Euler operator can be generated. Moreover these operators are of fundamental importance in the application of the Index Theorem to gauge theory. We have two decompositions (where $\Lambda^{\pm}(X) := (1\pm\tau)\Lambda^{*}\left((T_{\mathbb{C}}X)^{*}\right)$) of $\Lambda(X) := \Lambda^{*}\left((T_{\mathbb{C}}X)^{*}\right)$

$$\Lambda(X) = \Lambda^{+}(X) \oplus \Lambda^{-}(X) \text{ and } \Lambda(X) = \Lambda^{even}(X) \oplus \Lambda^{odd}(X)$$

By forming the intersections $\Lambda^{even(odd),\pm}(X) := \Lambda^{even(odd)}(X) \cap \Lambda^{\pm}(X)$, we have the finer decomposition

$$\Lambda(X) = \left(\Lambda^{even,+}(X) \oplus \Lambda^{odd,-}(X)\right) \oplus \left(\Lambda^{odd,+}(X) \oplus \Lambda^{even,-}(X)\right)$$

which we abbreviate by $W^{e}(X) \oplus W^{o}(X)$. Consequently, we have two generalized Dirac operators

$$\begin{aligned}\mathcal{D}^{W^{e}} &= d+\delta \in \operatorname{End}\left(\Omega^{even,+}(X) \oplus \Omega^{odd,-}(X)\right)\\ \mathcal{D}^{W^{o}} &= d+\delta \in \operatorname{End}\left(\Omega^{odd,+}(X) \oplus \Omega^{even,-}(X)\right),\end{aligned}$$

where $\Omega^{even(odd),\pm}(X) := C^{\infty}\left(\Lambda^{even(odd),\pm}(X)\right)$. Note that

$$\begin{aligned}&\mathcal{D}^{W^{e}+} : \Omega^{even,+}(X) \to \Omega^{odd,-}(X) \text{ and}\\ &\mathcal{D}^{W^{o}+} : \Omega^{odd,+}(X) \to \Omega^{even,-}(X).\end{aligned}$$

The signature operator can be written as

$$(d+\delta)^{+} = \mathcal{D}^{W^{e}+} \oplus \mathcal{D}^{W^{o}+} : \Omega^{even,+}(X) \oplus \Omega^{odd,+}(X) \to \Omega^{odd,-}(X) \oplus \Omega^{even,-}(X),$$

while the Euler operator is

$$(d+\delta)^{\chi} : \mathcal{D}^{W^{e}+} \oplus \left(\mathcal{D}^{W^{o}+}\right)^{*} : \Omega^{even,+}(X) \oplus \Omega^{even,-}(X) \to \Omega^{odd,-}(X) \oplus \Omega^{odd,+}(X).$$

We can also twist with a Hermitian vector bundle $E \to X$, to obtain the twisted operators

$$\mathcal{D}^{W^e,E} = (d+\delta)^E \in \operatorname{End}\left(\Omega^{even,+}(E) \oplus \Omega^{odd,-}(E)\right)$$
$$\mathcal{D}^{W^o,E} = (d+\delta)^E \in \operatorname{End}\left(\Omega^{odd,+}(E) \oplus \Omega^{even,-}(E)\right)$$

For reasons explained below, we call $\mathcal{D}^{W^e,E}$ and $\mathcal{D}^{W^o,E}$ **Yang-Mills-Dirac operators**. The following is equivalent to the twisted signature theorem and the twisted Gauss-Bonnet theorem combined.

Theorem 5.30 (The Yang-Mills-Dirac Index Theorem) *Let X be a compact, oriented Riemannian $2m$-manifold, where m is even, and let $E \to X$ be a Hermitian vector bundle with a covariant derivative ∇^E arising from a unitary connection 1-form ε on $\mathrm{U}(E)$. Then*

$$\operatorname{index}(\mathcal{D}^{W^e,E+}) = \tfrac{1}{2}\left(\mathbf{ch}_2(E) \smile \mathbf{L}(X)\right)[X] + \tfrac{1}{2}\dim E \cdot \chi(X) \text{ and}$$
$$\operatorname{index}(\mathcal{D}^{W^o,E+}) = \tfrac{1}{2}\left(\mathbf{ch}_2(E) \smile \mathbf{L}(X)\right)[X] - \tfrac{1}{2}\dim E \cdot \chi(X).$$

Proof Using $(d+\delta)^{E,+} = \mathcal{D}^{W^e,E+} \oplus \mathcal{D}^{W^o,E+}$, we have

$$\operatorname{index}((d+\delta)^{E,+}) = \operatorname{index}(\mathcal{D}^{W^e,E+}) + \operatorname{index}(\mathcal{D}^{W^o,E+})$$
$$\operatorname{index}(d+\delta)^{E,\chi} = \operatorname{index}(\mathcal{D}^{W^e,E+}) - \operatorname{index}(\mathcal{D}^{W^o,E+}).$$

The results follow from adding, subtracting and division by 2, since

$$\operatorname{index}((d+\delta)^{E,+}) = \left(\mathbf{ch}_2(E) \smile \mathbf{L}(X)\right)[X]$$
$$\operatorname{index}((d+\delta)^{E,\chi}) = \dim E \cdot \chi(X),$$

by Theorems 5.27 and 5.29. □

We now explain the "Yang-Mills-Dirac" nomenclature. Let $E = P \times_G \mathfrak{g}_{\mathbb{C}}$, where $P \to X$ is a principal G-bundle over a compact Riemannian 4-manifold X and $\mathfrak{g}_{\mathbb{C}}$ is the complexification of the Lie algebra of G. The kernel of the operator

$$\mathcal{T} : \Omega^1(E) \to \Omega^0(E) \oplus \Omega^2_-(E), \text{ given by } \mathcal{T}(\alpha) := \left(\delta^\omega \alpha, \tfrac{1}{2}(1-*)\, D^\omega \alpha\right),$$

can be regarded as the formal dimension of tangent space (at the class of the connection ω) of the manifold of moduli (space of orbits under the action of gauge transformations) of connections on P with self-dual curvature. The operator $\mathcal{T}$ bears a strong resemblance to the operator

$$\mathcal{D}^{W^o,E+} : \Omega^{odd,+}(E) \to \Omega^{even,-}(E),$$

which is the restriction of $(d+\delta)^E := D^\omega + \delta^\omega \in \operatorname{End}(\Omega^*(E))$. Indeed, for

$$\pi_\pm := \tfrac{1}{2}(1 \pm \tau) : \Omega^*(E) \to \Omega^{*,\pm}(E),$$

we have $\pi_- \circ \mathcal{T} = \mathcal{D}^{W^o,E} \circ \pi_+$. We also have isomorphisms

$$\pi_+|_{\Omega^1(E)} : \Omega^1(E) \cong \Omega^{odd,+}(E) \subset \Omega^1(E) \oplus \Omega^3(E), \text{ and}$$
$$\pi_-|_{\Omega^0(E)\oplus\Omega^2_-(E)} : \Omega^0(E) \oplus \Omega^2_-(E) \cong \Omega^{even,-}(E).$$

Thus, using $\mathbf{ch}_2(E) = \dim E + 2ch_1(E) + 4ch_2(E)$, it follows that

$$\begin{aligned}
\operatorname{index}(\mathcal{T}) &= \operatorname{index}\left(\pi_-|_{\Omega^0(E)}\right)^{-1} \circ \mathcal{D}^{W^o,E} \circ \left(\pi_-|_{\Omega^0(E)\oplus\Omega^2_-(E)}\right) \\
&= \operatorname{index}(\mathcal{D}^{W^o,E+}) = \tfrac{1}{2}\left(\mathbf{ch}_2(E) \smile \mathbf{L}(X)\right)[X] - \tfrac{1}{2}\dim E \cdot \chi(X) \\
&= \tfrac{1}{2}\left((\dim E + 2ch_1(E) + 4ch_2(E)) \smile \mathbf{L}(X)\right)[X] - \tfrac{1}{2}\dim E \cdot \chi(X) \\
&= \tfrac{1}{2}\left(4ch_2(E)[X] + \dim E \cdot \mathbf{L}(X)[X]\right) - \tfrac{1}{2}\dim E \cdot \chi(X) \\
&= 2ch_2(E)[X] - \tfrac{1}{2}\dim E \cdot \left(\chi(X) - \operatorname{sig}(X)\right),
\end{aligned}$$

in agreement with the computation in the proof of Theorem 6.1 in [7].

The Hirzebruch-Riemann-Roch Formula

Let X be a *complex manifold* with $n = \dim X = 2\dim_{\mathbb{C}} X = 2m$. In other words, X is a smooth n-manifold, and there is a covering $\{U\}$ of X and a collection $\{\varphi_U\}$ of coordinate charts $\varphi_U : U \to \mathbb{C}^m$, such that $\varphi_V \circ \varphi_U^{-1} : \varphi_U(U \cap V) \to \varphi_V(U \cap V)$ is holomorphic (i.e., $\left(\varphi_V \circ \varphi_U^{-1}\right)_* : T_{\varphi_U(x)}\mathbb{C}^m \to T_{\varphi_V(p)}\mathbb{C}^m$ is complex linear for each $x \in U \cap V$). The tangent spaces T_xX then possess a well-defined map $J_x \in T_xX$ (with $J_x^2 = -\operatorname{Id}_x$) which corresponds to multiplication by $i = \sqrt{-1}$ under $(\varphi_U)_* : T_xX \to T_{\varphi_U(x)}\mathbb{C}^m \cong \mathbb{C}^m$ (i.e., $J_x(X) = (\varphi_U)_*^{-1}(i\varphi_U(X))$). The bundle automorphism $J \in \operatorname{End}(TX)$ is known as the *complex structure* of the complex manifold X. While it is tempting to explicitly make T_xX a complex vector space by defining iX to be JX for $X \in T_xX$, this ultimately leads to profound confusion, since it is customary (and of great utility) to consider the complexification of the real vector space T_xX, namely

$$(T_{\mathbb{C}}X)_x := \mathbb{C} \otimes T_xX = \{V + iW : V, W \in T_xX\}.$$

of complex dimension $2m$. The problem is that multiplication by i in $T_{\mathbb{C}}X$ is not the same as the complex linear extension of J to $T_{\mathbb{C}}X$. In particular while J preserves the real subspace $T_xX \subset (T_{\mathbb{C}}X)_x$, multiplication by i does not. Thus, we use J_x instead of i for the complex structure on T_xX. Let

$$\varphi_U(x) = \left(z^1(x), \ldots, z^m(x)\right) = \left(x^1(x) + iy^1(x), \ldots, x^m(x) + iy^m(x)\right).$$

We define $\mathbb{C}$-valued, $\mathbb{R}$-linear functionals dz^j and $d\bar{z}^j$ on T_xX via

$$dz^j := dx^j + idy^j : T_xX \to \mathbb{C} \text{ and } d\bar{z}^j := dx^j - idy^j : T_xX \to \mathbb{C}$$

Any $\mathbb{R}$-linear functional on T_xX, such as dz^k or $d\bar{z}^k$, extends uniquely to a $\mathbb{C}$-linear functional on $(T_{\mathbb{C}}X)_x$, and we use the same symbols to denote these extensions; i.e., dz^j, $d\bar{z}^j \in (T_{\mathbb{C}}X)_x^*$. The local complex vector fields (local sections of $T_{\mathbb{C}}X$) $\partial_{z^k} := \frac{1}{2}\left(\partial_{x^k} - i\partial_{y^k}\right)$ and $\partial_{\bar{z}^k} := \frac{1}{2}\left(\partial_{x^k} + i\partial_{y^k}\right)$ are dual to dz^j and $d\bar{z}^j \in (T_{\mathbb{C}}X)^*$, in the sense that $dz^j(\partial_{z^k}) = \delta^j_k$, $d\bar{z}^j(\partial_{\bar{z}^k}) = \delta^j_k$ and $d\bar{z}^j(\partial_{z^k}) = dz^j(\partial_{\bar{z}^k}) = 0$. There is complex-linear extension of J_x to $(T_{\mathbb{C}}X)_x$. We denote this extension by the same symbol J_x. Since $J_x^2 = -\operatorname{Id}$, the eigenvalues of J_x are i and $-i$, and the eigenspaces of $J_x \in \operatorname{End}((T_{\mathbb{C}}X)_x)$ are

$$T_x^{1,0}X := \{V - iJV : V \in T_xX\} \text{ and } T_x^{0,1}X := \{V + iJV : V \in T_xX\},$$

respectively. Note that $\{\partial_{z^1},\ldots,\partial_{z^m}\}$ and $\{\partial_{\bar{z}^1},\ldots,\partial_{\bar{z}^m}\}$ are local framings of $C^\infty\left(T_x^{1,0}X\right)$ and $C^\infty\left(T_x^{0,1}X\right)$, respectively. We set

$$\Lambda^{p,0}\left(T_{\mathbb{C}}X^*\right)_x := \text{the vector space of all anti-symmetric multi-complex-linear functionals defined on } T_x^{1,0}X\times\overset{p}{\cdots}\times T_x^{1,0}X.$$

The $\Lambda^{p,0}\left(T_{\mathbb{C}}X^*\right)_x$ are the fibers of a complex vector bundle $\Lambda^{p,0}\left(T_{\mathbb{C}}X^*\right)\to X$. Let $\Omega^{p,0}(X)$ be the space of C^∞ sections of $\Lambda^{p,0}\left(T_{\mathbb{C}}X^*\right)$; i.e., $\Omega^{p,0}(X) := C^\infty\left(\Lambda^{p,0}\left(T_{\mathbb{C}}X^*\right)\right)$. On a coordinate neighborhood U, such a section is of the form $\frac{1}{p!}\sum_{(j)} f_{j_1\cdots j_p}dz^{j_1}\wedge\cdots\wedge dz^{j_p}$, where the $f_{j_1\cdots j_p}\in C^\infty(U,\mathbb{C})$ are anti-symmetric in $j_1\cdots j_p$. Similarly, we may define the bundles $\Lambda^{0,q}\left(T_{\mathbb{C}}X^*\right)$, and the space $\Omega^{0,q}(X) := C^\infty\left(\Lambda^{0,q}\left(T_{\mathbb{C}}X^*\right)\right)$ of sections which locally are of the form $\frac{1}{q!}\sum_{(k)} f_{k_1\cdots k_q}d\bar{z}^{k_1}\wedge\cdots\wedge d\bar{z}^{k_q}$. More generally, one has the bundles $\Lambda^{p,q}\left(T_{\mathbb{C}}X^*\right)$ whose sections in $\Omega^{p,q}(X) := C^\infty\left(\Lambda^{p,q}\left(T_{\mathbb{C}}X^*\right)\right)$ are called forms of bidegree (p,q), locally of the form

$$\frac{1}{p!q!}\sum_{(j)(k)} f_{j_1\cdots j_p;k_1\cdots k_q}dz^{j_1}\wedge\cdots\wedge dz^{j_p}\wedge d\bar{z}^{k_1}\cdots\wedge d\bar{z}^{k_q}.$$

Note that $\Omega^l(X,\mathbb{C}) := \mathbb{C}\otimes\Omega^l(X,\mathbb{R})\cong\sum_{p+q=l}\Omega^{p,q}(X)$. There is an operator $\bar{\partial}:\Omega^{0,q}(X)\to\Omega^{0,q+1}(X)$ given locally by

$$\bar{\partial}(\textstyle\sum_{(k)} f_{k_1\cdots k_q}d\bar{z}^{k_1}\wedge\cdots\wedge d\bar{z}^{k_q}) := \textstyle\sum_{(k)}\displaystyle\sum_{k_0=1}^{m}\partial_{\bar{z}^{k_0}}(f_{k_1\cdots k_q})d\bar{z}^{k_0}\wedge d\bar{z}^{k_1}\wedge\cdots\wedge d\bar{z}^{k_q}.$$

More generally, one analogously defines $\bar{\partial}:\Omega^{p,q}(X)\to\Omega^{p,q+1}(X)$, as well as $\partial:\Omega^{p,q}(X)\to\Omega^{p+1,q}(X)$. While we have given these operators locally, they are independent of local holomorphic coordinates. The operator $\partial+\bar{\partial}$ is the restriction of the usual exterior derivative on $\Omega^{p,q}(X)\subset\Omega^{p+q}(X,\mathbb{C})$, namely

$$\partial+\bar{\partial} = d|_{\Omega^{p,q}(X)}:\Omega^{p,q}(X)\to\Omega^{p+1,q}(X)+\Omega^{p,q+1}(X).$$

We have $\partial^2=0$, $\partial\bar{\partial}+\bar{\partial}\partial=0$ and $\bar{\partial}^2=0$, since $0=d^2=\left(\partial+\bar{\partial}\right)^2=\partial^2\oplus\left(\partial\bar{\partial}+\bar{\partial}\partial\right)\oplus\bar{\partial}^2$. In particular, since $\bar{\partial}^2=0$, we have a chain complex $0\to\Omega^{0,0}(X)\xrightarrow{\bar{\partial}}\Omega^{0,1}(X)\xrightarrow{\bar{\partial}}\cdots\xrightarrow{\bar{\partial}}\Omega^{0,m}(X)$ and **Dolbeault cohomology spaces**

$$H^{0,q}(X) := \frac{\operatorname{Ker}\left(\bar{\partial}|_{\Omega^{0,q}(X)}\right)}{\bar{\partial}\left(\Omega^{0,q-1}(X)\right)}.$$

In order to define harmonic representatives of Dolbeault cohomology classes, we need a Hermitian metric to define an adjoint $\bar{\partial}^*:\Omega^{0,q+1}(X)\to\Omega^{0,q}(X)$, for $\bar{\partial}:\Omega^{0,q}(X)\to\Omega^{0,q+1}(X)$. Suppose that we are given a Riemannian metric h on X, so that for all $x\in X$ and $V,W\in T_xX$, we have

$$h(JV,W) = -h(V,JW),\ \text{or equivalently}\ h(JV,JW) = h(V,W).$$

Such can always found by setting $h(V,W) = h_0(JV,JW) + h_0(V,W)$ for an arbitrary metric h_0. Then h_x uniquely extends to a complex bilinear form $h_{\mathbb{C}}$ on $\mathbb{C}\otimes_{\mathbb{R}} T_xX$. Define a Hermitian metric H_x (complex linear in first slot and conjugate linear in the second slot) on $\mathbb{C}\otimes_{\mathbb{R}} T_xX$ by $H(V,W) := (h_{\mathbb{C}})(V,\overline{W})$. There is a conjugate-linear bundle isomorphism $\flat : (T_{\mathbb{C}}X)_x \to (T_{\mathbb{C}}X)^*_x$ given, for $V,W \in (T_{\mathbb{C}}X)_x$, by $\flat_x(V)(W) = H(W,V)$. We have $\flat\left(T^{1,0}X\right) = \Lambda^{1,0}(TX^*)$ and $\flat\left(T^{0,1}X\right) = \Lambda^{0,1}(TX^*)$. There is a conjugate-linear inverse to $\flat$, denoted by $\# : (T_{\mathbb{C}}X)^*_x \to (T_{\mathbb{C}}X)_x$. Now H on $(T_{\mathbb{C}}X)_x$ induces a Hermitian inner product (still denoted by H) on $\Lambda^1\left((T_{\mathbb{C}}X)^*_x\right)$, given for $\alpha,\beta \in \Lambda^1\left((T_{\mathbb{C}}X)^*_x\right)$, by $H(\alpha,\beta) = H(\#\beta,\#\alpha)$, where the switch in the order is due to the conjugate linearity of $\#$. We then have an induced Hermitian inner product on $\Lambda^k\left((T_{\mathbb{C}}X)^*_x\right)$ for $1 \le k \le m$, where $\{e_{i_1}\wedge\cdots\wedge e_{i_k} : i_1 < \cdots < i_k\}$ is an orthonormal basis for $\Lambda^k\left((T_{\mathbb{C}}X)^*_x\right)$ if $e_1,\ldots,e_m$ is an orthonormal basis of $\Lambda^1\left((T_{\mathbb{C}}X)^*_x\right)$. By restriction, we have Hermitian inner products on $\Lambda^{p,q}\left((T_{\mathbb{C}}X)^*_x\right)$ as well.

Let $(\cdot,\cdot)$ be the Hermitian L^2 inner product on $\Omega^{0,q}(X)$ induced by H :

$$(\alpha,\beta) := \int_X H(\alpha,\beta)\, v_h, \text{ with } \|\alpha\| := \sqrt{(\alpha,\alpha)}.$$

We may then speak of the formal adjoint $\bar\partial^* : \Omega^{p,q+1}(X) \to \Omega^{p,q}(X)$ of $\bar\partial : \Omega^{p,q}(X) \to \Omega^{p,q+1}(X)$, having the property $\left(\bar\partial\alpha,\beta\right) = \left(\alpha,\bar\partial^*\beta\right)$. In order to exhibit a formula for $\bar\partial^*$, we have a $\mathbb{C}$-linear star operator $* : \Omega^k(X,\mathbb{C}) \to \Omega^{n-k}(X,\mathbb{C})$ determined by $\alpha\wedge*\beta := h_{\mathbb{C}}(\alpha,\beta)\,v_h$ for $\alpha,\beta\in\Omega^k(X,\mathbb{C})$. Since $\dim_{\mathbb{R}} X = 2m$ is even, $*^2|_{\Omega^k(X,\mathbb{C})} = (-1)^k\,\mathrm{Id}$ and the formal adjoint of $d : \Omega^k(X,\mathbb{R}) \to \Omega^{k+1}(X,\mathbb{R})$ is $\delta := -*d*$. By the $\mathbb{C}$-linearity of $*$ and the fact that d is the $\mathbb{C}$-linear extension of $\partial+\bar\partial$, δ has the $\mathbb{C}$-linear extension

$$\delta = -*d* = -*\left(\partial+\bar\partial\right)* = (-*\partial*)\oplus\left(-*\bar\partial*\right)$$

We have $*\Omega^{p,q}(X) = \Omega^{m-q,m-p}(X)$, from which it follows that

$$\left(-*\bar\partial*\right)\left(\Omega^{p+1,q}(X)\right) \subseteq \Omega^{p,q}(X), \text{ while } (-*\partial*)\left(\Omega^{p+1,q}(X)\right) \subseteq \Omega^{p+1,q-1}(X).$$

Thus, for $\alpha\in\Omega^{p,q}(X)$ and $\beta\in\Omega^{p+1,q}(X)$, we have $H(d\alpha,\beta) = H(\partial\alpha+\bar\partial\alpha,\beta) = H(\partial\alpha,\beta)$, and $H(\alpha,\delta\beta) = H(\alpha,-*\partial*\beta - *\bar\partial*\beta) = H(\alpha,-*\bar\partial*\beta)$. Thus,

$$\begin{aligned}\int_X H(\partial\alpha,\beta)v_h &= \int_X H(d\alpha,\beta)v_h = \int_X h_{\mathbb{C}}(d\alpha,\bar\beta)v_h = \int_X h_{\mathbb{C}}(\alpha,\delta\bar\beta)v_h \\ &= \int_X h_{\mathbb{C}}(\alpha,\overline{\delta\beta})v_h = \int_X H(\alpha,\delta\beta)v_h = \int_X H(\alpha,\left(-*\bar\partial*\right)\beta)v_h.\end{aligned}$$

Similarly, $\int_X H(\bar\partial\alpha,\beta)v_h = \int_X H(\alpha,(-*\partial*)\beta)v_h$. Thus, the formal adjoints of ∂ and $\bar\partial$ are given by $\partial^* = -*\bar\partial*$ and $\bar\partial^* = -*\partial*$. We define the space of **harmonic** $(0,q)$-forms by

$$\mathcal{H}^{0,q}(X) := \left\{\alpha\in\Omega^{0,q}(X) : \left(\bar\partial+\bar\partial^*\right)\alpha = 0\right\}.$$

We omit the standard proof (see [29, p.84]) of

Theorem 5.31 (Hodge Decomposition) *There is an orthogonal decomposition*

$$\begin{aligned}\Omega^{0,q}(X) &= \mathcal{H}^{0,q}(X)\oplus\bar\partial\left(\Omega^{0,q-1}(X)\right)\oplus\bar\partial^*\left(\Omega^{0,q+1}(X)\right) \\ &= \mathcal{H}^{0,q}(X)\oplus\left(\bar\partial\bar\partial^*\right)\left(\Omega^{0,q}(X)\right)\oplus\left(\bar\partial^*\bar\partial\right)\left(\Omega^{0,q}(X)\right).\end{aligned}$$

Corollary 5.32 *Suppose that* $\bar{\partial}\gamma = 0$ *for some* $\gamma \in \Omega^{0,q}(X)$*. There is a unique* $\alpha \in \mathcal{H}^{0,q}(X)$*, such that for some* $\beta \in \Omega^{0,q-1}(X)$*,* $\gamma = \alpha + \bar{\partial}\beta$*. In other words, every cohomology class in* $H^{0,q}(X)$ *has a unique harmonic representative.*

Proof Theorem 5.31 yields a unique $\alpha \in \mathcal{H}^{0,q}(X)$ such that

$$\gamma = \alpha + \bar{\partial}\beta + \bar{\partial}^*\beta'$$

for some $\beta \in \Omega^{0,q-1}(X)$ and $\beta' \in \Omega^{0,q+1}(X)$. Now

$$\begin{aligned} 0 &= \bar{\partial}\gamma = \bar{\partial}\alpha + \bar{\partial}^2\beta + \bar{\partial}\bar{\partial}^*\beta' = \bar{\partial}\bar{\partial}^*\beta' \\ &\Rightarrow \left(\bar{\partial}\bar{\partial}^*\beta', \beta'\right) = 0 \Rightarrow \left\|\bar{\partial}^*\beta'\right\|^2 = 0 \Rightarrow \bar{\partial}^*\beta' = 0. \end{aligned}$$

□

Suppose that ∇ is the covariant derivative for the Levi-Civita connection of the Riemannian metric h. It is always possible to choose coordinates about a point $x \in X$ such that the coordinate vector fields have vanishing ∇-derivatives at x. However, it is not necessarily the case that such coordinates can be chosen of the form $\left(x^1, y^1, \ldots, x^m, y^m\right)$, where $\left(z^1, \ldots, z^m\right) = \left(x^1 + iy^1, \ldots, x^m + iy^m\right)$ is a complex-analytic coordinate chart. If for each $x \in X$ such coordinates can be found, then the complex manifold X with Riemannian metric h is a Kähler manifold. While one can take this to be the definition of a Kähler manifold, usually one of the other equivalent conditions in the following theorem is taken to be the definition.

Theorem 5.33 *Let* X *be a complex manifold with complex structure* J*, and Riemannian metric* h*, with Levi-Civita covariant derivative* ∇*. The following are equivalent.*

I*. About each* $x \in X$*, there is a complex chart* $\left(x^1 + iy^1, \ldots, x^m + iy^m\right)$*,*
such that $\nabla\left(\partial_{x^i}\right) = \nabla\left(\partial_{y^i}\right) = 0$ *at* x*.*

II*.* $\nabla J = 0$ *(i.e.,* $(\nabla J)(X) = J(\nabla X) - \nabla(J(X)) = 0$*).*

III*. The 2-form Kähler 2-form* $\kappa \in \Omega^2(X, \mathbb{R})$*, given by*
$\kappa(X, Y) := h(X, JY)$*, is closed (i.e.,* $d\kappa = 0$*).*

Note that $\kappa \in \Omega^{1,1}(X)$, since locally

$$\kappa\left(\partial_{z^j}, \partial_{\bar{z}^k}\right) = h_{\mathbb{C}}\left(J\partial_{z^j}, \partial_{\bar{z}^k}\right) = h_{\mathbb{C}}\left(i\partial_{z^j}, \partial_{\bar{z}^k}\right) = i\left(h_{\mathbb{C}}\right)_{j\bar{k}}, \text{ and}$$

$$\kappa\left(\partial_{\bar{z}^j}, \partial_{\bar{z}^k}\right) = \kappa\left(\partial_{z^j}, \partial_{z^k}\right) = 0 \Rightarrow \kappa = i\sum\nolimits_{j,k}\left(h_{\mathbb{C}}\right)_{j\bar{k}} dz^j \wedge d\bar{z}^k.$$

Until further notice, we assume that X (with Riemannian metric h) is a Kähler manifold. Then, it turns out that $\sqrt{2}\left(\bar{\partial} + \bar{\partial}^*\right) : \Omega^{0,even}(X) \to \Omega^{0,odd}(X)$, strongly resembles a standard Dirac operator, just as the bundle $\Lambda^{0,*}(X)$ strongly resembles a spinor bundle $\Sigma(X)$, with $\Lambda^{0,even}(X)$ and $\Lambda^{0,odd}(X)$ corresponding to $\Sigma^+(X)$ and $\Sigma^-(X)$. However, a Kähler manifold need not admit a spin structure (e.g., $\mathbb{C}P^2$), and so the analogy cannot be precise. The underlying problem is that $\Lambda^{0,*}\left(T_{\mathbb{C}}X^*\right)$ is associated to the unitary frame bundle UX with group $\mathrm{U}(m)$, instead of the $\mathrm{Spin}(2m)$-bundle P, where $P \to FX$ is a spin structure over the orthonormal frame bundle FX. However, it is the case that for any coordinate ball $B \subseteq X$, the bundle $\Lambda^{0,*}\left(T_{\mathbb{C}}X^*\right)|_B$ is isomorphic to a twist of the spinor bundle $\Sigma(B) = P \times_{\mathrm{Spin}(2m)} \Sigma_{2m}$ (associated to a spin structure $P \to FX|_B$) with a line bundle L such that $L \otimes L \cong \Lambda^{0,m}(T_{\mathbb{C}}X^*)$; i.e., L is a square-root of the *canonical line bundle* $\Lambda^{0,m}(T_{\mathbb{C}}X^*)$ of the complex manifold X. Moreover, under this isomorphism,

$$\sqrt{2}\left(\bar{\partial} + \bar{\partial}^*\right)^{even} : \Omega^{0,even}(B) \to \Omega^{0,odd}(B)$$

is of the form of the standard Dirac operator over B twisted by L, namely

$$\mathcal{D}^{L+} : C^\infty(L \otimes \Sigma^+(B)) \to C^\infty(L \otimes \Sigma^-(B)).$$

While we will not go into the details here, there is a nice treatment of the necessary machinery (e.g., Spin^c structures) in [34, Appendix D] that allows us to apply the Local Index Theorem (Theorem 5.21) to obtain

Theorem 5.34 (Hirzebruch-Riemann-Roch Theorem) *Let X be a compact Kähler manifold. Then*

$$\mathrm{Index}\left(\left(\bar{\partial} + \bar{\partial}^*\right) : \Omega^{0,even}(X) \to \Omega^{0,odd}(X)\right) = \mathbf{Td}(TX)[X].$$

There is also a twisted version of the Hirzebruch-Riemann-Roch Theorem. Let $E \to X$ be a Hermitian vector bundle with a covariant derivative ∇^E arising from a unitary connection 1-form ε on $\mathrm{U}(E)$. We have the spaces $\Omega^{p,q}(X,E) := C^\infty(E \otimes \Lambda^{p,q}(T_{\mathbb{C}}X^*))$ of E-valued forms of bidegree (p,q), which are locally of the form (in multi-index notation) $\phi = \frac{1}{p!q!}\sum_{(j),(k)} f_{(j)_p;(k)_q} \otimes dz^{(j)_p} \wedge d\bar{z}^{(k)_q}$, where $f_{(j)_p;(k)_q}$ is a local section of E. Moreover, there are global operators

$$\partial^E : \Omega^{p,q}(X,E) \to \Omega^{p+1,q}(X,E) \text{ and } \bar{\partial}^E : \Omega^{p,q}(X,E) \to \Omega^{p,q+1}(X,E)$$

determined locally by

$$\partial^E\phi = \frac{1}{p!q!}\sum_{h,(j),(k)} \nabla^E_{\partial_{z^h}}\left(f_{(j)_p;(k)_q}\right) dz^h \wedge dz^{(j)_p} \wedge d\bar{z}^{(k)_q}, \text{ and}$$
$$\bar{\partial}^E\phi = \frac{1}{p!q!}\sum_{h,(j),(k)} \nabla^E_{\partial_{\bar{z}^h}}\left(f_{(j)_p;(k)_q}\right) d\bar{z}^h \wedge dz^{(j)_p} \wedge d\bar{z}^{(k)_q}. \tag{28}$$

Note that just as $\bar{\partial} + \bar{\partial}^*$ was shown to be locally a twisted Dirac operator (twisted locally by L), $\bar{\partial}^E + \bar{\partial}^{E*}$ is also locally a twisted Dirac operator (twisted locally by $E \otimes L$). As a consequence, we have

Theorem 5.35 (Twisted Hirzebruch-Riemann-Roch Theorem) *Let X be a compact Kähler manifold and let $E \to X$ be a Hermitian vector bundle. Then*

$$\mathrm{Index}\left(\bar{\partial}^E + \bar{\partial}^{E*} : \Omega^{0,even}(X,E) \to \Omega^{0,odd}(X,E)\right) = \left(\mathbf{ch}(E) \smile \mathbf{Td}(X)\right)[X].$$

There is a version of Theorem 5.35 which applies when $E \to X$ is holomorphic. We say that the complex vector bundle $\pi_E : E \to X$ of complex fiber dimension N over the complex manifold X is *holomorphic* if E has the structure of a complex manifold, and about each point $x \in X$ there is a neighborhood B and a biholomorphic vector bundle map $\phi_B : \pi_E^{-1}(B) \to B \times \mathbb{C}^N$. If E is a holomorphic Hermitian bundle over a Kähler manifold and ∇^E is the covariant derivative of a complex connection 1-form on GL(E) which is a Hermitian connection and compatible with the complex structure on E, then $\bar{\partial}^E$ defined in (28) satisfies $\bar{\partial}^E \circ \bar{\partial}^E = 0$. By definition, the kernel of $\bar{\partial}_E$ consists of the *holomorphic sections of* E. We have an exact sequence

$$\Omega^{0,0}(X,E) \overset{\bar{\partial}_E}{\to} \cdots \overset{\bar{\partial}_E}{\to} \Omega^{0,q}(X,E) \overset{\bar{\partial}_E}{\to} \Omega^{0,q+1}(X,E) \overset{\bar{\partial}_E}{\to} \cdots ,$$

along with cohomology groups

$$H^q(\mathcal{O}_E) := \frac{\operatorname{Ker}\left(\bar{\partial}_E|_{\Omega^{0,q}(X,E)}\right)}{\bar{\partial}_E\left(\Omega^{0,q-1}(X,E)\right)}.$$

Hence, by a Hodge theoretic proof strictly analogous to that of Theorem 5.31, when E is Hermitian and holomorphic we have

$$H^q(\mathcal{O}_E) \cong \mathcal{H}^q(E) := \operatorname{Ker}\left(\left(\bar{\partial}^E + \bar{\partial}^{E\,*}\right)|_{\Omega^{0,q}(X,E)}\right).$$

The following theorem is then immediate.

Theorem 5.36 (Holomorphic Hirzebruch-Riemann-Roch Theorem) *Let X be a compact Kähler manifold and let $E \to X$ be a Hermitian holomorphic vector bundle. Then the holomorphic Euler characteristic of E is given by*

$$\begin{aligned}
&\chi(E) := \sum\nolimits_{q=0}^{m} (-1)^q \dim H^q(\mathcal{O}_E) \\
&= \operatorname{Index}\left(\bar{\partial}^E + \bar{\partial}^{E*} : \Omega^{0,even}(X,E) \to \Omega^{0,odd}(X,E)\right) \\
&= (\mathbf{ch}(E) \smile \mathbf{Td}(X))[X].
\end{aligned}$$

In particular, $\chi(E)$, which seems to depend on the holomorphic structure of E, actually only depends on the topology of E.

6 Generalizations

6.1 The G-equivariant Index Theorem

Let G be a compact Lie group group which acts smoothly to the left on a compact n-manifold X, via a C^∞ map $L : G \times X \to X$. We write $g \cdot x = L_g(x) = L(g,x)$. Let $\pi_E : E \to X$ be a C^∞ complex vector bundle over X and suppose that there is a left action of G on E such that for all $g \in G$ and $e \in E$, we have that $\pi(g \cdot e) = g \cdot \pi(e)$ and $e \mapsto g \cdot e$ defines a linear map $E_x \to E_{g \cdot x}$. Then $\pi_E : E \to X$ is called a G-vector bundle. For a section $u \in C^\infty(E)$ and $g \in G$, we have a section $g \cdot u \in C^\infty(E)$ defined by

$$(g \cdot u)(x) = g \cdot \left(u(g^{-1} \cdot x)\right) \text{ for } x \in X.$$

Let $\pi_F : F \to X$ be another G-vector bundle.

Definition 6.1 An operator $P : C^\infty(E) \to C^\infty(F)$ is a G**-operator** if for all $g \in G$, we have $P(g \cdot u) = g \cdot P(u)$.

If P is an elliptic (pseudo-) differential G-operator, then $\operatorname{Ker} P$ and $\operatorname{Coker} P$ are preserved by the action of G, and hence are representation spaces for G. Recall that $R(G)$ is the Grothendieck ring obtained from the abelian semi-group of equivalence classes of finite-dimensional representations of G with addition induced by the direct sum. Tensor product of representations induces a multiplication on $R(G)$ making it a ring.

Definition 6.2 The **index** of an elliptic G-operator $P : C^\infty(E) \to C^\infty(F)$ is defined by

$$\operatorname{ind}_G P = [\operatorname{Ker} P] - [\operatorname{Coker} P] \in R(G).$$

Moreover, for $g \in G$, we define

$$\operatorname{ind}_g P = \text{trace}\,(g : \operatorname{Ker} P \to \operatorname{Ker} P) - \text{trace}\,(g : \operatorname{Coker} P \to \operatorname{Coker} P)\,.$$

To formulate the G-index Theorem, we need to define the topological index of an elliptic G-operator P in terms of its principal symbol, say $\sigma(P) \in C^\infty\left(\operatorname{Hom}(\pi^* E, \pi^* F)\right)$. Note that the action of G on X induces an action on $T^*X \cong TX$ (via a G-invariant metric on X). Using the fact that P is a G-operator, $\sigma(P)$ defines an element of $[\sigma(P)] \in K_G(TX)$. We proceed along the same lines of the case where G is trivial. One first selects a G-equivariant embedding $f : X \to \mathbb{R}^{n+m}$ where $\mathbb{R}^{n+m}$ is a representation space for G. The existence of such is a consequence of the Peter-Weyl Theorem (see [39]). Although it is more difficult to prove (especially if G is nonabelian), we have a Thom isomorphism $\varphi_{TN\to TX} : K_G(TX) \to K_G(TN)$ and an extension homomorphism $h : K_G(TN) \to K_G(T\mathbb{R}^{n+m})$. The composition $h \circ \varphi_{TN\to TX}$ gives us a homomorphism

$$f_! := h \circ \varphi_{TN\to TX} : K_G(TX) \to K_G(T\mathbb{R}^{n+m}).$$

Moreover, for $i : \{\mathbf{0}\} \to \mathbb{R}^{n+m}$, we have $i_!^{-1} : K_G(T\mathbb{R}^{n+m}) \to K_G(\{\mathbf{0}\}) = R(G)$.

Definition 6.3 For an elliptic G-operator $P : C^\infty(E) \to C^\infty(F)$ with symbol class $[\sigma(P)] \in K_G(T^*X)$, the **topological G-index** of P is defined by

$$\operatorname{top}_G\text{-}\operatorname{ind}(P) := i_!^{-1} \circ f_!([\sigma(P)]) \in R(G).$$

For $g \in G$, the **topological g-index** of P is defined by

$$\operatorname{top}_g\text{-}\operatorname{ind}(P) := \text{trace}\left((\operatorname{top}_G\text{-}\operatorname{ind}(P))\,(g)\right).$$

The proof of the Index Formula (Theorem 5.9) generalizes (see [12]) without difficulty to yield the following

Theorem 6.4 (G-index formula) *For an elliptic G-operator $P : C^\infty(E) \to C^\infty(F)$ over a compact manifold X, we have*

$$\operatorname{ind}_G P = \operatorname{top}_G\text{-}\operatorname{ind}(P).$$

As when G is trivial there is a cohomological form of $\operatorname{top}_G\text{-}\operatorname{ind}(P)$. This is particularly easy to deduce when G acts trivially on X. In that case, there is an isomorphism $K(X) \otimes R(G) \cong K_G(X)$ induced by tensoring bundles over X on which G acts trivially with product G-bundles $X \times V_i$ where V_i is an irreducible G-module (see [43]). Then we have

$$ch_G : K_G(X) \to H^*(X;\mathbb{C}) \otimes R(G) \text{ given by } ch \otimes \operatorname{Id} \text{ on } K(X) \otimes R(G) \cong K_G(X),$$

and we also have (using compact supports) $ch_G : K_G(T^*X) \to H^*(T^*X;\mathbb{C}) \otimes R(G)$.

Theorem 6.5 *For an elliptic differential G-operator $P : C^\infty(E) \to C^\infty(F)$, arising from a G-action which is trivial on X, the topological index of P is given by*

$$\operatorname{ind}_G P = \operatorname{top}_G\text{-}\operatorname{ind}(P) = (-1)^n \left\{ch_G([\sigma(P)])\mathbf{Td}(TX \otimes \mathbb{C})\right\}[TX],$$

where the symbol class $[\sigma(P)]$ is regarded as in $K_G(T^*X)$. Moreover, for $g \in G$,

$$\operatorname{ind}_g P = (-1)^n \left\{\operatorname{tr}\left(ch_G([\sigma(P)])g\right)\mathbf{Td}(TX \otimes \mathbb{C})\right\}[TX],$$

where tr *is the trace of $ch_G([\sigma(P)])(g)$ in the $R(G)$ factor of $K(X) \otimes R(G) \cong K_G(X)$, which results in an element of $K(X)$.*

Now suppose that the action of G on X is not trivial. For each $g \in G$, let $X^g := \{x \in X : g \cdot x = x\}$ denote the set of fixed points of g. Since there is a metric on X such that G acts by isometries, it follows that X^g is a union of finitely many compact connected submanifolds of X, say $X_1, \ldots, X_{k_g}$, of possibly different dimensions $d_1, \ldots, d_{k_g}$. For $k = 1, \ldots, k_g$, let $i_k : X_k \to X$ be the inclusion and let $N_k \to X_k$ be the normal bundle of X_k in X. We have $i_{k*} : TX_k \to TX$ and the normal bundle of TX_k in TX is $TN_k \to TX_k$, which recall has a complex structure. We wish to compute $\operatorname{ind}_g(P)$ as in Definition 6.2, and for this we may assume that G is the cyclic group generated by g. Note that $\pi_k : TN_k \to TX_k$ is a possibly nontrivial G-bundle. Recall that we have Thom element $\lambda_{TN_k} \in K_G(TN_k)$ which provides the Thom isomorphism $\varphi_k : K_G(TX_k) \to K_G(TN_k)$ via $\varphi_k(a) = \pi_k^*(a)\lambda_{TN_k}$. Then the extension homomorphism $h_k : K_G(TN_k) \to K_G(TX)$ yields $(i_k)_! := h_k \circ \varphi_k : K_G(TX_k) \to K_G(TX)$.

Theorem 6.6 (Atiyah-Segal-Singer Fixed- Point Formula) *In the above notation, we have*

$$\operatorname{ind}_g(P) = \sum\nolimits_{k=1}^{k_g} (-1)^{d_k} \left(\frac{ch_g((i_{k*})^*[\sigma(P)])}{ch_g(\lambda_{TN_k})}\mathbf{Td}(TX_k \otimes \mathbb{C})\right)[TX_k],$$

where the quotient has meaning in the context of the localization of the ring $R(G)$ at g since the trace of g in the representation defining λ_{TN_k} is nonzero.

For a more thorough discussion of the proof, see [10], [34, p.259f] and [44, p.120f]. The references [4], [5], [34, p.259f] and [44, p.120f] also contain the major instances of Theorem 6.6 obtained by using various standard elliptic operators P.

Elliptic Operator P	**Corresponding G-Theorem**
$d + \delta : \Omega^{even}(X) \to \Omega^{odd}(X)$	Lefschetz Fixed-Point Theorem
$d + \delta : \Omega^{+}(X) \to \Omega^{-}(X)$	G-Signature Theorem
$\partial + \overline{\partial} : \Omega^{0,even}(X) \to \Omega^{0,odd}(X)$	Holomorphic Lefschetz Theorem
$\mathcal{D}^{\pm} : C^\infty(\Sigma^{\pm}(X)) \to C^\infty(\Sigma^{\mp}(X))$	G-Spin Theorem

6.2 The Atiyah–Patodi–Singer Index Theorem

Let X be a compact, oriented Riemannian manifold with boundary $Y = \partial X$ with $\dim X = n = 2m$ even. Let $\mathcal{D} : C^\infty(X, S) \to C^\infty(X, S)$ be a compatible operator of Dirac type where $S \to X$ is a bundle of Clifford modules. Relative to the splitting $S = S^+ \oplus S^-$ into chiral halves, we have the operators $\mathcal{D}^+ : C^\infty(X, S^+) \to C^\infty(X, S^-)$ and $\mathcal{D}^- : C^\infty(X, S^-) \to C^\infty(X, S^+)$ which are formal adjoints on sections with support in $X \setminus Y$. We assume that all structures (e.g., Riemannian metric, Clifford module, connection) are products on some collared neighborhood $N \cong [-1,1] \times Y$ of Y. Then $\mathcal{D}^+|_N := \mathcal{D}^+ : C^\infty(N, S^+|N) \to C^\infty(N, S^-|N)$ has the form

$$\mathcal{D}^+|_N = \sigma(\partial_u + \mathcal{B}).$$

Here $u \in [-1,1]$ is the normal coordinate (i.e., $N = \{(u,y) \mid y \in Y,\ u \in [-1,1]\}$) with $\partial_u = \frac{\partial}{\partial u}$ the inward normal), $\sigma = \mathbf{c}(du)$ is the (unitary) Clifford multiplication by du with $\sigma(S^+|N) = S^-|N$, and

$$\mathcal{B} : C^\infty\left(Y, S^+|_Y\right) \to C^\infty\left(Y, S^+|_Y\right)$$

denotes the canonically associated (elliptic, self-adjoint) Dirac operator over Y, called the *tangential operator*. Note that due to the product structure, σ and $\mathcal{B}$ do not depend on u. Let $P_{\geq}(\mathcal{B})$ denote the spectral (Atiyah–Patodi–Singer) projection onto the subspace $L_+(\mathcal{B})$ of $L^2(Y, S^+|_{\partial X})$ spanned by the eigensections corresponding to the nonnegative eigenvalues of $\mathcal{B}$. Let

$$C^\infty\left(X, S^+; P_{\geq}\right) := \left\{\psi \in C^\infty\left(X, S^+\right) \mid P_{\geq}(\mathcal{B})\left(\psi|_Y\right) = 0\right\}, \text{ and}$$
$$\mathcal{D}^+_{P_{\geq}} := \mathcal{D}^+|_{C^\infty\left(X,S^+;P_{\geq}\right)} : C^\infty\left(X, S^+; P_{\geq}\right) \to C^\infty\left(X, S^-\right).$$

The eta function for $\mathcal{B}$ is defined by

$$\eta_{\mathcal{B}}(s) := \sum\nolimits_{\lambda \in \operatorname{spec}\mathcal{B} - \{0\}} (\operatorname{sign}\lambda)\, m_\lambda\, |\lambda|^{-s},$$

for $\Re(s)$ sufficiently large, where m_λ is the multiplicity of λ. Implicit in the following result (originating in [8]) is that $\eta_{\mathcal{B}}$ extends to a meromorphic function on all $\mathbb{C}$, which is holomorphic at $s = 0$ so that $\eta_{\mathcal{B}}(0)$ is finite.

Theorem 6.7 (Atiyah-Patodi-Singer Index Formula) *The above operator $\mathcal{D}^+_{P_{\geq}}$ has finite index given by*

$$\operatorname{index}\mathcal{D}^+_{P_{\geq}} = \int_X \left(\mathbf{ch}(S,\varepsilon) \wedge \widetilde{\mathbf{A}}(X,\theta)\right) - \frac{m_0 + \eta_{\mathcal{B}}(0)}{2}.$$

where $m_0 = \dim(\operatorname{Ker}\mathcal{B})$, $\mathbf{ch}(S,\varepsilon) \in \Omega^*(X,\mathbb{R})$ is the total Chern character form of the complex vector bundle S with compatible, unitary connection ε, and $\widetilde{\mathbf{A}}(X,\theta) \in \Omega^*(X,\mathbb{R})$ is closely related to the total $\widehat{\mathbf{A}}(X,\theta)$ form relative to the Levi-Civita connection θ, namely $\widetilde{\mathbf{A}}(X,\theta)_{4k} = 2^{2k-m}\widehat{\mathbf{A}}(X,\theta)_{4k}$.

Bibliography

[1] M. F. Atiyah: *Collected Works* (5 vols., Oxford: Oxford University Press, 1988)

[2] M. F. Atiyah: *K-Theory* (New York: Benjamin, 1967)

[3] M. F. Atiyah and R. Bott: The index problem for manifolds with boundary *In: Differential Analysis* (papers presented at the Bombay Colloquium, Oxford: Oxford University Press, 1964) 175–186; in [1], vol. 3, 25–37

[4] M. F. Atiyah and R. Bott: The Lefschetz fixed point theorem for elliptic complexes I *Ann. of Math.* **86** (1967) 374–407; in [1], vol. 3, 91–125

[5] M. F. Atiyah and R. Bott: The Lefschetz fixed point theorem for elliptic complexes II *Ann. of Math.* **88** (1968) 451–491; in [1], vol. 3, 127–169

[6] M. F. Atiyah, R. Bott and V. K. Patodi: On the heat equation and the index theorem *Invent. Math.* **19** (Russ. transl., Moscow, 1973), (1973) 279–330; *Errata* **28** (1975) 277–280; in [1], vol. 4, 11–63

[7] M. F. Atiyah, N. J. Hitchen and I. M. Singer: Self-duality in four dimensional Riemannian geometry *Proc. Roy. Soc. London* **362** (1978) 425–461; in [1], vol. 5, 27–65

[8] M. F. Atiyah, V. K. Patodi and I. M. Singer: Spectral asymmetry and Riemannian geometry *Bull. London Math. Soc.* **5** (1973) 229–234; in [1], vol. 4, 71–79

[9] M. F. Atiyah, V. K. Patodi and I. M. Singer: Spectral asymmetry and Riemannian geometry I *Math. Proc. Camb. Phil. Soc.* **77** (1975) 43–69; in [1], vol. 4, 81–107

[10] M. F. Atiyah and G. B. Segal: The index of elliptic operators II *Ann. of Math.* **87** (1968) 531–545; in [1], vol. 3, 223–237

[11] M. F. Atiyah and I. M. Singer: The index of elliptic operators on compact manifolds *Bull. Amer. Math. Soc.* **69** (1963) 422–433; in [1], vol. 3, 13–24

[12] M. F. Atiyah and I. M. Singer: The index of elliptic operators I *Ann. of Math.* **87** (1968) 484–530; in [1], vol. 3, 171–219

[13] M. F. Atiyah and I.M. Singer: The index of elliptic operators IV *Ann. of Math.* **93** (1971) 119–138; in [1], vol. 3, 301–322

[14] F. W. Atkinson: The normal solubility of linear equations in normed spaces *Mat. Sb.* **28(70)** (1951) 3–14

[15] N. Berline, E. Getzler and M. Vergne: *Heat Kernels and Dirac Operators* (Berlin-Heidelberg-New York: Springer-Verlag, 1992)

[16] D. Bleecker: The supertrace of the steady asymptotic of the spinorial heat kernel *J. Math. Phys.* **33** (1992) 2053–2070

[17] D. Bleecker: The behavior of the analytic index under nontrivial embedding *In: Analysis, geometry and topology of elliptic operators* (Hackensack, NJ: World Sci. Publ., 2006) 41–62

[18] B. Booss–Bavnbek and D. Bleecker: *Topology and Analysis* (Berlin-Heidelberg-New York: Springer-Verlag, 1985)

[19] D. Bleecker and B. Booss-Bavnbek: Spectral invariants of operators of Dirac type on partitioned manifolds *In: Aspects of boundary problems in analysis and geometry* (Oper. Theory Adv. Appl. 151, Basel: Birkhäuser, 2004) 1–130

[20] B. Booss–Bavnbek and K. P. Wojciechowski: *Elliptic Boundary Problems for Dirac Operators* (Boston: Birkhäuser, 1993)

[21] J. Dieudonné: Sur les homomorphismes d'espace normés *Bull. Sci. Math.* **67(2)** (1943) 72–84

[22] E. I. Fredholm: Sur une classe d'equations fonctionnelles *Acta Math.* **27** (1903) 365–390

[23] I. M. Gelfand: On elliptic equations *Russ. Math. Surveys* **15(3)** (1960) 113–123; *translation from Uspehi Mat. Nauk* **15(3)** (1960) 121–132

[24] E. Getzler: Pseudodifferential operators on supermanifolds and the index theorem *Comm. Math. Phys.* **92** (1983) 163–178

[25] E. Getzler: A short proof of the Atiyah-Singer index theorem *Topology* **25** (1986) 111–117

[26] P. B. Gilkey: Curvature and eigenvalues of the Laplace operator for elliptic complexes *Advances in Math.* **10** (1973) 344–382

[27] P. B. Gilkey: *Invariance Theory, the Index Theorem and the Heat Equation* (Boca Ratan-Ann Arbor-London-Tokyo: CRC Press, 1995)

[28] I. Z. Gohberg and M. G. Krein: The basic propositions on defect numbers, root numbers, and indices of linear operators *translation from Upehi Mat. Nauk* **12(2)** (1957) 44–118; *Amer. Math. Soc. Transl.* **13(2)** (1960) 185–264

[29] P. Griffiths and J. Harris: *Principles of Algebraic Geometry* (New York: Wiley, 1978)

[30] E. Guentner: K-homology and the Index Theorem *In: Index Theory and Operator Algebras* (Contemp. Math. 148, 1993) 47–66

[31] K. Jänich: Vektorraumbündel und der Raum der Fredholm-Operatoren *Math. Ann.* **161** (1965) 129–142

[32] M. Karoubi: *K-theory – An Introduction* (Berlin-Heidelberg-New York: Springer-Verlag, 1978)

[33] N. H. Kuiper: The homotopy type of the unitary group of Hilbert space *Topology* **3** (1965) 19–30

[34] H. B. Lawson and M. L. Michelsohn: *Spin Geometry* (Princeton: Princeton Univ. Press, 1989)

[35] H. P. Jr. McKean and I. M. Singer: Curvature and the eigenvalues of the Laplacian *J. Differential Geometry* **1** (1967) 43–69

[36] R. B. Melrose: *The Atiyah-Patodi-Singer Index Theorem* (Boston: A. K. Peters, 1993)

[37] J. Milnor J. Stasheff: *Characteristic Classes* (Ann. Math. Studies 76, Princeton: Princeton Univ. Press, 1974)

[38] F. Noether: Ober eine Klasse singulärer Integralgleichungen *Math. Ann.* **82** (1921) 42–63

[39] R. S. Palais: Imbedding of compact, differentiable transformation groups in orthogonal representations *J. Math. Mech.* **6** (1957) 673–678

[40] R. S. Palais: *Foundations of Global Non-Linear Functional Analysis* (New York: Benjamin, 1968)

[41] R. S. Palais, (ed.): *Seminar on the Atiyah-Singer index theorem* (Ann. of Math. Studies 57, Princeton: Princeton Univ. Press, 1965)

[42] V. K. Patodi: Curvature and the eigenforms of the Laplace operator *J. Diff. Geom.* **5** (1971) 233–249

[43] G. B. Segal: Equivariant K-theory *Inst. Hautes Études Sci. Publ. Math.* **34** (1968) 129–151

[44] P. Shanahan: *The Atiyah-Singer Index Theorem: An Introduction* (Lecture Notes in Mathematics 638, Berlin-Heidelberg-New York: Springer-Verlag, 1978)

[45] I. N. Vekua: *Generalized Analytic Functions* (transl. from Russian, Oxford: Pergamon Press, 1962)

[46] T. Voronov: Quantization of forms on the cotangent bundle *Comm. Math. Phys.* **205** (1999) 315–336

[47] H. Weyl: *The Classical Groups* (Princeton Univ. Press, Princeton, 1946)

[48] B. Yood: Properties of linear transformations preserved under additions of completely continuous transformations *Duke Math J.* **18** (1951) 599–612

DAVID BLEECKER

Department of Mathematics, University of Hawaii, Honolulu, HI 96822, U.S.A.

E-mail: `bleecker@math.hawaii.edu`

This page intentionally left blank

Handbook of Global Analysis
Demeter Krupka and David Saunders

Partial differential equations on closed and open manifolds

Jürgen Eichhorn

Contents

1 Introduction

Global analysis is essentially the global theory of differential equations on manifolds. It naturally splits into a theory of ordinary (ODEs) and one of partial differential equations (PDEs). The theory of ordinary differential equations can be embedded into the theory of dynamical systems. In this contribution, we are concerned with PDEs on manifolds.

As very well known, there are many important PDEs of mathematical physics and geometry which have been intensively studied by global analysts. We will discuss a selection of them and exhibit essential approaches of treatment and solving. Clearly, the methods depend on the equations under consideration, on the choice of functional spaces, the compactness or non-compactness of the underlying manifold, its geometry and topology. On compact manifolds, one often uses (functional analytic) compactness arguments of Arzela-

Ascoli type. On open manifolds these are not available, and versions of the continuity method are a good approach. The intension of our contribution was to present as much as possible approaches for solving, not to discuss as much as possible PDEs. The latter is absolutely senseless. Too many important and interesting PDEs exist. Only for the reason of space, we did not discuss Einstein's equations, the Yamabe problem, Seiberg-Witten theory, Melrose's B-calculus (cf. [87]), B.-W. Schulze's approach to manifolds with singularities (cf. [106], [107]) etc. Nevertheless, the theory of solving for these equations is in principal contained in our contribution.

The contribution is organized as follows. Section 2 and 3 are devoted to linear and non-linear Sobolev structures and give hence the functional analytic "frame" for many approaches. The spectral properties of linearized differential operators play an essential role in the theory of solutions. Therefore we sketch in sections 4 and 5 the spectral theory of self-adjoint operators on manifolds. Section 6 is devoted to the heat equation and the heat kernel and section 7 to the wave equation, Huygens' principle and the Hamiltonian approach for the wave equation. Here we already indicate an important approach for solving PDEs, namely to reformulate it as an ODE on an infinite-dimensional manifold. In section 8 we outline index theory on open manifolds. The continuity method is one of the key approaches for non-linear PDEs. We present a basic version of it in section 9 and apply this in the 10th section on Teichmüller theory. Concerning harmonic maps, we present the heat flow method as established by Eells/Sampson. For more on harmonic maps, we refer to John Wood's contribution. In section 12, we discuss some non-linear field equations and sketch the method of Agricola/Friedrich/Ivanow/Kim. They do not solve these equations by purely analytical methods – e.g. by deformations – but discuss these equations in purely geometrical terms and search step by step for geometries which represent a solution or single out others, respectively. Section 13 is devoted to gauge theory and section 14 contains a reformulation of equations of fluid dynamics as ODE on an infinite-dimensional manifold. Finally, section 15 is devoted to the spectacular Ricci flow.

For reasons of space, we had to make a rigorous choice. Therefore we apologize to whom we could not mention.

2 Sobolev spaces

Sobolev spaces or – more general – non-linear Sobolev structures provide the frame for the solution of PDEs. There are still some other choices of functional spaces possible, but in this contribution we essentially restrict to the Sobolev case. The theory of Sobolev spaces for Euclidean bounded C^∞-domains and that for closed manifolds are nearly parallel. In distinction to the compact case, this theory on open manifolds has its own features. We consider Sobolev spaces on open complete manifolds with C^k-bounded curvature tensor and – sometimes – with positive injectivity radius. All presented theorems also hold for compact manifolds since they have bounded geometry of infinite order.

Let (M^n, g) be a smooth Riemannian manifold, $T = TM$ its tangent bundle and $T^* = T^*M$ its cotangent bundle, $T^u_v = T^{\otimes u} \otimes T^{*\otimes v}$ the bundle of u-contravariant and v-covariant tensors and $\Lambda^q = \Lambda^q(M) = \Lambda^q T^*M$ the qth exterior power. We denote by $\Gamma(\cdot) = C^\infty(\cdot)$ the vector space of smooth sections and by $\Gamma_c(\cdot) = C^\infty_c(\cdot)$ the subspace of smooth sections with compact support. Let $\Omega^q = \Omega^q(M) = C^\infty(\Lambda^q)$ denote the smooth differential forms and $\Omega^* = \sum\limits_{q\geq 0} \Omega^q$. By $\nabla = \nabla^g$ we denote the Levi-Civita connection,

i. e. the unique metric and torsion-free connection with respect to g. $R = R^g$ denotes the curvature tensor, $R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z$ and $K(X,Y) = g(R(X,Y)Y,X)/(|X|^2 \cdot |Y|^2 - g(X,Y)^2)$ the sectional curvature for X and Y linearly independent.

Assume that (M^n, g) is oriented. If $\varphi, \omega \in C_c^\infty(\Lambda^q) = \Omega_c^q$ then there is a pre-Hilbert scalar product defined by $\langle \varphi, w \rangle = \int\limits_M \varphi^{i_1 \cdots i_q} \omega_{i_1 \ldots i_q} \,\mathrm{dvol}_x(g) = \int\limits_M g(\varphi,\omega) \,\mathrm{dvol}_x(g) = \int\limits_M (\varphi,\omega)_x \,\mathrm{dvol}_x(g)$, where $|\varphi|^2_{L_2} \equiv |\varphi|_2^2 = \langle \varphi, \varphi \rangle = \int\limits_M |\varphi|^2_{g,x} \,\mathrm{dvol}_x(g)$. We denote $\Omega^{q,2} \equiv L_2(\Lambda^q, g) = \overline{\Omega_c^q}^{|\,|_2}$.

If $\omega \in \Omega^q$, $\omega|_U = \sum\limits_{i_1 < \cdots i_q} \omega_{i_1 \cdots i_q} du^{i_1} \wedge \cdots \wedge du^{i_q}$, then $d\omega = d_q \omega$ is defined by $d\omega|_U = \sum\limits_{i_1 < \cdots < i_q} (d\omega_{i_1 \ldots i_q}) \wedge du^{i_1} \wedge \cdots \wedge du^{i_q}$. It is known that $d^2 = d_{q+1} d_q = 0$, whence we obtain the de Rham complex

$$0 \longrightarrow \Omega^0 \longrightarrow \Omega^1 \xrightarrow{d_1} \cdots \longrightarrow \Omega^q \xrightarrow{d_q} \Omega^{q+1} \xrightarrow{d_{q+1}} \cdots \longrightarrow \Omega^n \longrightarrow 0.$$

The map $\delta = \delta_{q+1} : \Omega^{q+1} \longrightarrow \Omega^q$ is defined as the adjoint to d_q w. r. t. $\langle\, , \rangle$, i. e. $\langle d\varphi, \omega \rangle = \langle \varphi, \delta\omega \rangle$ for $\varphi \in \Omega_c^q$, $\omega \in \Omega_c^{q+1}$. δ_{q+1} can be expressed as $\delta_{q+1} = (-1)^{n(q+1)+n+1} * d_{n-(q+1)} *$, where $*$ is the Hodge $*$-operator $*(e^{i_1} \wedge \cdots \wedge e^{i_{q+1}}) = \mathrm{sign}\left({}^{\;1 \;\; 2 \;\, \ldots \;\, n}_{i_1 i_2 \ldots i_{q+1} j_1 \ldots j_{q-1}} \right) e^{j_1} \wedge \cdots \wedge e^{j_{q-1}}$, $e^1, \ldots e^n$ being an orthonormal cobasis in $T_x^* M$. $\Delta = \Delta_q = d_{q-1}\delta_q + \delta_{q+1} d_q : \Omega^q \longrightarrow \Omega^q$ is the Laplace operator. It is symmetric on $\Omega_c^q \subset \Omega^{q,2} \equiv L_2(\Lambda^q, g)$ and ≥ 0.

If $(E, h, \nabla) \longrightarrow (M^n, g)$ is a Riemannian or Hermitean fibre bundle with fibre metric $h = h(\cdot,\cdot)$ and metric connection $\nabla = \nabla^E = \nabla^h$, then we can similarly define its curvature $R = R^E \in \Omega^2(\mathrm{End}\, E)$, $R^E(X,Y)\varphi = \nabla_X \nabla_Y \varphi - \nabla_Y \nabla_X \varphi - \nabla_{[X,Y]} \varphi$, where for a vector bundle $F \longrightarrow M$, $\Omega^q(F) = C^\infty(\Lambda^q \otimes F)$ is the space of $q-$forms with values in F. If E is flat, i. e. $R^E = 0$, then we obtain the de Rham complex with values in E_X.

$$0 \longrightarrow \Omega^0(E) \longrightarrow \Omega^1(E) \longrightarrow \cdots \longrightarrow \Omega^q(E) \xrightarrow{d_q} \Omega^{q+1}(E) \longrightarrow \cdots \longrightarrow \Omega^n(E) \longrightarrow 0$$

with $d_{q+1}^E d_q^E = d^2 = 0$, where $d_q^E(\omega \otimes \varphi) = d_q \omega \otimes \varphi + (-1)^q \omega \wedge \nabla^E \varphi$. g, h and the Riemannian measure define a global scalar product by $\langle \omega \otimes \varphi, \omega' \otimes \varphi' \rangle = \int\limits_M (\omega, \omega')_{g,x} \cdot (\varphi, \varphi')_{h,x} \,\mathrm{dvol}_x(g)$, $\omega \otimes \varphi, \omega' \otimes \varphi' \in \Omega_c^q(E)$, that extends linearly. Thus there is an adjoint operator $\delta = \delta_{q+1} = \delta_{q+1}^E : \Omega^{q+1}(E) \longrightarrow \Omega^q(E)$. $\Delta_q = \Delta_q^E = d_{q-1}^E \delta_q^E + \delta_{q+1}^E d_q^E : \Omega^q(E) \longrightarrow \Omega^q(E)$ is the Laplace operator acting on q-forms with values in E. Δ_q and Δ_q^E satisfy Weitzenboeck formulas. In local coordinates these are

$$(\Delta_1 \omega)_i = -\nabla^r \nabla_r \omega_i + R_{ir} \omega^r \text{ for } q = 1, \tag{2.1}$$

$$(\Delta_q \omega)_{i_1 \ldots i_q} = -\nabla^r \nabla_r \omega_{i_1 \ldots i_q} + \sum_{a=1}^{q} R^r_{i_a} \omega_{i_1 \ldots r \ldots i_q} + \sum_{a<b} R^{rs}_{i_a i_b} \omega_{i_1 \ldots r \ldots s \ldots i_q} \text{ for } n \geq q \geq 2. \tag{2.2}$$

Here $R_{ir} = \mathrm{Ric}\left(\frac{\partial}{\partial u^i}, \frac{\partial}{\partial u^r}\right)$ etc.

For Δ_q^E there is actually a more general formula which includes (2.1) and (2.2) as special cases. Let R denote the curvature of the tensor product connection on $\Lambda^q T^*M \otimes E$. For $\Phi \in \Omega^q(E)$ and an orthonormal basis $e_1, \dots, e_n \in T_xM$, $v_1, \dots, v_q \in T_xM$, define

$$\mathcal{R}(\Phi)v_1, \dots v_q \equiv R(\Phi)(v_1, \dots, v_q) = \sum_{k=1}^{q} \sum_{j=1}^{n} R(e_j, v_k)\Phi_{v_1,\dots,v_{k-1},e_j,v_{k+1},\dots,v_q} \quad (2.3)$$

Then

$$\Delta_q^E = \nabla^*\nabla + \mathcal{R}.$$

The proof uses the explicit expressions for d^E and δ^E,

$$(d^E\Phi)_{v_0,\dots,v_q} = \sum_{k=0}^{q} (-1)^k (\nabla_{v_k}\Phi)(v_0, \dots, \hat{v}_k, \dots, v_q), \quad (2.4)$$

$$(\delta^E\Phi)_{v_2,\dots,v_q} = -\sum_{j=1}^{n} (\nabla_{e_j}\Phi)(e_j, v_2, \dots, v_q). \quad (2.5)$$

Inserting (2.4), (2.5) into $\Delta_q^E = d_{q-1}^E \delta_q^E + \delta_{q+1}^E d_q^E$ yields (2.3). An easy calculation exhibits (2.1), (2.2) as special cases of (2.3) for $E = M \times \mathbb{R}$.

We define now Clifford bundles and the generalized Dirac operator. For $x \in M$ let $\mathrm{CL}_x = \mathrm{CL}(T_xM, g_x)$ be the Clifford algebra at x. CL_x could be complexified, depending on the other bundles and structures under consideration. A Hermitean vector bundle $(E, h, \nabla^E) \longrightarrow (M^n, g)$ is called a bundle of Clifford modules if each fibre E_x is a Clifford module over (the complexified) algebra CL_x with skew-symmetric Clifford multiplication. The metric connection ∇^E is a Clifford connection if it satifies the Leibniz rule

$$\nabla_x^E(Y \cdot \varphi) = (\nabla_x^g Y) \cdot \varphi + Y \cdot \nabla_x^E \varphi, \quad X, Y \in T_xM, \quad \varphi \in X^\infty(E). \quad (2.6)$$

Then $(E, h, \nabla^E, \cdot) \longrightarrow (M^n, g)$ is called a Clifford bundle. The composition

$$D = \cdot \circ g^{-1} \circ \nabla : \Gamma(E) \xrightarrow{\nabla} \Gamma(T^*M \otimes E) \xrightarrow{g^{-1}} \Gamma(TM \otimes E) \xrightarrow{\cdot} \Gamma(E)$$

is called the generalized Dirac operator D. Set $D = D(E, h, \nabla^E, \cdot, g)$. If $e_1, \dots, e_n \in T_xM$ is an orthonormal basis of T_xM, then $D = \sum_{i=1}^{n} e_i \cdot \nabla_{e_i}$. D is a first order elliptic and symmetric operator on $C_c^\infty(E) = \Gamma_c(E)$.

Examples 2.1 **1)** The simplest standard example is $(E = \Lambda^*T^*M \otimes \mathbb{C}, g_{\Lambda^*}, \nabla^{g_{\Lambda^*}}) \longrightarrow (M^n, g)$ with Clifford multiplication

$$X \otimes \omega \in T_xM \otimes (\Lambda^*T^* \otimes \mathbb{C}) \longrightarrow X \cdot \omega = \omega_x \wedge \omega - i_x\omega, \quad (2.7)$$

where $\omega_x = g(\cdot, X)$. It is well known that in this case

$$\begin{aligned} D &= d + d^* = d + \delta, \\ D^2 &= (d+\delta)^2 = \Delta = \text{graded Laplace operator} = \Delta_0 \oplus \cdots \oplus \Delta_n. \end{aligned} \quad (2.8)$$

Any result concerning generalized Dirac operators D^2 is simultaneously a result for Laplace operators.

If we consider a twisted Clifford structure $E \otimes F$, $(F, h_F, \nabla^F) \longrightarrow (M^n, g)$, $\nabla(e \otimes f) = (\nabla^E e) \otimes f + e \otimes \nabla^F f$, $X \cdot_{E \otimes F} (e \otimes f) = (X \cdot_E e) \otimes f$, we get a twisted generalized Dirac operator. In (2.7) above, D^2 is then the graded Laplace operator acting on forms with values in F.

Let $\Delta^E = (\nabla^E)^* \nabla^E$ be the Bochner-Laplace operator,

$$(\Delta^E \varphi)(x) = -\sum_{i=1}^{n} (\nabla^E_{e_i} \nabla^E_{e_i} \varphi)(x),$$

where $e_1, \ldots, e_n$ is an orthonormal basis in $T_x M$ satisfying $\nabla e_i|_x = 0$. Then D^2 satisfies a Weitzenboeck formula

$$D^2 = \Delta^E + \mathcal{R} \tag{2.9}$$

where $\mathcal{R}\varphi = \frac{1}{2} \sum\limits_{j=1}^{n} e_i \cdot e_j \cdot R^E(e_i \cdot e_j)\varphi$.

2) If (M^n, g) admits a spin structure, and S is the associated (graded or not) spinor bundle with fibre metric h_S and spin connection ∇^S, then $(S, h_S, \nabla^S) \longrightarrow (M^n, g)$ is a Clifford bundle and its generalized Dirac operator is the classical Dirac operator. □

Next we will define bounded geometry. Given $x \in M$, the map $\exp_x : T_x M \longrightarrow M$, $\exp_x(X) = c(1)$, where $c(t)$, $0 \leq t \leq 1$, is called the geodesic solution of $\nabla_{\dot{c}} \dot{c} = 0$, $c(0) = x$, $\dot{c}(0) = X$, is the exponential map. If it is defined for all $x \in M$ and all $X \in T_x M$, (M^n, g) is called complete. According to the Hopf-Rinow theorem this is equivalent to the completeness of $(M^n, \mathrm{dist}_g(\cdot, \cdot))$ as metric space or to the fact that every bounded set is relatively compact. The number $\sup\{r \mid \exp_x : B_r(0) \subset T_x M \longrightarrow M \text{ is a diffeomorphism}\}$ is the injectivity radius of (M^n, g) at x, $r_{\mathrm{inj}}(x)$. $r_{\mathrm{inj}}(M^n, g) = \inf\limits_x r_{\mathrm{inj}}(x)$ is the injectivity radius of (M^n, g). We say (M^n, g) has bounded geometry up to order k if it satifies the conditions (I) and $(B_k(M, g))$,

$$r_{\mathrm{inj}}(M, g) > 0 \tag{I}$$

$$|\nabla^i R| \leq C_i, \quad \forall i = 0, \ldots, k. \tag{$B_k(M,g)$}$$

The condition $(B_\infty(M, g))$ means $|\nabla^i R| \leq C_i$, $i = 1, 2, \ldots$. Every closed Riemannian manifold satifies (I) and (B_∞). Examples of open manifolds satifying (I) and (B_∞) are homogeneous spaces or Riemannian coverings of closed manifolds. Greene has proven in [66] that every open manifold admits a metric g satisfying (I) and (B_∞), i. e. bounded geometry does not affect the topological type. We restrict in most of our considerations to bounded geometry. The reason for this is the fact that then Sobolev analysis is available, e. g. embedding theorems, module structure theorems and many invariance properties. If we give up (I) for instance, then these theorems do not apply. Parts of them still hold by using weighted Sobolev spaces, but this requires additional effort.

We list some important consequences of (I) and (B_k).

Proposition 2.2 **a)** *(I) implies completeness of (M^n, g).*

b) *If (M^n, g) satifies (I) and (B_k) and $\mathfrak{U} = \{(U_\alpha, \Phi_\alpha)\}_\alpha$ is a locally finite cover by normal charts, then there exist constants C_β, C'_β, C'_γ, multi-indexed by β, γ, such that*

$$|D^\beta g_{ij}| \leq C_\beta, \quad |D^\beta g^{ij}| \leq C'_\beta, \quad \text{for } |\beta| \leq k \tag{2.10}$$

and

$$|D^\gamma \Gamma_{ij}^m| \le C'_\gamma, \quad |\gamma| \le k-1, \tag{2.11}$$

all constants are independent of α.

c) *If* $(E, h, \nabla^E) \longrightarrow (M^n, g)$ *is a Riemannian vector bundle satifying* (I), $(B_k(M,g))$, $(B_k(E,\nabla))$, *then additionally to (2.10), (2.11) there holds for the connection coefficients* $\Gamma_{i\lambda}^\mu$ *defined by* $\nabla_{\frac{\partial}{\partial u^i}} \varphi_\lambda = \Gamma_{i\lambda}^\mu \varphi_\mu$, $\{\varphi_\mu\}_\mu$ *a local orthonormal frame obtained by radial parallel translation,*

$$|D^\beta \Gamma_{i\lambda}^\mu| \le D_\beta, \quad |\beta| \le k-1. \tag{2.12}$$

Proof Under the assumption of (I) any Cauchy sequence $(x_\nu)_\nu$ in M can be considered, up to quasi-isometry, as contained in a small closed Euclidean ball, omitting only a finite number of the x_ν's. This proves a). b) and c) are the content of [46]. □

Proposition 2.3 *Assume* (M^n, g) *satisfies* (I) *and* (B_0). *There exists* $\varepsilon_0 > 0$ *such that for any* $\varepsilon \in]0, \varepsilon_0[$ *there is a countable cover of* M *by geodesic balls* $B_\varepsilon(x_i)$, $\bigcup_i B_\varepsilon(x_i) = M$, *such that the cover of* M *by the balls* $B_{2\varepsilon}(x_i)$ *with double radius and same centers is still uniformly locally finite.*

We refer to [82] for the proof. □

Proposition 2.3 implies the existence of an associated uniform partition of unity.

Proposition 2.4 *Assume* (M^n, g) *open with* (I) *and* (B_k) *and* $r \in]0, r_{\text{inj}}[$. *For every* $0 < \varepsilon < \frac{r}{2}$ *there exists a partition of unity* $1 = \sum_{i=1}^\infty \psi_i$ *on* M *such that*

1) $\psi_i \ge 0$, $\psi_i \in C_c^\infty(M)$, $\operatorname{supp} \psi_i \subset B_{2\varepsilon}(x_i)$, *where the sequence* $\{x_i\}_i$ *comes from proposition 2.3,*

2) $|D_u^\beta \psi_i(u^1, \dots, u^n)| \le C_\beta$, $|\beta| \le k+2$, *where* $(u^1, \dots, u^n)$ *are normal coordinates in* $B_{2\varepsilon}(x_i)$.

We refer to [82], [110] for the proof. □

Bounded geometry also ensures a good approximation $\tilde{d}$ for the Riemannian distance $\operatorname{dist}(\cdot,\cdot) = d$, $d(x,y) = \inf\{\text{length}\,(c) \mid c \text{ joins } x \text{ and } y\}$. d itself is only Lipschitz.

Proposition 2.5 *Take* (M^n, g) *with* (I) *and* (B_k). *There exists a function* $\tilde{d} : M \times M \longrightarrow [0, \infty[$ *satisfying the following properties.*

1) *There exists* $\varrho > 0$ *s. t.*

$$|\tilde{d}(x,y) - d(x,y)| < \varrho$$

for every $x, y \in M$.

2) *For every multi-index* β, $1 \le |\beta| \le k+1$, *there exists* $C_\beta > 0$, *s. t. in normal coordinates* $(y^1, \dots, y^n)$

$$|D_y^\beta \tilde{d}(x,y)| \le C_\beta.$$

Moreover, for every $\varepsilon > 0$ *there exists a function* $\tilde{d} : M \times M \longrightarrow [0, \infty[$ *satisfying 1) with* $\varrho < \varepsilon$.

We refer to [82] for the proof. □

Let $(E, h, \nabla^h) \longrightarrow (M^n, g)$ be a Riemannian vector bundle. Then the Levi-Civita connection ∇^g and ∇^h define metric connections ∇ in all tensor bundles $T^u_v \otimes E$. Denote smooth sections as above by $C^\infty(T^u_v \otimes E)$, by $C^\infty_c(T^u_v \otimes E)$ those with compact support. In the sequel we shall write E instead of $T^u_v \otimes E$, keeping in mind that E can be an arbitrary vector bundle. Now we define for $p \in \mathbb{R}$, $1 \leq p < \infty$ and r a non-negative integer

$$|\varphi|_{p,r} := \left(\int \sum_{i=0}^{r} |\nabla^i \varphi|^p_x \, \mathrm{dvol}_x(g) \right)^{1/p},$$

$$\begin{aligned}
&\Omega^{0,p}_r(E) \equiv \Omega^p_r(E) = \{\varphi \in C^\infty(E) \mid |\varphi|_{p,r} < \infty\},\\
&\bar{\Omega}^{0,p,r}(E) \equiv \bar{\Omega}^{p,r}(E) = \text{completion of } \Omega^p_r(E) \text{ with respect to } |\cdot|_{p,r},\\
&\mathring{\Omega}^{0,p,r}(E) \equiv \mathring{\Omega}^{p,r}(E) = \text{completion of } C^\infty_c(E) \text{ with respect to } |\cdot|_{p,r} \text{ and}\\
&\Omega^{0,p,r}(E) \equiv \Omega^{p,r}(E) = \{\varphi \mid \varphi \text{ measurable distributional section with } |\varphi|_{p,r} < \infty\}.
\end{aligned}$$

Here we use the standard identification of sections of a vector bundle E with E-valued zero-forms. $\Omega^{q,p,r}(E)$ stands for a Sobolev space of q-forms with values in E.

For $p = 2$, we often use the notations $|\,|_{2,0} = |\,|_{L_2} = |\,|_2 = \|\,\|$. Furthermore, we define

$$\begin{aligned}
{}^{b,m}|\varphi| &:= \sum_{i=0}^{m} |\nabla^i \varphi|_x,\\
{}^{b,m}\Omega(E) &= \{\varphi \mid \varphi \; C^m\text{-section and } {}^{b,m}|\varphi| < \infty\}, \text{ and}\\
{}^{b,m}\mathring{\Omega}(E) &= \text{completion of } C^\infty_c(E) \text{ with respect to } {}^{b,m}|\cdot|.
\end{aligned}$$

${}^{b,m}\Omega(E)$ equals the completion of

$${}^{b}_{m}\Omega(E) = \{\varphi \in C^\infty(E) \mid {}^{b,m}|\varphi| < \infty\}$$

with respect to ${}^{b,m}|\cdot|$.

Denote by ${}^{b,\infty}\Omega(E)$ the locally convex space of smooth sections φ such that $\nabla^s \varphi$ is bounded for $s = 0, 1, 2, \ldots$.

Proposition 2.6 *The spaces $\mathring{\Omega}^{p,r}(E)$, $\bar{\Omega}^{p,r}(E)$, $\Omega^{p,r}(E)$, ${}^{b,m}\mathring{\Omega}(E)$, ${}^{b,m}\Omega(E)$ are Banach spaces and there are inclusions*

$$\begin{aligned}
&\mathring{\Omega}^{p,r}(E) \subseteq \bar{\Omega}^{p,r}(E) \subseteq \Omega^{p,r}(E),\\
&\quad {}^{b,m}\mathring{\Omega}(E) \subseteq {}^{b,m}\Omega(E).
\end{aligned}$$

If $p = 2$, then $\mathring{\Omega}^{2,r}(E)$, $\bar{\Omega}^{2,r}(E)$, $\Omega^{2,r}(E)$ are Hilbert spaces. □

$\mathring{\Omega}^{p,r}(E)$, $\bar{\Omega}^{p,r}(E)$, $\Omega^{p,r}(E)$ are different from one another in general.

Proposition 2.7 *If (M^n, g) satisfies (I) and (B_k), then*

$$\mathring{\Omega}^{p,r}(E) = \bar{\Omega}^{p,r}(E) = \Omega^{p,r}(E), \quad 0 \leq r \leq k+2.$$

We refer to [45] for the proof. □

Embedding theorems are of great importance in non-linear global analysis and even more the module structure theorem which we present now.

Theorem 2.8 **a)** *Assume* $r - \frac{n}{p} \geq s - \frac{n}{q}$, $r \geq s$. *Let* $B \subset \mathbb{R}^n$ *be a Euclidean ball. Then*

$$\mathring{\Omega}^{p,r}(B \times \mathbb{R}^n) \hookrightarrow \mathring{\Omega}^{q,s}(B \times \mathbb{R}^n)$$

continuously.

b) *If* $r - \frac{n}{p} > s$, $s \in \mathbb{Z}_+$, *then*

$$\mathring{\Omega}^{p,r}(B \times \mathbb{R}^n) \hookrightarrow {}^{b,s}\mathring{\Omega}(B \times \mathbb{R}^n)$$

continuously. □

The global version of a) looks slightly different.

Theorem 2.9 *Let* $(E, h, \nabla^E) \longrightarrow (M^n, g)$ *be a Riemannian vector bundle satisfying* (I), $(B_k(M^n, g))$, $(B_k(E, \nabla))$, $k \geq 1$.

a) *Assume* $k \geq r$, $r - \frac{n}{p} \geq s - \frac{n}{q}$, $r \geq s$, $q \geq p$. *Then*

$$\Omega^{p,r}(E) \hookrightarrow \Omega^{q,s}(E) \tag{2.13}$$

continuously.

b) *If* $r - \frac{n}{p} > s$, *then*

$$\Omega^{p,r}(E) \hookrightarrow {}^{b,s}\Omega(E) \tag{2.14}$$

continuously.

We refer to [43] for a proof. □

Now we come to the module structure theorem.

Theorem 2.10 *Let* $(E_i, h_i, \nabla_i) \longrightarrow (M^n, g)$ *be vector bundles with* (I), $(B_k(M^n, g))$, $(B_k(E_i, \nabla_i))$, $i = 1, 2$. *Assume* $0 \leq r \leq r_1, r_2 \leq k$. *If* $r = 0$ *assume*

$$\left\{ \begin{array}{rcl} r - \frac{n}{p} & < & r_1 - \frac{n}{p_1} \\ r - \frac{n}{p} & < & r_2 - \frac{n}{p_2} \\ r - \frac{n}{p} & \leq & r_1 - \frac{n}{p_1} + r_2 - \frac{n}{p_2} \\ \frac{1}{p} & \leq & \frac{1}{p_1} + \frac{1}{p_2} \end{array} \right\} \text{or}$$

$$\left\{ \begin{array}{rcl} r - \frac{n}{p} & \leq & r_1 - \frac{n}{p_1} \\ 0 & < & r_2 - \frac{n}{p_2} \\ \frac{1}{p} & \leq & \frac{1}{p_1} \end{array} \right\} \text{or} \left\{ \begin{array}{rcl} 0 & < & r_1 - \frac{n}{p_1} \\ r - \frac{n}{p} & \leq & r_2 - \frac{n}{p_2} \\ \frac{1}{p} & \leq & \frac{1}{p_2} \end{array} \right\}.$$

If $r > 0$, *assume* $\frac{1}{p} \leq \frac{1}{p_1} + \frac{1}{p_2}$ *and*

$$\left\{ \begin{array}{rcl} r - \frac{n}{p} & < & r_1 - \frac{n}{p_1} \\ r - \frac{n}{p} & < & r_2 - \frac{n}{p_2} \\ r - \frac{n}{p} & \leq & r_1 - \frac{n}{p_1} + r_2 - \frac{n}{p_2} \end{array} \right\} \text{or} \left\{ \begin{array}{rcl} r - \frac{n}{p} & \leq & r_1 - \frac{n}{p_1} \\ r - \frac{n}{p} & \leq & r_2 - \frac{n}{p_2} \\ r - \frac{n}{p} & < & r_1 - \frac{n}{p_1} + r_2 - \frac{n}{p_2} \end{array} \right\}.$$

Then the tensor product of sections defines a continuous bilinear map

$$\Omega^{p_1,r_1}(E_1, \nabla_1) \times \Omega^{p_2,r_2}(E_2, \nabla_2) \longrightarrow \Omega^{p,r}(E_1 \otimes E_2, \nabla_1 \otimes \nabla_2). \tag{2.15}$$

We refer to [43] for the proof. □

Corollary 2.11 *Assume* $r = r_1 = r_2$, $p = p_1 = p_2$, $r > \frac{n}{p}$.

(a) *If* $E_1 = M \times \mathbb{R}$, $E_2 = E$, *then* $\Omega^{p,r}(E)$ *is a* $\Omega^{p,r}(M \times \mathbb{R})$*-module.*

(b) *If* $E_1 = M \times \mathbb{R} = E_2$, *then* $\Omega^{p,r}(M \times \mathbb{R})$ *is a commutative, associative Banach algebra.*

(c) *If* $E_1 = E = E_2$, *then the tensor product of sections defines a continuous map*

$$\Omega^{p,r}(E) \times \Omega^{p,r}(E) \longrightarrow \Omega^{p,r}(E \otimes E).$$

□

We extend our considerations to weighted Sobolev spaces, of which there are several different definitions. We consider here two definitions which coincide in the case of bounded geometry. Denote by $\varrho_y = d(y,x)$ the Riemannian distance between the points $y, x \in M$. Although $\varrho_y(x)$ is not C^1 we can achieve differentiability by approximation, see proposition 2.5.

Lemma 2.12 *Assume* (M^n, g) *with* (I) *and* (B_k). *Then there exist a* C^{k+2}*-function* $\tilde{\varrho}_y$ *and constants* $C, C_1, \ldots, C_{k+2}$, *such that*

$$|\varrho_y(x) - \tilde{\varrho}_y(x)| \le C \tag{2.16}$$

and

$$|\nabla^i \tilde{\varrho}_y| \le C_i, \quad 1 \le i \le k+2. \tag{2.17}$$

We refer to proposition 2.5 or [8], [82] for a proof. □

Now we define for $\varepsilon > 0$, $y \in M$

$$\Omega^{p,r}_{\varepsilon,y}(E) := \{\varphi \in \mathcal{D}'(E) \mid e^{\varepsilon\tilde{\varrho}_y}\varphi \in \Omega^{p,r}(E)\}.$$

and

$$\Omega'^{p,r}_{\varepsilon,y}(E) := \{\varphi \in \mathcal{D}'(E) \mid |e^{\varepsilon\varrho_y}\nabla^i\varphi|_{p,0} < \infty, i = 0, \ldots, r\}.$$

Both are Banach spaces endowed with the norms

$$|\varphi|_{p,r,\varepsilon,y} := |e^{\varepsilon\tilde{\varrho}_y}\varphi|_{p,r}$$

and

$$|\varphi|'_{p,r,\varepsilon,y} := \left(\int \sum_{i=0}^{r} |e^{\varepsilon\varrho_y}\nabla^i\varphi|^p_x \,\mathrm{dvol}_x(g)\right)^{1/p},$$

respectively. Both norms depend on the base point, but different base points give equivalent norms. The independence of the choice of $\tilde{\varrho}$ (up to equivalence) is expressed by

Proposition 2.13 *For* $r \le k$, $\Omega^{p,r}_{\varepsilon,y}(E)$ *and* $\Omega'^{p,r}_{\varepsilon,y}(E)$ *are equivalent Banach spaces.*

Proposition 2.14 *Assume* (E, h, ∇), (M^n, g) *as above,* $r, s \le k$. *If* $r - \frac{n}{p} \ge s - \frac{n}{q}$, $r \ge s$, $q \ge p$, $\varepsilon_1 \ge \varepsilon_2$, *then*

$$\Omega^{p,r}_{\varepsilon_1,y}(E) \hookrightarrow \Omega^{q,s}_{\varepsilon_2,y}(E)$$

continuously.

Theorem 2.15 *Assume* (E_i, h_i, ∇_i), $i = 1, 2$, (M^n, g) *as above,* $0 \le r_1, r_2 \le r \le k$, $r - \frac{n}{p} \le r_i - \frac{n}{p_i}$, $r - \frac{n}{p} \le (r_1 - \frac{n}{p_1}) + (r_2 - \frac{n}{p_2})$, $\frac{1}{p} \le \frac{1}{p_1} + \frac{1}{p_2}$. *Then the tensor product of sections defines a bilinear map*

$$\Omega^{p_1, r_1}_{\varepsilon_1, y}(E_1) \times \Omega^{p_2, r_2}_{\varepsilon_2, y}(E_2) \longrightarrow \Omega^{p, r}_{\varepsilon_1 + \varepsilon_2, y}(E_1 \otimes E_2).$$

Corollary 2.16 *Assume* (E_i, ∇_i), (M^n, g), k, r_i, r, p_i, p *as above,* $\varepsilon_1 = \varepsilon_2 = \varepsilon$. *Then the tensor product of sections defines a continuous map*

$$\Omega^{p_1, r_1}_{\varepsilon, y}(E_1) \times \Omega^{p_2, r_2}_{\varepsilon, y}(E_2) \longrightarrow \Omega^{p, r}_{\varepsilon, y}(E_1 \otimes E_2).$$

Corollary 2.17 *Assume* (M^n, g) *with* (I) *and* (B_k), $r \le k$, $r > \frac{n}{p}$, $p \ge 1$, $E = M \times \mathbb{R}$. *Then* $\Omega^{p,r}_{\varepsilon, y}(M^n \times \mathbb{R})$ *is an algebra.* □

Given $(E, h, \nabla^E) \longrightarrow (M^n, g)$, for fixed $E \longrightarrow M$, $r \ge 0$, $p \ge 1$, the Sobolev space $\Omega^{p,r}(E) = \Omega^{p,r}(E, h, \nabla^E, g, \nabla^g, \mathrm{dvol}_x(g))$ depends on h, $\nabla = \nabla^E$ and g. Moreover, if we choose another sequence of differential operators with injective symbol, e. g. $D, D^2, \ldots$ in case of a Clifford bundle, we should get other Sobolev spaces. Hence two questions arise, namely

1) the dependence on the choice of h, ∇^E, g,

2) the dependence on the sequence of differential operators.

We start with the first issue and investigate the dependence upon the metric connection $\nabla = \nabla^E$ of (E, h). If $\nabla' = \nabla'^E$ is another metric connection then $\eta = \nabla' - \nabla$ is a 1-form with values in $\mathfrak{G}_E$, $\nabla' - \nabla \in \Omega^1(\mathfrak{G}) = \Omega(T^*M \otimes \mathfrak{G}_E)$. Here $\mathfrak{G}$ is the bundle of the skew-symmetric endomorphisms. $\nabla = \nabla^E$ induces a connection $\nabla = \nabla^{\mathfrak{G}_E}$ in $\mathfrak{G}_E$ and hence a Sobolev norm $|\nabla' - \nabla|_{\nabla, p, r} = |\nabla' - \nabla|_{h, \nabla, g, \nabla^g, p, r}$.

Theorem 2.18 *Assume* $(E, h, \nabla^E) \longrightarrow (M^n, g)$ *with* (I), $(B_k(M))$, $(B_k(E, \nabla^E))$, $k \ge r > \frac{n}{p} + 1$. *Let* $\nabla' = \nabla'^E$ *be a second metric connection with* $(B_k(E, \nabla'^E))$ *and suppose*

$$|\nabla' - \nabla|_{\nabla, p, r-1} < \infty.$$

Then

$$\Omega^{p, \varrho}(E, h, \nabla, g) = \Omega^{p, \varrho}(E, h, \nabla', g), \quad 0 \le \varrho \le r$$

as Sobolev spaces.

2.18 can be extended to a more general

Theorem 2.19 *Let* $(E, h, \nabla) \longrightarrow (M^n, g)$ *be a Riemannian vector bundle with* (I), $(B_k(M^n, g))$, $(B_k(E, \nabla))$, $k \ge r > \frac{n}{p} + 1$. *Suppose* h' *is a fibre metric on* E *with metric connection* ∇' *and* g' *a metric on* M^n *with* (I), $(B_k(M^n, g'))$, $(B_k(E, \nabla'))$ *satisfying* $C \cdot h \le h' \le D \cdot h$, $C_1 \cdot g \le g' \le C_2 \cdot g$, $|\nabla' - \nabla|_{h, \nabla, g, p, r-1} < \infty$, $|\nabla^{g'} - \nabla^g|_{g, p, r-1} < \infty$. *Then*

$$\Omega^{p, \varrho}(E, h, \nabla, g) = \Omega^{p, \varrho}(E, h', \nabla', g'), \quad 0 \le \varrho \le r,$$

as equivalent Sobolev spaces. □

We are left with the dependence on the sequence of differential operators. This can be answered by the following two theorems.

Theorem 2.20 *Let (M^n, g) be an open Riemannian manifold satisfying $(B_\infty(M^n, g))$. Then*

$$\overset{\circ}{\Omega}{}^{q,2,2s}(M, \nabla) = \overset{\circ}{\Omega}{}^{q,2,2s}(M, \Delta), \quad 0 \le q \le n, \quad s = 0, 1, 2, \dots.$$

as equivalence of Sobolev spaces. □

Here the Ω's are Sobolev spaces of forms.

Theorem 2.21 *Let $(E, h, \nabla, \cdot) \longrightarrow (M^n, g)$ be a Clifford bundle satisfying $(B_\infty(M^n, g))$ and $(B_\infty(E, \nabla))$. If (M^n, g) is complete then*

$$\Omega^{2,r}(E, \nabla) = \Omega^{2,r}(E, D), \quad r = 0, 1, 2, \dots$$

as equivalent Sobolev spaces.

We refer to [43], [100] for the proof. □

3 Non-linear Sobolev structures

If there is given a PDE for non-linear objects like metrics, connections or maps, then one should define Sobolev structures also for such a class of objects. On closed manifolds, there arises no problem. One simply defines such a structure by means of a finite cover by charts and grafts the Euclidean definitions. Another cover yields an equivalent structure. For open manifolds, this is totally wrong. We developed a natural, intrinsic approach by means of uniform structures which we sketch below and which yields back the classical compact case.

The key notation is that of a uniform structure. Let X be a set. A filter F on X is a system of subsets which satisfies

(F_1) $M \in F, M_1 \supseteq M$ implies $M_1 \in F$.

(F_2) $M_1, \dots, M_n \in F$ implies $M_1 \cap \dots \cap M_n \in F$.

(F_3) $\emptyset \notin F$.

A system $\mathfrak{U}$ of subsets of $X \times X$ is called a uniform structure on X if it satisfies (F_1), (F_2) and

(U_1) Every $U \in \mathfrak{U}$ contains the diagonal $\Delta \subset X \times X$.

(U_2) $V \in \mathfrak{U}$ implies $V^{-1} \in \mathfrak{U}$.

(U_3) If $V \in \mathfrak{U}$ then there exists $W \in \mathfrak{U}$ such that $W \circ W \subset V$.

The sets of $\mathfrak{U}$ are called neighbourhoods of the uniform structure and $(X, \mathfrak{U})$ is called the uniform space.

$\mathfrak{B} \subset \mathfrak{P}(X \times X)$ (= sets of all subsets of $X \times X$) is a basis for a uniquely determined uniform structure if and only if it satisfies the following conditions:

(B_1) If $V_1, V_2 \in \mathfrak{B}$ then $V_1 \cap V_2$ contains an element of $\mathfrak{B}$.

(U_1') Each $V \in \mathfrak{B}$ contains the diagonal $\Delta \subset X \times X$.

(U_2') For each $V \in \mathfrak{B}$ there exists $V' \in \mathfrak{B}$ such that $V' \subseteq V^{-1}$.

(U_3') For each $V \in \mathfrak{B}$ there exists $W \in \mathfrak{B}$ such that $W \circ W \subset V$.

Every uniform structure $\mathfrak{U}$ induces a topology on X. Let $(X, \mathfrak{U})$ be a uniform space. Then for every $x \in X$, $\mathfrak{U}(x) = \{V(x)\}_{V \in \mathfrak{U}}$ is the neighbourhood filter for a uniquely determined topology on X. This topology is called the uniform topology generated by the uniform structure $\mathfrak{U}$. We refer to [105] for the proofs and further informations on uniform structures. We ask under which conditions $\mathfrak{U}$ is metrizable. A uniform space $(X, \mathfrak{U})$ is called Hausdorff if $\mathfrak{U}$ satisfies the condition

$(U_1 H)$ The intersection of all sets $\in \mathfrak{U}$ is the diagonal $\Delta \subset X \times X$.

Then the uniform space $(X, \mathfrak{U})$ is Hausdorff if and only if the corresponding topology on X is Hausdorff. The following criterion answers the question above.

Proposition 3.1 *A uniform space $(X, \mathfrak{U})$ is metrizable if and only if $(X, \mathfrak{U})$ is Hausdorff and $\mathfrak{U}$ has a countable basis $\mathfrak{B}$.* □

We present now some important examples. Let $(E, h, \nabla) \longrightarrow (M^n, g)$ be a Riemannian vector bundle, $1 \leq p < \infty$, $r > 0$, $\delta > 0$. Set

$$V_\delta = \{(\varphi, \varphi') \in C^\infty(E)^2 \,|\, |\varphi' - \varphi|_{p,r} = \left(\int\limits_M \sum_{i=0}^{r} |\nabla^i(\varphi' - \varphi)|_\alpha^p \,\mathrm{dvol}_x(g) \right)^{\frac{1}{2}} < \delta\}.$$

Then it is evident that $\mathfrak{B} = \{V_\delta\}_{\delta > 0}$ is a basis for a metrizable uniform structure $\mathfrak{U}^{p,r}(C^\infty(E))$, $(C^\infty(E), \mathfrak{U}^{p,r}(C^\infty(E)))$ is a uniform space. Let $(\overline{C^\infty(E)}^{\mathfrak{U}^{p,r}}, \overline{\mathfrak{U}^{p,r}})$ be the completion. We started with the local metric $d_\varphi(\varphi, \varphi') = d(\varphi, \varphi') = |\varphi - \varphi'|_{p,r}$ in the dense subspace $C^\infty(E)$ which can be extended to a complete local metric $|\cdot - \cdot|_{p,r}$ what is locally equivalent to the existence of a global one.

Lemma 3.2 *$\overline{C^\infty(E)}$ is locally arcwise connected.*

Corollary 3.3 **a)** *In $\overline{C^\infty(E)}^{\mathfrak{U}^{p,r}}$ coincide components and arc components.*

b) *$\overline{C^\infty(E)}^{\mathfrak{U}^{p,r}}$ has a representation as a topological sum*

$$\overline{C^\infty(E)}^{\mathfrak{U}^{p,r}} = \sum_{i \in I} \mathrm{comp}^{p,r}(\varphi_i)$$

c)

$$\mathrm{comp}^{p,r}(\varphi) = \{\varphi' \in \overline{C^\infty(E)}^{\mathfrak{U}^{p,r}} \,|\, |\varphi - \varphi'|_{p,r}^{ext} < \infty\} = \varphi + \overline{\Omega}^{p,r}(E, \nabla),$$

i.e. each component is an affine Sobolev space.

Corollary 3.4 *On a compact manifold there is only one (arc) component, namely*

$$\mathrm{comp}^{p,r}(0) = \overline{\Omega}^{p,r}(E, \nabla).$$

We write $\Omega^{p,r}(C^\infty(E))$ for $\overline{C^\infty(E)}^{\mathfrak{U}^{p,r}(C^\infty(E))}$.

Remarks 3.5 **a)** We see for $\varphi = 0$ that the zero component $\mathrm{comp}^{p,r}(0)$ coincides with the Sobolev space $\overline{\Omega}^{p,r}(E, \nabla)$. Insofar our approach yields a generalization of Sobolev spaces.

b) It is very easy to see that the index set I is uncountable if (M^n, g) is open. "Each growth generalies its component." On compact manifolds, there is only one growth, namely no growth, hence there is only one component.

c) Let $\{X_i\}_{i\in I}$ be a family of disjoint metric spaces, d_i the metric on X_i. Then there exists a metric d on the topological sum $X = \sum_{i\in I} X_i$ s.t. d induces the uniform structure on X_i which belongs to d_i (cf. [105], p. 120). This is the situation in corollary 3.3 and c). □

Let us consider other choices of V_δ. Set $V_\delta = \{(\varphi, \varphi') \in L_{1,\mathrm{loc}}(E)^2 \,|\, |\varphi - \varphi'|_{p,r} < \delta\}$. In $|\varphi - \varphi'|_{p,r}$ we take distributional derivatives. Then $\mathfrak{B} = \{V_\delta\}_{\delta>0}$ is a basis for a metrizable uniform structure $\mathfrak{U}^{p,r}(L_{1,\mathrm{loc}}(E))$ which is already complete, $(\overline{L_{1,\mathrm{loc}}(E)}, \overline{\mathfrak{U}}^{p,r}(L_{1,\mathrm{loc}})) = (L_{1,\mathrm{loc}}, \mathfrak{U}^{p,r}(L_{1,\mathrm{loc}}(E))$. The background for completeness is the fact that the Sobolev spaces of distributions are already complete. We write $\Omega^{p,r}(L_{1,\mathrm{loc}}(E))$ for $(L_{1,\mathrm{loc}}(E), \mathfrak{U}^{p,r}(L_{1,\mathrm{loc}}))$ which is a complete uniform space.

Proposition 3.6 **a)** $\Omega^{p,r}(L_{1,\mathrm{loc}}(E))$ *is locally arcwise connected.*

b) *In* $\Omega^{p,r}(L_{1,\mathrm{loc}}(E))$ *coincide components and arc components.*

c) $\Omega^{p,r}(L_{1,\mathrm{loc}}(E))$ *has a representation as a topological sum*

$$\Omega^{p,r}(L_{1,\mathrm{loc}}(E)) = \sum_{j\in J} \mathrm{comp}^{p,r}(\varphi_j).$$

d)

$$\mathrm{comp}^{p,r}(\varphi) = \{\varphi' \in \Omega^{p,r}(L_{1,\mathrm{loc}}(E)) \,|\, |\varphi - \varphi'|_{p,r} < \infty\} = \varphi + \Omega^{p,r}(E, \nabla).$$

□

Consider finally

$$V_\delta = \{(\varphi, \varphi') \in C_c^\infty(E)^2 \,|\, |\varphi - \varphi'|_{p,r} < \delta\}.$$

We obtain $\mathfrak{U}^{p,r}(C_c^\infty(E))$, $(\overline{C_c^\infty(E)}^{\mathfrak{U}^{p,r}}, \overline{\mathfrak{U}}^{p,r}(C_c^\infty(E))) = \Omega^{p,r}(C_c^\infty(E))$, locally arcwise connectedness, a topological sum representation and

$$\mathrm{comp}^{p,r}(\varphi) = \varphi + \mathring{\Omega}^{p,r}(E, \nabla).$$

But $\Omega^{p,r}(C_c^\infty(E))$ consists of one component as the following remark shows.

Remark 3.7 If $\varphi \in C_c^\infty(E)$ then in $\Omega^{p,r}(C_c^\infty(E))$

$$\mathrm{comp}(\varphi) = \mathrm{comp}^{p,r}(0) = \mathring{\Omega}^{p,r}(E, \nabla).$$

□

From section 2 follows immediately

Corollary 3.8 *If* $(E, h, \nabla) \longrightarrow (M^n, g)$ *satisfies* (I), $(B_k(M^n, g))$, $(B_k(E, \nabla))$ *then*

$$\Omega^{p,r}(L_{1,\mathrm{loc}}(E)) = \Omega^{p,r}(C^\infty(E))$$

for $0 \leq r \leq k + 2$. □

Remark 3.9 A long further series of equivalences for uniform spaces can be stated if we apply all of our invariance properties of Sobolev spaces. □

The advantage of this approach is that we can develop e.g. a Sobolev theory of PDEs without decay conditions for the sections. The classical theory is a theory between the zero components $\mathrm{comp}(0)$. Our framework allows a quite parallel theory as maps between other components. Clearly ∇ maps $\mathrm{comp}^{p,r}(\varphi)$ into $\mathrm{comp}^{p,r-1}(\varphi)$. If A is a differential operator which maps $\Omega^{p,r}(E)$ into $\Omega^{p,r-m}(F)$ then A maps $\mathrm{comp}^{p,r}(\varphi) \subset \Omega^{p,r}(L_{1,\mathrm{loc}}(E))$ into $\mathrm{comp}^{p,r-m}(A\varphi) \subset \Omega^{p,r-m}(L_{1,\mathrm{loc}}(F))$. A necessary condition for the solvability of $A\varphi = \psi$, $\varphi \in \mathrm{comp}^{p,r}(\varphi_0)$ to find, $\psi \in \Omega^{p,r-m}(L_{1,\mathrm{loc}}(F))$ or $\in \Omega^{p,r-m}(C^\infty(F))$ given, is that $A\varphi_0 \in \mathrm{comp}^{p,r-m}(\psi)$ etc. We will here not establish the complete PDE-theory for this setting. It should appear elsewhere.

A similar setting can be established for the Banach-Hölder theory. Set

$$V_\delta = \{(\varphi, \varphi') \in C^\infty(E)^2 \,|\, {}^{b,m}|\varphi - \varphi'| = \sum_{i=0}^{m} \sup_x |\nabla^i(\varphi - \varphi')|_x < \delta\}$$

and $\mathfrak{B} = \{V_\delta\}_{\delta>0}$. $\mathfrak{B}$ is a basis for a metrizable uniform structure ${}^{b,m}\mathfrak{U}(C^\infty(E))$. Let $(\overline{C^\infty(E)}^{\,b,m\mathfrak{U}}, {}^{b,m}\overline{\mathfrak{U}})$ be the completion. Then we get properties absolutely parallel to the assertions 3.2 – 3.6,

$$\begin{aligned} \overline{C^\infty(E)}^{\,b,m\mathfrak{U}} &\equiv {}^{b,m}\Omega(C^\infty(E)) = \sum_{j\in J} {}^{b,m}\mathrm{comp}(\varphi_j), \\ {}^{b,m}\mathrm{comp}(\varphi) &= \{\varphi' \in {}^{b,m}\Omega(C^\infty(E)) \,|\, |\varphi - \varphi'| < \infty\} = \varphi + {}^{b,m}\Omega(E, \nabla). \end{aligned}$$

Similarly

$$V_\delta = \{(\varphi, \varphi') \in C^\infty(E)^2 \,|\, {}^{b,m,\alpha}|\varphi - \varphi'| < \delta\}$$

and $\mathfrak{B} = \{V_\delta\}_{\delta>0}$ define ${}^{b,m,\alpha}\mathfrak{U}(C^\infty(E))$, the completion ${}^{b,m,\alpha}\Omega(C^\infty(E)) = \sum_{j\in J} {}^{b,m,\alpha}\mathrm{comp}(\varphi_j)$ and

$${}^{b,m,\alpha}\mathrm{comp}(\varphi) = \varphi + {}^{b,m,\alpha}\Omega(E, \nabla). \tag{3.1}$$

Here ${}^{b,m,\alpha}|\varphi|$ is defined as

$${}^{b,m}|\varphi| + \sup_{x,y\in M} \sup_{c\in G(x,y)} \frac{|\tau(c)\nabla^m\varphi(x) - \nabla^m\varphi(y)|}{d(x,y)^\alpha},$$

where $G(x,y) = \{$ length minimizing geodesics joining x and $y\}$, $\tau(c)$ is parallel translation along c from $\pi^{-1}(x)$ to $\pi^{-1}(y)$ and $d(x,y)$ is the distance from x to y.

Remark 3.10 Our Sobolev embedding theorems from section 2 induce embedding theorems for components of corresponding uniform spaces, e.g. if we have (I), (B_k), $r > \frac{n}{p} + m$ then

$$\Omega^{p,r}(C^\infty(E)) \supset \mathrm{comp}^{p,r}(\varphi) \hookrightarrow {}^{b,m}\mathrm{comp}(\varphi) \subset {}^{b,m}\Omega(C^\infty(E)).$$

□

We discuss another example, which is important in Teichmüller theory for open surfaces. That is the space of bounded conformal factors, adapted to a Riemannian metric g.

Let

$$\mathcal{P}_m(g) = \{\varphi \in C^\infty(M) \,|\, \inf_{x\in M} \varphi(x) > 0, \sup_{x\in M} \varphi(x) < \infty,\ |\nabla^i \varphi|_{g,x} \le C_i,\ 0 \le i \le m\}.$$

Set for $1 \le p < \infty$, $r \in \mathbb{Z}_+$, $\delta > 0$,

$$V_\delta = \{(\varphi, \varphi') \in \mathcal{P}_m(g)^2 \,|\, |\varphi - \varphi'|_{g,p,r} := \left(\int \sum_{i=0}^{r} |(\nabla^g)^i(\varphi - \varphi')|^p_{g,x}\, \mathrm{dvol}_x(g)\right)^{\frac{1}{2}} < \delta\}.$$

Then $\mathfrak{B} = \{V_\delta\}_{\delta>0}$ is a basis for a metrizable uniform structure.

Let $\overline{\mathcal{P}}^p_{m,r}(g)$ the completion,

$$C^1\mathcal{P} = \{\varphi \in C^1(M) \,|\, \inf_{x\in M} \varphi(x) > 0, \sup_{x\in M} \varphi(x) < \infty\}$$

and set

$$\mathcal{P}^{p,r}_m(g) = \overline{\mathcal{P}}^p_{m,r}(g) \cap C^1\mathcal{P}.$$

$\mathcal{P}^{p,r}_m(g)$ is locally contractible, hence locally arcwise connected and hence components coincide with arc components. Let

$$U^{p,r}_m(\varphi) = \{\varphi' \in \mathcal{P}^{p,r}_m(g) \,|\, |\varphi - \varphi'|_{g,p,r} < \infty\}$$

and denote by $\mathrm{comp}(\varphi) = \mathrm{comp}^{p,r}_m(\varphi)$ the component of φ in $\mathcal{P}^{p,r}_m(g)$. $|\,|_{g,p,r}$ in (3.1) means the local extended metric, i.e. it is defined by taking distributional derivatives.

Theorem 3.11 *$\mathcal{P}^{p,r}_m(g)$ has a representation as topological sum*

$$\mathcal{P}^{p,r}_m(g) = \sum_{i\in I} \mathrm{comp}(\varphi_i)$$

and

$$\mathrm{comp}(\varphi) = U^{p,r}_m(\varphi).$$

□

Remark 3.12 On a compact manifold there is only one component, the component $\mathrm{comp}(1)$.

□

We come back to this example in section 10 when we discuss Teichmüller theory on open manifolds.

Let M^n be an open smooth manifold, $\mathcal{M} = \mathcal{M}(M)$ be the space of all Riemannian metrics. We want to endow $\mathcal{M}$ with a canonical intrinsic topology either in the C^m- or Sobolev setting, depending on the subsequent investigation.

Let $g \in \mathcal{M}$. We define

$$^{b}U(g) = \{g' \in \mathcal{M} \mid {}^{b}|g - g'| := \sup_{x \in M} |g - g'|_{g,x} < \infty, {}^{b}|g - g'|_{g'} < \infty\}.$$

Then, it is easy to see that ${}^{b}U(g)$ coincides with the quasi isometry class of g, i.e., $g' \in {}^{b}U(g)$ if and only if there exist $C, C' > 0$ such that

$$C \cdot g' \leq g \leq C' \cdot g' \tag{3.2}$$

holds in the sense of positive definite forms. In particular $g' \in {}^{b}U(g)$ if and only if $g \in {}^{b}U(g')$.

To endow $\mathcal{M}$ with canonical topologies we use the language of uniform structures.

Set for $m \geq 1$, $\delta > 0$, $C(n, \delta) = 1 + \delta + \delta\sqrt{2n(n-1)}$

$$\begin{aligned} V_\delta \quad = \quad & \{(g, g') \in \mathcal{M} \mid g' \in {}^{b}U(g), C(n,\delta)^{-1} \cdot g \leq g' \leq C(n,\delta) \cdot g \text{ and} \\ & {}^{b,m}|g - g'|_g := {}^{b}|g - g'|_g + \sum_{j=0}^{m-1} {}^{b}|(\nabla^g)^j (\nabla^g - \nabla^{g'})|_g < \delta\}. \end{aligned}$$

Proposition 3.13 *The set $\mathfrak{B} = \{V_\delta\}_{\delta > 0}$ is a basis for a metrizable uniform structure on $\mathcal{M}$.*

Denote ${}^{b}_{m}\mathcal{M} = (\mathcal{M}, {}^{b,m}\mathfrak{U}(\mathcal{M}))$ and by ${}^{b,m}\mathcal{M}$ the completion. It has been proven by Salomonsen that ${}^{b,m}\mathcal{M}$ still consists of positive definite elements, which are of class C^m.

Remark 3.14 We endowed by our procedure $\mathcal{M}$ with a canonical intrinsic C^m-topology without choice of a cover or a special g to define the C^m-distance. According to the definition of the uniform topology, for $g \in {}^{b,m}\mathcal{M}$

$$\{{}^{b,m}U_\varepsilon\}_{\varepsilon > 0} = \{g' \in {}^{b,m}\mathcal{M} \mid {}^{b}|g - g'|_{g'} < \infty, {}^{b,m}|g - g'|_g < \infty\}$$

is a neighborhood basis in this topology. □

Set

$${}^{b,m}U(g) = \{g' \in {}^{b,m}\mathcal{M} \mid {}^{b}|g - g'|_{g'} < \infty, {}^{b,m}|g - g'|_g < \infty\}.$$

Proposition 3.15 *Denote by* $\operatorname{comp}(g)$ *the component of* $g \in {}^{b,m}\mathcal{M}$. *Then*

$$\operatorname{comp}(g) \equiv {}^{b,m}\operatorname{comp}(g) = {}^{b,m}U(g).$$

□

Theorem 3.16 *The space ${}^{b,m}\mathcal{M}$ has a representation as a topological sum*

$${}^{b,m}\mathcal{M} = \sum_{i \in I} {}^{b,m}U(g_i). \tag{3.3}$$

□

Remark 3.17 On a compact manifold $\mathcal{M}$, the index set I consists of one element. One has only one component. On compact manifolds the notion of growth at infinity does not exist. □

Theorem 3.18 *Each component of* ${}^{b,m}\mathcal{M}$ *is a Banach manifold.*

For later use we restrict ourselves additionally to metrics with bounded geometry. Let (M^n, g) be open. Consider the conditions (I) and (B_k) and

$$\begin{aligned} \mathcal{M}(I) &= \{g \in \mathcal{M} \,|\, g \text{ satisfies}(I)\}, \\ \mathcal{M}(B_k) &= \{g \in \mathcal{M} \,|\, g \text{ satisfies}(B_k)\}, \\ \mathcal{M}(I, B_k) &= \mathcal{M}(I) \cap \mathcal{M}(B_k). \end{aligned}$$

Now we want to introduce Sobolev uniform structures into the space of metrics. Let now $k \geq r > \frac{n}{p} + 1$, $\delta > 0$, $C(n,\delta) = 1 + \delta + \delta\sqrt{2n(n-1)}$,

$$\begin{aligned} V_\delta &= \Big\{(g, g') \in \mathcal{M}(I, B_k) \times \mathcal{M}(I, B_k) \,| \\ &\quad |g - g'|_{g,p,r} = \Big(\int (|g - g'|^p_{g,x} + \sum_{i=0}^{r-1} |(\nabla^g)^i(\nabla^{g'} - \nabla^g)|^p_{g,x}) \,\mathrm{dvol}_x(g)\Big)^{\frac{1}{p}} < \delta\Big\}. \end{aligned}$$

Proposition 3.19 *The set* $\{V_\delta\}_{\delta>0}$ *is a basis for a metrizable uniform structure on* $\mathcal{M}(I, B_k)$.

Denote $\mathcal{M}^p_r(I, B_k)$ as $(\mathcal{M}(I, B_k), \mathfrak{U}^{p,r}(\mathcal{M}(I, B_k)))$ and by $\mathcal{M}^{p,r}(I, B_k)$ the completion. It was proven by Salomonsen that the completion yields only positive definite elements, i.e. we still remain in the space of C^1 Riemannian metrics.

For $g \in \mathcal{M}^{p,r}(I, B_k)$

$$\begin{aligned} \{U^{p,r}_\varepsilon(g)\}_{\varepsilon>0} &= \big\{\{g' \in \mathcal{M}^{p,r}(I, B_k) \,|\, {}^b|g - g'|_g < \infty, {}^b|g - g'|_{g'} < \infty, \\ &\quad |g - g'|_{g,p,r} < \varepsilon\}\big\}_{\varepsilon>0} \end{aligned}$$

is a neighborhood basis in the uniform topology. There arises a small difficulty. $g \in \mathcal{M}^{p,r}(I, B_k)$ must not be smooth and hence $|g - g'|_{g,p,r}$ must not be defined immediately. But in this case we use the density of $\mathcal{M}(I, B_k) \subset \mathcal{M}^{p,r}(I, B_k)$ and apply a suitable approximations procedure (cf. [43]).

Proposition 3.20 *The space* $\mathcal{M}^{p,r}(I, B_k)$ *is locally contractible.*

For the proof we refer to [44], lemma 3.8. □

Proposition 3.21 *In* $\mathcal{M}^{p,r}(I, B_k)$ *components and arc components coincide.* □

Set for $g \in \mathcal{M}^{p,r}(I, B_k)$

$$\begin{aligned} U^{p,r}_\varepsilon(g) &= \{g' \in \mathcal{M}^{p,r}(I, B_k) \,|\, {}^b|g - g'|_g < \infty, {}^b|g - g'|_{g'} < \infty, \\ &\quad |g - g'|_{g,p,r} < \infty\}. \end{aligned}$$

Proposition 3.22 *Denote by* $\mathrm{comp}(g)$ *the component of* $g \in \mathcal{M}^{p,r}(I, B_k)$. *Then,*

$$\mathrm{comp}(g) \equiv \mathrm{comp}^{p,r}(g) = U^{p,r}(g).$$

Theorem 3.23 *Let* M^n *be open,* $k \geq r > \frac{n}{p} + 1$. *Then* $\mathcal{M}^{p,r}(I, B_k)$ *has a representation as a topological sum*

$$\mathcal{M}^{p,r}(I, B_k) = \sum_{i \in I} U^{p,r}(g_i).$$

□

We can reformulate theorems 2.18 and 2.19 as

Proposition 3.24 *Let $g \in \mathcal{M}(I, B_k)$, $k \geq r > \frac{n}{p} + 1$, $g' \in \mathrm{comp}(g) \subset \mathcal{M}^{p,r}(I, B_k)$. Then*

$$\Omega^{p,r}(T^u_v, g) = \Omega^{p,r}(T^u_v, g')$$

as equivalence of Sobolev spaces. □

Theorem 3.25 *Assume $k \geq r > \frac{n}{p} + 1$. Then, each component of the space $\mathcal{M}^{p,r}(I, B_k)$ is a Banach manifold and for $i = 2$ it is a Hilbert manifold.*

We introduce now in quite analogous manner uniform structures of connections. Let $(E, h) \longrightarrow (M^n, g)$ be a Riemannian vector bundle. Denote by $\mathcal{C}_E$ the set of all metric connections in E and set for $m \in \mathbb{Z}_+$, $\delta > 0$

$$V_\delta = \{(\nabla, \nabla') \in \mathcal{C}_E^2 | \, {}^{b,m}|\nabla' - \nabla|_\nabla < \delta\},$$

where, according to our definitions in section 2,

$${}^{b,m}|\nabla' - \nabla|_\nabla = \sum_{\mu=0}^{m} \sup_{x \in M} |\nabla^\mu(\nabla' - \nabla)|_x \sim \sup_{\substack{x \in M \\ 0 \leq \mu \leq m}} |\nabla^\mu(\nabla' - \nabla)|_x.$$

Proposition 3.26 $\mathfrak{B} = \{V_\delta\}_{\delta > 0}$ *is a basis for a metrizable uniform structure ${}^{b,m}\mathfrak{U}(\mathcal{C}_E)$ on $\mathcal{C}_E$.*

Denote ${}^b_m\mathcal{C}_E = (\mathcal{C}_E, {}^{b,m}\mathfrak{U}(\mathcal{C}_E))$ and by ${}^{b,m}\mathcal{C}_E$ the completion.

Proposition 3.27 **a)** *${}^{b,m}\mathcal{C}_E$ is locally arcwise connected, hence components coincide with arc components.*

b) *${}^{b,m}\mathcal{C}_E$ has a representation as topological sum*

$${}^{b,m}\mathcal{C}_E = \sum_{i \in I} {}^{b,m}\mathrm{comp}(\nabla_i).$$

c) *For $\nabla \in {}^{b,m}\mathcal{C}_E$*

$${}^{b,m}\mathrm{comp}(\nabla) = \{\nabla' \in {}^{b,m}\mathcal{C}_E | \, {}^{b,m}|\nabla' - \nabla|_\nabla < \infty\} = \nabla + {}^{b,1}\Omega^1(\mathfrak{g}_E),$$

where $\mathfrak{g}_E$ are the skew symmetric endomorphisms of E and the connection in $\mathfrak{g}_E$ is defined by $\nabla^{\mathfrak{g}}\varphi = [\nabla^E, \varphi]$. □

Remark 3.28 On a compact manifold we have only one component. □

Suppose that (M^n, g) satisfies (B_k) and consider the set $\mathcal{C}_E(B_k) = \{\nabla \in \mathcal{C}_E | (E, \nabla) \text{ satisfies } (B_k)\}$. Restricting ${}^{b,m}\mathfrak{U}$ to $\mathcal{C}_E(B_k)$ yields ${}^b_m\mathcal{C}(B_k)$ and the completion ${}^{b,m}\mathcal{C}_E(B_k)$.

Proposition 3.29 *Suppose $m \geq k + 1$.*

a) *${}^{b,m}\mathcal{C}_E(B_k)$ is locally arcwise connected, hence components coincide with arc components.*

b) *${}^{b,m}\mathcal{C}_E(B_k)$ has a representation as topological sum*

$${}^{b,m}\mathcal{C}_E(B_k) = \sum_{j \in J} {}^{b,m}\mathrm{comp}(\nabla_j).$$

c) *For* $\nabla \in {}^{b,m}\mathcal{C}_E(B_k)$*,*

$${}^{b,m}\mathrm{comp}(\nabla) = \{\nabla' \in {}^{b,m}\mathcal{C}_E(B_k) | \, {}^{b,m}|\nabla' - \nabla|_\nabla < \infty\} = \nabla + {}^{b,m}\Omega^1(\mathfrak{g}_E).$$

□

Corollary 3.30 *If a component* ${}^{b,m}\mathrm{comp}(\nabla) \subset {}^{b,m}\mathcal{C}_E$ *contains an element of* ${}^{b,m}\mathcal{C}_E(B_k)$*,* $m \geq k+1$*, then the whole component* ${}^{b,m}\mathrm{comp}(\nabla)$ *is contained in* ${}^{b,m}\mathcal{C}_E(B_k)$ *and coincides with the corresponding component in* ${}^{b,m}\mathcal{C}_E(B_k)$*.* □

Assume now that (M^n, g) satisfies (I) and (B_k), $k > \frac{n}{p} + 1$ and consider again $\mathcal{C}_E(B_k)$. Set for $k \geq r > \frac{n}{p} + 1$, $\delta > 0$

$$V_\delta = \{(\nabla, \nabla') \in \mathcal{C}_E^2(B_k) | |\nabla' - \nabla|_{\nabla,p,r} < \delta\},$$

where

$$|\nabla' - \nabla|_{\nabla,p,r} = \left(\int \sum_{i=0}^{r} |\nabla^i(\nabla' - \nabla)|_x^p \, \mathrm{dvol}_x(g) \right)^{\frac{1}{p}}.$$

Proposition 3.31 $\mathfrak{B} = \{V_\delta\}_{\delta > 0}$ *is a basis for a metrizable uniform structure* $\mathfrak{U}^{p,r}(\mathcal{C}_E(B_k))$*.* □

Denote $\mathcal{C}_{E,r}^p(B_k) = (\mathcal{C}_E(B_k), \mathfrak{U}^{p,r}(\mathcal{C}_E(B_k)))$ and by $\overline{\mathcal{C}}_E^{p,r}(B_k)$ its completion.

Proposition 3.32 **a)** $\overline{\mathcal{C}}_E^{p,r}(B_k)$ *is locally arcwise connected, hence components coincide with arc components.*

b) $\overline{\mathcal{C}}_E^{p,r}(B_k)$ *has a representation as topological sum*

$$\overline{\mathcal{C}}_E^{p,r}(B_k) = \sum_{i \in I} \mathrm{comp}^{p,r}(\nabla_i).$$

c) *For* $\nabla \in \mathcal{C}_E(B_k)$*,*

$$\mathrm{comp}^{p,r}(\nabla) = \left\{\nabla' \in \overline{\mathcal{C}}_E^{p,r}(B_k) \middle| |\nabla' - \nabla|_{\nabla,p,r} < \infty\right\} = \nabla + \Omega^{1,p,r}(\mathfrak{g}_E, \nabla).$$

□

Finally, we define

$$\mathcal{C}(E, B_k, f, p) = \mathcal{C}_E(B_k, f, p) = \{\nabla \in \mathcal{C}_E(B_k) | \int |R^\nabla|^p \, \mathrm{dvol} < \infty\} \subset \mathcal{C}_E(B_k).$$

f, p stands for finite p-action. Restriction of $\mathfrak{U}^{p,r}(\mathcal{C}_E(B_k))$ to $\mathcal{C}_E(B_k, f, p)$ yields $\mathcal{C}_{E,r}^p(B_k, f, p)$ and the completion $\overline{\mathcal{C}}_E^{p,r}(B_k, f, p)$. We describe the structure of $\overline{\mathcal{C}}_E^{p,r}(B_k, f, p)$, where we suppose as above $k \geq r > \frac{n}{p} + 1$.

Proposition 3.33 **a)** $\overline{\mathcal{C}}_E^{p,r}(B_k, f, p)$ *is locally arcwise connected, hence components coincide with arc components.*

b) $\overline{\mathcal{C}}_E^{p,r}(B_k, f, p)$ *has a representation as a topological sum*

$$\overline{\mathcal{C}}_E^{p,r}(B_k, f, p) = \sum_{j \in J} \mathrm{comp}(\nabla_j).$$

c) *For* $\nabla \in \mathcal{C}_E(B_k, f, p)$,

$$\text{comp}(\nabla) = \{\nabla' \in \mathcal{C}_E^{p,r}(B_k, f, p) | |\nabla' - \nabla|_{\nabla,p,r} < \infty\} = \nabla + \Omega^{1,p,r}(\mathfrak{g}_E, \nabla).$$

□

Corollary 3.34 *If* $\nabla \in \mathcal{C}_E^{p,r}(B_k, f, p)$ *then its component in* $\mathcal{C}_E^{p,r}(B_k, f, p)$ *coincides with its corresponding component in* $\mathcal{C}_E^{p,r}(B_k)$. □

Our last class of examples for non-linear Sobolev structures are manifolds of maps and diffeomorphism groups.

Let (M^n, g), $(N^{n'}, h)$ be open, complete, satisfying (I) and (B_k) and let $f \in C^\infty(M, N)$. Then the differential $f_* = df$ is a section of $T^*M \otimes f^*TN$. f^*TM is endowed with the induced connection $f^*\nabla^h$ which is locally given by

$$\Gamma_{i\mu}^\nu = \partial_i f^\alpha(x) \Gamma_{\alpha,\mu}^{h,\nu}(f(x)), \quad \partial_i = \frac{\partial}{\partial x^i}.$$

∇^g and $f^*\nabla^h$ induce metric connections ∇ in all tensor bundles $T_s^q(M) \otimes f^*T_\nu^u(N)$. Therefore $\nabla^m df$ is well defined. Since (I) and (B_0) imply the boundedness of the g_{ij}, g^{km}, $h_{\mu\nu}$ in normal coordinates, the conditions df to be bounded and $\partial_i f$ to be bounded are equivalent.

In local coordinates

$$\sup_{x \in M} |df|_x = \sup \text{tr}^g(f^*h) = \sup g^{ij} h_{\mu\nu} \partial_j f^\mu \partial_i f^\nu.$$

For (M^n, g), $(N^{n'}, h)$ of bounded geometry up to order k and $m \le k$ we denote by $C^{\infty,m}(M, N)$ the set of all $f \in C^\infty(M, N)$ satisfying

$$^{b,m}|df| := \sum_{\mu=0}^{m-1} \sup_{x \in M} |\nabla^\mu df|_x < \infty.$$

Assume (M^n, g), $(N^{n'}, h)$ are open, complete, and of bounded geometry up to order k, $r \le m \le k$, $1 < p < \infty$, $r > \frac{n}{p} + 1$. Consider $f \in C^{\infty,m}(M, N)$. According to chapter I, theorem 2.9 b) for $r > \frac{n}{p} + s$

$$\Omega^{p,r}(f^*TN) \hookrightarrow {}^{b,s}\Omega(f^*TN), \tag{3.4}$$

$$^{b,s}|Y| \le D \cdot |Y|_{p,r}, \tag{3.5}$$

where $|Y|_{p,r} = \left(\int \sum_{i=0}^{r} |\nabla^i Y|^p \,\text{dvol} \right)^{\frac{1}{p}}$. Set for $\delta > 0$, $\delta \cdot D \le \delta_N < r_{\text{inj}}(N)/2$, $1 < p < \infty$, $V_\delta = \{(f, g) \in C^{\infty,m}(M, N) \times C^{\infty,m}(M, N) \,|\exists Y \in \Omega_r^p(f^*TN)$ such that $g = g_Y = \exp Y$ and $|Y|_{p,r} < \delta\}$.

Proposition 3.35 $\mathfrak{B} = \{V_\delta\}_{0<\delta<r_{\text{inj}}(N)/2D}$ *is a basis for uniform structure* $\mathfrak{U}^{p,r}(C^{\infty,m}(M, N))$.

□

$\mathfrak{U}^{p,r}(C^{\infty,m}(M, N))$ is metrizable. Let $^m\Omega^{p,r}(M, N)$ be the completion of $C^{\infty,m}(M, N)$. From now on we assume $r = m$ and denote $^r\Omega^{p,r}(M, N) = \Omega^{p,r}(M, N)$.

Theorem 3.36 *Let (M^n, g), (N^n, h) be open, complete, of bounded geometry up to order k, $1 < p < \infty$, $r \le k$, $r > \frac{n}{p} + 1$. Then each component of $\Omega^{p,r}(M,N)$ is a C^{1+k-r}-Banach manifold, and for p = 2 it is a Hilbert manifold.* □

Let (M^n, g) be open, complete, oriented, of bounded geometry up to order k, $1 < p < \infty$, $k > r > \frac{n}{p} + 1$. Set

$$\mathcal{D}^{p,r}(M) = \left\{ \begin{array}{ll} f \in \Omega^{p,r}(M,M) & \Big| \quad \begin{array}{l} f \text{ is injective, surjective, preserves} \\ \text{orientation and } |\lambda|_{\min}(df) > 0 \end{array} \end{array} \right\}.$$

Theorem 3.37 *$\mathcal{D}^{p,r}(M)$ is open in $\Omega^{p,r}(M,M)$, in particular, each component is a C^{1+k-r}-Banach manifold.* □

We use the completed Sobolev structures of this section later in the sections 8, 9, 10, 13 and 14.

4 Self-adjoint linear differential operators on manifolds and their spectral theory

For the sake of clarity, we recall in a few words the basic notations of spectral theory.

Let $X = (X, \langle,\rangle)$ be a Hilbert space over $\mathbb{C}$, $A : \mathcal{D}_A \longrightarrow X$, $\mathcal{D}_A \subset X$, a densely defined linear operator. If $\mathcal{D}_A = X$ and $|Ax| \le c \cdot |x|$, $c \in \mathbb{R}$, for all $x \in X$, then A is bounded and $|A| = \inf\{c \,|\, |Ax| \le c \cdot |x| \text{ for all } x \in X\}$ is called the operator norm of A, $|A| = |A|_{X \to X}$. Here $|y| = \langle y, y\rangle^{\frac{1}{2}}$. $L(X,X)$ shall denote the Banach algebra of bounded operators $A : X \longrightarrow X$. Similarly one defines $|A|_{X \to Y}$ for a bounded $A : X \longrightarrow Y$. The most important case for us is $A : \mathcal{D}_A \longrightarrow X$, $\mathcal{D}_A \ne X$. But we assume once and for all that $\mathcal{D}_A \subset X$ be dense. $A : \mathcal{D}_A \longrightarrow X$ is called closed if $x_i \longrightarrow x$, $x_i \in \mathcal{D}_A$, $Ax_i \longrightarrow y$ imply $x \in \mathcal{D}_A$ and $Ax = y$. This can be reformulated as follows. Let $\Gamma(A) = \{(x, Ax) | x \in \mathcal{D}_A\}$ be the graph of A. Then A is closed if and only if $\Gamma(A)$ is closed in the Hilbert space $(X \times X, \langle,\rangle \oplus \langle,\rangle)$. A fundamental theorem states that A closed and $\mathcal{D}_A = X$ imply A bounded. For later applications we have in mind mainly unbounded differential operators, $\mathcal{D}_A \subset X$. Set for $x, y \in \mathcal{D}_A$ $\langle x, y\rangle_A := \langle x, y\rangle + \langle Ax, Ay\rangle$, $|x|_A = \langle x, x\rangle_A^{\frac{1}{2}}$. $\langle x, y\rangle_A$ is a scalar product in $\mathcal{D}_A$ and $|\;|_A$ a norm. Then A is closed if and only if $(\mathcal{D}_A, \langle,\rangle_A)$ is a Hilbert space, i.e. $|\;|_A$ is complete. We recall that $B \supseteq A$ means $D_B \supseteq \mathcal{D}_A$ and $B|_{\mathcal{D}_A} = A$. A is called closable if there exists a closed $B \supseteq A$.

Theorem 4.1 *If A is closable, then there exists a unique minimal closed extension $\bar{A}$, called the closure of A. Furthermore $D_{\bar{A}}$ is the closure of $\mathcal{D}_A$ with respect to $|\;|_A$ and $\Gamma(\bar{A}) = \overline{\Gamma(A)}$.*

□

An intrinsic criterion for closability of A goes as follows: A is closable if and only if $x_i \longrightarrow 0$, $x_i \in \mathcal{D}_A$, $(Ax_i)_i$ a Cauchy sequence, imply $Ax_i \longrightarrow 0$.

Warning. The closure of $\mathcal{D}_A$ with respect to $|\;|_A$ or the closure of $\Gamma(A)$ need not necessarily define a closed operator.

If $A : \mathcal{D}_A \longrightarrow X$ is densely defined (as always assumed) then for every $y \in X$ there exists at most one element $y^* \in X$ such that $\langle Ax, y\rangle = \langle x, y^*\rangle$ for all $x \in \mathcal{D}_A$. Then we define $D_{A^*} = \{y \in X \,|\exists y^* \in X \text{ such that } \langle Ax, y\rangle = \langle x, y^*\rangle \text{ for all } x \in \mathcal{D}_A\}$ and $A^*y := y^*$. D_{A^*} is a linear space, A^* is a linear closed operator and is called adjoint to A. $B \supseteq A$ implies $A^* \supseteq B^*$. If A is closable then $A^* = (\bar{A})^*$.

A special class of closable operators are symmetric operators. $A : \mathcal{D}_A \longrightarrow X$ is called symmetric (or formally self-adjoint) if $\langle Ax, y\rangle = \langle x, Ay\rangle$ for all $x, y \in \mathcal{D}_A$. A is called self-adjoint if $A = A^*$. Clearly, a self-adjoint operator is symmetric. The converse is not necessarily true.

Theorem 4.2 **a)** *If A is symmetric then A is closable and $\bar{A}$ is also symmetric.*

b) *If $B \supseteq A$ is a symmetric extension of the symmetric operator A then $B \subseteq A^*$.*

c) *If A is self-adjoint then A is closed, and there are no non-trivial symmetric extensions.*

d) *$A : \mathcal{D}_A \longrightarrow X$ is symmetric if and only if $\langle Ax, x\rangle$ is real for all $x \in \mathcal{D}_A$.*

e) *If $\mathcal{D}_A = X$ and A is symmetric then A is self-adjoint and bounded.* □

Theorem 4.3 *(Main criterion for self-adjointness.) Let $A : \mathcal{D}_A \longrightarrow X$ be symmetric. Then the following conditions are equivalent*

a) *A is self-adjoint.*

b) *A is closed and $\ker(A^* \pm i) = \{0\}$.*

c) $\mathrm{Im}\,(A \pm i) = X$. □

Assume $A : \mathcal{D}_A \longrightarrow X$ symmetric. A is called essentially self-adjoint if the closure $\bar{A}$ is self-adjoint.

Theorem 4.4 *Assume $A : \mathcal{D}_A \longrightarrow X$ symmetric. Then the following conditions are equivalent.*

a) *A is essentially self-adjoint.*

b) $\ker(A^* \pm i) = \{0\}$.

c) $\mathrm{Im}\,(A \pm i)$ *is dense.* □

Remark 4.5 Theorems 4.3, 4.4 remain true if we replace $i \in \mathbb{C}$ by a complex number $\lambda \in \mathbb{C}$ with $\Im\lambda \neq 0$. □

We infer from theorem 4.2 that $A \subseteq A^{**} \subseteq A^*$ for any symmetric A, and $A = A^{**} \subseteq A^*$ for any closed symmetric A. Moreover, we conclude from the definition $A = A^{**} = A^*$ for any self-adjoint operator A. Hence a closed symmetric operator A is self-adjoint if and only if A^* is symmetric. A key role in all the proofs is played by the following

Lemma 4.6 *Let $A : \mathcal{D}_A \longrightarrow X$ be symmetric, $\lambda \in \mathbb{C}$. Then $(A - \lambda\, id)^* = A^* - \bar{\lambda}\, id$, $X = \overline{\mathrm{Im}\,(A - \lambda\, id)} \oplus \ker(A^* - \bar{\lambda}\, id)$ and $|Ax - \lambda x| \geq |\Im\lambda| \cdot |x|$, $x \in \mathcal{D}_A$.* □

The following criterion due to T. Kato produces many self-adjoint operators via perturbation theory.

Theorem 4.7 *Let $A : \mathcal{D}_A \longrightarrow X$ be self-adjoint, $B : D_B \longrightarrow X$ symmetric, $D_B \supseteq \mathcal{D}_A$. Assume for $0 \leq \delta < 1$ and $c \geq 0$ that*

$$|Bx| \leq \delta |Ax| + c|x| \text{ for all } x \in \mathcal{D}_A.$$

Then $A + B$, with $D_{A+B} = \mathcal{D}_A$, is self-adjoint. □

A particular important class of symmetric operators are the semi-bounded ones. $A : \mathcal{D}_A \longrightarrow X$ is called semi-bounded (or bounded from below) if there exists $c \in \mathbb{R}$ such that $\langle Ax, x\rangle \geq c \cdot |x|^2$ for all $x \in \mathcal{D}_A$. This inequality implies in particular that $\langle Ax, x\rangle$ is real, hence A is symmetric. If A is bounded from below then $-A$ is bounded from above, $\langle -Ax, x\rangle \leq -c|x|^2$ and vice versa. Hence we need to consider only operators bounded from below. Semi-bounded operators have a particular convenient self-adjoint extension due to K. O. Friedrichs, the so-called Friedrichs' extension. Assume $\langle Ax, x\rangle \geq c \cdot |x|^2$, choose $\lambda \in \mathbb{R}$ such that $\lambda + c > 0$ and set $\langle x, y\rangle_\lambda := \langle Ax, y\rangle + \lambda\langle x, y\rangle$, $|x|_\lambda = \langle x, x\rangle_\lambda^{\frac{1}{2}}$. Let X_λ be the completion of $\mathcal{D}_A$ with respect to $|\ \ |_\lambda$. Then X_λ becomes a Hilbert space with respect to the extended scalar product. If $\mu \in \mathbb{R}$ is such that $\mu + c > 0$ then X_λ, X_μ are equivalent as Hilbert spaces. For this reason one writes $X_\lambda = X_{E,A}$ and calls $X_{E,A}$ the energy space and $|\ \ |_\lambda$ an energy norm with respect to A.

Theorem 4.8 *Assume $A : \mathcal{D}_A \longrightarrow X$ semi-bounded, $\langle Ax, x\rangle \geq c \cdot |x|^2$. Then $A_F x := A^* x$, $D_{A_F} := X_{E,A} \cap D_{A^*}$, is a self-adjoint extension of A and $\langle A_F x, x\rangle \geq c \cdot |x|^2$, $x \in D_{A_F}$. A_F is called Friedrichs' extension. For the energy space, this equivalence of Hilbert spaces holds: $X_{E,A} = X_{E,\bar{A}} = X_{E,A_F}$.* □

From 4.8 one easily infers

Theorem 4.9 *Let $A : \mathcal{D}_A \longrightarrow X$ be densely defined and closed. Then A^*A, $D_{A^*A} = \{x \in \mathcal{D}_A | Ax \in D_{A^*}\}$, is densely defined and self-adjoint.* □

The type of the spectrum of generalized Dirac operators on open manifolds sees a considerable part of the underlying geometry and assumptions on the spectral type dis/allow many analytical constructions. For this reason we list the most important notions of spectral theory.

Let $A : \mathcal{D}_A \longrightarrow X$ be densely defined and closed. Then $\varrho(A) = \{\lambda \in \mathbb{C} |\ R_\lambda := (A - \lambda id)^{-1}$ is defined and belongs to $L(X)\}$ is called the resolvent set of A, while its complement $\sigma(A) := \mathbb{C} \setminus \varrho(A)$ is called the spectrum of A. $\varrho(A)$ is open, $\sigma(A)$ is closed. $\sigma_p(A) := \{\lambda\, | \exists x \in \mathcal{D}_A,\ |x| = 1,\ Ax = \lambda x\}$ is called the point spectrum and $\sigma_{c,R}(A) := \{\lambda | (A - \lambda id)^{-1}$ exists but does not belong to $L(X)\}$ is called the resolvent continuous spectrum. Below we define the continuous spectrum $\sigma_c(A)$. $\sigma_{c,R}(A)$ and $\sigma_c(A)$ need not coincide. A bounded sequence $(x_i)_i$ in $\mathcal{D}_A$ is called a Weyl sequence for λ if the set $\{x_i | i = 1, 2, \dots\}$ is not pre-compact and $\lim_{i\to\infty}(Ax_i - \lambda x_i) = 0$. $\sigma_e(A) = \{\lambda\,|$ there exists a Weyl sequence for $\lambda\}$ is called the essential spectrum. Denote $\sigma_{p,f}(A) = \{\lambda \in \sigma_p(A) | \mathrm{mult}(\lambda)$ is finite$\}$, where $\mathrm{mult}(\lambda)$ is the multiplicity of $\lambda = \dim\ker(A - \lambda I)$.

Theorem 4.10 *Let $A : \mathcal{D}_A \longrightarrow X$ be a self-adjoint operator.*

a) *$\sigma(A)$ is a subset of the real numbers. Furthermore*

$$|(A - \lambda\, id)^{-1}| \leq \frac{1}{|\Im\lambda|} \quad \text{for } \Im\lambda \neq 0.$$

b) *If $\lambda \in \sigma_{c,R}(A)$ then $\mathrm{Im}\,(A - \lambda id)$ is dense in X.*
Moreover, there exists a Weyl sequence for λ, so $\lambda \in \sigma_e(A)$.

c) *$\lambda \in \sigma_p(A)$ if and only if $\overline{\mathrm{Im}\,(A - \lambda id)}$ is a proper subspace of X.*

d) *$\lambda \in \sigma_{c,R}(A)$ if and only if $\overline{\mathrm{Im}\,(A - \lambda id)} = X$ and $\mathrm{Im}\,(A - \lambda id)$ is properly contained in X.*

e) *If $\lambda_1, \lambda_2 \in \sigma_p(A)$, $\lambda_1 \neq \lambda_2$, $Ax_1 = \lambda_1 x_1$, $Ax_2 = \lambda_2 x_2$, then $\langle x_1, x_2\rangle = 0$.* □

Theorem 4.11 *Let $A : \mathcal{D}_A \longrightarrow X$ be self-adjoint, $\lambda \in \mathbb{C}$. Then one of the following cases occurs:*

a) $\mathrm{Im}\,(A - \lambda id) = X$. *Then $\lambda \in \varrho(A)$.*

b) $\mathrm{Im}\,(A - \lambda id)$ *is properly contained in X, $\overline{\mathrm{Im}\,(A - \lambda id)} = X$. Then $\lambda \in \sigma_e(A)$, $\lambda \notin \sigma_{p,f}(A)$.*

c) $\mathrm{Im}\,(A - \lambda\,\mathrm{id}) = \overline{\mathrm{Im}\,(A - \lambda\,\mathrm{id})}$ *is properly contained in X and $X \ominus \overline{\mathrm{Im}\,(A - \lambda\,\mathrm{id})}$ is finite-dimensional. Then $\lambda \in \sigma_{p,f}(A)$ and $\lambda \notin \sigma_e(A)$.*

d) $\operatorname{Im}(A-\lambda\,\mathrm{id})$ *is properly contained in* $\overline{\operatorname{Im}(A-\lambda\,\mathrm{id})}$ *and* $\overline{\operatorname{Im}(A-\lambda\,\mathrm{id})}$ *is properly contained in* X. *If* $X \ominus \overline{\operatorname{Im}(A-\lambda\,\mathrm{id})}$ *is finite-dimensional, then* $\lambda \in \sigma_{p,f}(A) \cap \sigma_e(A)$.

e) $X \ominus \overline{\operatorname{Im}(A-\lambda\,\mathrm{id})}$ *is infinite-dimensional. Then* $\lambda \in \sigma_e(A)$ *and* $\lambda \notin \sigma_{p,f}(A)$. □

In theorems 4.10, 4.11 the subsets of $\sigma(A)$ are characterized by the mapping properties of $A - \lambda\,\mathrm{id}$. Other subsets can be defined by restriction of A to certain spectral subspaces of X. Since $A : \mathcal{D}_A \longrightarrow X$ is self-adjoint, there exists a spectral family $\{E_\mu\}_{\mu\in\mathbb{R}}$ of orthogonal projections such that $\lambda \le \mu$ implies $E_\lambda \le E_\mu$, $\lim\limits_{\mu\to\infty} E_\mu x = x$, $\lim\limits_{\mu\to-\infty} E_\mu x = 0$ for all $x \in X$, $\lim\limits_{\lambda\to\mu^+} E_\lambda = E_\mu$, and there is a spectral measure dE_μ such that

$$A = \int_{-\infty}^{+\infty} \lambda \, dE_\lambda.$$

Theorem 4.12 *Let* $A : \mathcal{D}_A \longrightarrow X$ *be self-adjoint,* $A = \int \lambda\, dE_\lambda$. *Then*

a) $\mu \in \sigma(A) \iff E_{\mu+\varepsilon} - E_{\mu-\varepsilon} \neq 0$ *for all* $\varepsilon > 0 \iff E_\mu - E_{\mu^-} \neq 0$.

b) *If* μ *is isolated in* $\sigma(A)$ *then* $\mu \in \sigma_p(A)$.

c) $\mu \in \sigma_e(A)$ *if and only if for every positive* ε $\operatorname{Im}(E_{\mu+\varepsilon} - E_{\mu-\varepsilon})$ *is infinite-dimensional.* □

Every Borel measure m on $\mathbb{R}$ admits a decomposition $m = m_{pp} + m_c = m_{pp} + m_{ac} + m_{sc}$, where m_{pp} is a pure point measure, $m_c = m - m_{pp}$ is characterized by $m_c(\{p\}) = 0$ for all points p, m_{ac} is absolutely continous with respect to the Lebesque measure and m_{sc} is singular with respect to the Lebesgue measure, i.e. $m_{sc}(S) = 0$ for a certain set S such that $\mathbb{R} \setminus S$ has zero Lebesgue measure. We consider for $x \in X$ the positive Borel measure $[a,b] \xrightarrow{m_x} \langle E_{[a,b]}x, x\rangle$ and set $X_{pp} = \{x \in X | m_x = m_{x,pp}\}$, $X_{ac} = \{x \in X | m_x = m_{x,ac}\}$, $X_{sc} = \{x \in X | m_x = m_{x,sc}\}$, $X_c = \{x \in X | m_x = m_{x,c}\}$ and $\sigma_{pp}(A) := \sigma(A|_{X_{pp}})$, $\sigma_c(A) := \sigma(A|_{X_c})$, $\sigma_{ac}(A) = \sigma(A|_{X_{ac}})$, $\sigma_{sc}(A) = \sigma(A|_{X_{sc}})$. Then $\sigma_c(A)$, σ_{ac} and σ_{sc} are called the continuous, the absolutely continuous and the singular continuous spectra. Finally we define $\sigma_{pd}(A) := \{\lambda \in \sigma_p(A) \,|\, \mathrm{mult}(\lambda) < \infty$ and λ is an isolated point in $\sigma(A)\}$, called is the purely discrete spectrum.

Theorem 4.13 *Let* $A : \mathcal{D}_A \longrightarrow X$ *be self-adjoint.*

a) $X = X_{pp} \oplus X_{ac} \oplus X_{sc}$.

b) $\sigma(A) = \sigma_{pd}(A) \cup \sigma_e(A)$, $\sigma_{pd}(A) \cap \sigma_e(A) = \emptyset$.

c) $\sigma_e(A) = \sigma_c(A) \cup \sigma_p(A)^1 \cup \{\lambda \in \sigma_p(A) \,|\, \mathrm{mult}(\lambda) = \infty\}$.

d) $\sigma_c(A) = \sigma_{ac}(A) \cup \sigma_{sc}(A)$.

e) $\sigma(A) = \sigma_{pp}(A) \cup \sigma_c(A) = \overline{\sigma_p(A)} \cup \sigma_c(A)$. □

Here M^1 denotes the set of accumulation points of M.

Remark 4.14 In general, $\sigma_p(A) \subset \sigma_{pp}(A)$, hence $\sigma_p(A) \cup \sigma_c(A) \cup \sigma_{sc}(A) \subset \sigma(A)$, but $\overline{\sigma_p(A)} = \sigma_{pp}(A)$. This shows also that in general $\sigma_c(A) \setminus (\sigma_c(A) \cap \overline{\sigma_p(A)}) \subset \sigma_{c,R}(A)$. □

The main objective of spectral geometry as part of analysis on open manifolds is to find and describe which components of the decompositions 4.13 for a generalized Dirac operator A are empty or not.

A special class of self-adjoint operators are those whose spectrum is purely discrete, i. e. $\sigma_e(A) = \emptyset$, $\sigma(A) = \sigma_{pd}(A)$.

Theorem 4.15 *Let A be an operator with a purely discrete spectrum.*

a) *A is not bounded.*

b) *The eigenvalues of the operator A can be ordered by their absolute value, taking into account multiplicities. If $\{\lambda_j\}_{j=1,2,\ldots}$ are the eigenvalues and $\{x_j\}_{j=1,2,\ldots}$ the corresponding orthonormal eigenvectors then the system $\{x_j\}_{j=1,2,\ldots}$ is complete and $|\lambda_j| \underset{j\to\infty}{\longrightarrow} \infty$,*

$$\mathcal{D}_A = \left\{ x \in X \,\middle|\, \sum_{j=1}^{\infty} \lambda_j^2 |\langle x, x_j\rangle|^2 < \infty \right\} \quad \textit{and} \quad Ax = \sum_{j=1}^{\infty} \lambda_j \langle x, x_j\rangle x_j.$$

□

Theorem 4.16 *Let $A : \mathcal{D}_A \longrightarrow X$ be positive-definite (i. e. $\langle Ax, x\rangle \geq c|x|^2$, $c > 0$, $x \in \mathcal{D}_A$), A_F the Friedrichs' extension, $X_{E,A}$ the corresponding energy space and $A^{\frac{1}{2}} = \int\limits_c^{\infty} \sqrt{\lambda}\, dE_\lambda$, $D_{A^{\frac{1}{2}}} = \{x \in X \mid \int\limits_c^{\infty} \lambda\, d|E_\lambda x|^2 < \infty\}$. Then*

$$D_{A_F^{\frac{1}{2}}} = X_{E,A} \quad \textit{and} \quad |x|_{X_{E,A}} = |A_F^{\frac{1}{2}} x|.$$

□

Theorem 4.17 *Let $A : \mathcal{D}_A \longrightarrow X$ be positive definite with purely discrete spectrum and $X_{E,A}$ the corresponding energy space.*

a) *If $0 < \lambda_1 \leq \lambda_2 \leq \cdots$ are the ordered eigenvalues of A (with multiplicities) and $\{x_j\}_{j=1,2,\ldots}$ is a corresponding system of orthonormal eigenvectors then*

$$X_{E,A} = \left\{ x \in X \,\middle|\, \sum_{j=1}^{\infty} \lambda_j |\langle x, x_j\rangle|^2 < \infty \right\}.$$

$\{x_j \cdot \lambda_j^{\frac{1}{2}}\}_{j=1,2,\ldots}$ is a complete orthonormal system in $X_{E,A}$. $A^{\frac{1}{2}}$ provides a unitary mapping of $X_{E,A}$ onto X.

b) *The operators A^{-1} and $A^{-\frac{1}{2}} = (A^{\frac{1}{2}})^{-1}$ in $L(X)$ are compact. If A^{-1} is regarded as mapping of X into $X_{E,A}$, then it is compact too.* □

The purely discrete spectrum property of self-adjoint elliptic operators on closed manifolds is an immediate consequence of Rellich's criterion.

Theorem 4.18 *(Rellich's criterion). A self-adjoint positive-definite operator A has purely discrete spectrum if and only if the embedding $X_{E,A} \longrightarrow X$ is compact.* □

It is spectral theory's aim to compute or estimate the spectrum. There are not too many standard methods available, and the majority is related to the minimax formula.

Theorem 4.19 *Let $A : \mathcal{D}_A \longrightarrow X$ be self-adjoint and semi-bounded, $A \geq c \cdot \mathrm{id}$. Let*

$$\lambda_n(A) := \sup_{x_1,\ldots,x_{n-1}} \inf_{\substack{y \in \mathcal{D}_A, |y|=1 \\ y \in L(x_1,\ldots,x_{n-1})^{\perp}}} \langle Ay, y\rangle.$$

Then for fixed n either

a) *there exists n eigenvalues smaller than the bottom of the essential spectrum (taking into account their mulitplicity) and $\lambda_n(A)$ is the n-th eigenvalue, or*

b) *$\lambda_n(A)$ is the bottom of the essential spectrum. In this case $\lambda_n = \lambda_{n+1} = \lambda_{n+2} = \cdots$ and there exists $n-1$ eigenvalues below λ_n.* □

Let A, B be self-adjoint operators in X, $X = X_{pp}(B) \oplus X_{ac}(B) \oplus X_{sc}(B)$ the decomposition of X of 4.13 a), $P_{ac}(B) : X \longrightarrow X_{ac}(B)$ the projector. Assume that the so-called wave operators

$$W_{\pm}(A, B) := s - \lim_{t \to \pm\infty} e^{-iAt} e^{iBt} P_{ac}(B)$$

are defined and that im $W_{\pm}(A, B) = X_{ac}(A)$. Then $W_{\pm}$ are called complete. In this case,

$$AW_{\pm}(A, B) = W_{\pm}(A, B)B.$$

$S = W_+^* W_-$ is called the scattering matrix and plays a decisive role in scattering theory.

We obtain

Corollary 4.20 *If $W_+(A, B)$ exists and is complete, then $A|_{X_{ac}(A)}$ and $B|_{X_{ac}(B)}$ are unitarily equivalent, so in particular $\sigma_{ac}(A) = \sigma_{ac}(B)$.* □

Sufficient criteria for this hypothesis are given by

Theorem 4.21 *Let A, B be self-adjoint.*

a) *If $e^{-tA^2} - e^{-tB^2}$ is of trace class for one $t > 0$, then $W_{\pm}(A^2, B^2)$ exist and are complete.*

b) *If $e^{-tA^2} A - e^{-tB^2} B$ are of trace class for all $t > 0$, then $W_{\pm}(A, B)$ exist and are complete.*

□

Here $A \in L(X)$ is of trace class if $\operatorname{tr}|A| < \infty$, by writing $A = U|A|$, U a partial isometry, in the polar decomposition and $|A|_1 := \operatorname{tr}|A| = \sum_{j=1}^{\infty} \langle |A| x_j, x_j \rangle$, with $\{x\}_{j=1,2,\ldots}$ an arbitrary complete orthonormal system. If A is of trace class one defines $\operatorname{tr} A = \sum_{j=1}^{\infty} \langle A x_j, x_j \rangle$, a convergent series. $A \in L(X)$ is called Hilbert-Schmidt operator if $\operatorname{tr} A^* A < \infty$. Then $|A|_2 := (\operatorname{tr} A^* A)^{\frac{1}{2}}$.

Theorem 4.22 **a)** *A is of trace class if and only if $A = B \cdot C$, B and C Hilbert-Schmidt. Then $|A|_1 \le |B|_2 \cdot |C|_2$.*

b) *An integral operator with square integrable kernel is Hilbert-Schmidt.* □

Now we apply these notions to differential operators on manifolds. Their natural initial domain is C_c^{∞}. We ask, under which assumptions are they self-adjoint? The key is given by

Lemma 4.23 *Let $A : \mathcal{D}_A \longrightarrow X$ be a densely defined symmetric operator.*

a) *If there exists a unitary 1-parameter group U_t s.t. $U_t \mathcal{D}_A \subseteq \mathcal{D}_A$, $U_t A = A U_t$ and $\frac{d}{dt} U_t x = i A U_t x$, $x \in \mathcal{D}_A$, then every power A^n, $n \ge 0$, is essentially self-adjoint.*

b) *If A has equal deficiency indices $\kappa_{\pm} = \dim\ker(A^* \pm i)$ and both equations*

$$\frac{du}{dt} \pm (iA)^* u = 0, \quad 0 \le t < \infty,$$

have a unique solution of the Cauchy problem then A is essentially self-adjoint.

c) *If A is semibounded from above and the Cauchy problem for the equation*

$$\frac{du}{dt} - A^*u = 0 \quad or \quad \frac{d^2u}{dt^2} - A^*u = 0$$

has a unique solution then A is essentially self-adjoint. □

Proposition 4.24 **a)** *Let $(E, h, \nabla, \cdot) \longrightarrow (M^n, g)$ be a Clifford bundle, (M^n, g) complete and D the generalized Dirac operator. Then all powers D^n, $n \geq 0$, are essentially self-adjoint.*

b) *Let (M^n, g) be complete, $\Delta = \Delta_q$ be the Laplace operator acting on q-forms with values in a flat vector bundle E, $V = V(x)$ an endomorphism of $\Lambda^q T^*M \otimes \mathbb{C}$, $V(x) \geq Q(x)$, $1 \leq Q(x) \leq \infty$, $Q^{-\frac{1}{2}}(x)$ a Lipschitz function on M such that $|Q^{-\frac{1}{2}}(x) - Q^{-\frac{1}{2}}(y)| \leq \kappa \cdot \mathrm{dist}(x, y)$ for all $x, y \in M$. If for any piecewise smooth curve $\gamma : [0, \infty[\longrightarrow M$ such that $\lim_{t \to \infty} \mathrm{dist}(p, \gamma(t)) = \infty$ the integral*

$$\int_\gamma Q^{-\frac{1}{2}}(x)\, d\gamma = \infty$$

then the operator $H_0 = (\Delta_q + V)|_{C_c^\infty}$ is essentially self-adjoint. □

Corollary 4.25 *Let $(E, h, \nabla) \longrightarrow (M^n, g)$ be a Riemannian vector bundle, (M^n, g) complete and Δ_q the Laplace operator acting on q-forms with values in E. Then $(\Delta_q)^n$, $n = 1, 2, \ldots$ are essentially self-adjoint. In particular this holds for the Laplace operators acting on ordinary q-forms.* □

For non-complete manifolds, there exists at least one self-adjoint extension, namely Friedrichs' extension. But there can be many others. We will discuss them in section 4.

In any case we denote in the sequel by $\Delta = \Delta_q$ a self-adjoint extension (which is unique if (M^n, g) is complete). We start with the simplest decomposition

$$\sigma(\Delta_q) = \sigma_e(\Delta_q) \cup \sigma_{pd}(\Delta_q), \quad \sigma_e(\Delta_q) \cap \sigma_{pd}(\Delta_q) = \emptyset.$$

If (M^n, g) is closed then $\sigma_e(\Delta_q)$ ist empty: We know from theorem 4.18 that $\sigma_e(\Delta_q) = \emptyset$ if and only if the embedding $(L_2(\Lambda^q))_{E,\Delta_q} \hookrightarrow L_2(\Lambda^q) \equiv \Omega^{q,2,0}$ is compact. But $(L_2(\Lambda^q))_{E,\Delta_q} = \Omega^{q,2,1}$ (for (M^n, g) closed) and the compactness of the embedding $\Omega^{q,2,1} \hookrightarrow \Omega^{q,2,0} = L_2$ is a well known fact and can easily be proven, $\{\Omega^{q,2,i}\}_i$ is a Rellich chain on closed manifolds. For M^n open, the embedding $(L_2(\Lambda^q))_{E,\Delta_q} \hookrightarrow L_2(\Lambda^q) = \Omega^{2,q,0}$ is far from being compact, at least if M^n has "enough space". Then there exist sequences of sections with "nearly disjoint" support, bounded together with their derivatives which have in L_2 no convergent subsequence (since the supports are "nearly disjoint"). Such sequences often constitute Weyl sequences and converse, i. e. if there is "enough space" then should be essential spectrum. We will support this by a series of exact propositions.

Proposition 4.26 *Let $(E, h, \nabla^h, \cdot) \longrightarrow (M^n, g)$ be a Clifford bundle, M^n open and complete, $K \subset M$ a compact subset, $D_F(E|_{M \setminus K})$ Friedrichs' extension of $D|_{C_c^\infty(E|_{M\setminus K})}$. Then there hold*

$$\sigma_e(D) = \sigma_e(D_F) = \sigma_e(D_F(E|_{M\setminus K}))$$

and

$$\sigma_e(D^2) = \sigma_e(D_F^2) = \sigma_e((D_F(E|_{M\setminus K}))^2).$$

Corollary 4.27 *The essential spectrum of D and D^2 remains invariant under compact perturbations of the topology and the metric. In particular this holds for the Laplace operators acting on forms with values in a vector bundle.* □

As for compact manifolds, we can define the Riemannian connected sum for open Riemannian manifolds, even for Riemannian vector bundles $(E_i, h_i, \nabla^{h_i}) \longrightarrow (M_i^n, g_i)$, where at the compact glueing domain the metric and connection are not uniquely determined. Another corollary is then given by

Proposition 4.28 *Suppose $(E_i, h_i, \nabla^{h_i}) \longrightarrow (M_i^n, g_i)$, $i = 1, \ldots, r$ Riemannian vector bundles of the same rank, (M_i^n, g_i) complete, and let $\Delta = \Delta_q$ be the Laplace operator acting on q-forms with values in E_i (resp. E). Then*

$$\sigma_e\Delta_q\Big(\mathop{\#}_{i=1}^{r}(E_i \longrightarrow M_i)\Big) = \bigcup_{i=1}^{r}\sigma_e(\Delta_q(E_i \longrightarrow M_i)).$$

□

4.26 can be reformulated as the statement that the essential spectrum for an isolated end ε is well defined. We denote it by $\sigma_e(D_F(\varepsilon))$, $\sigma_e(D_F^2(\varepsilon))$.

Proposition 4.29 *If (M^n, g) is complete and has finitely many ends $\varepsilon_1, \ldots, \varepsilon_r$ then*

$$\sigma_e(D) = \bigcup_{i=1}^{r}\sigma_e(D_F(\varepsilon_i)), \quad \sigma_e(D^2) = \bigcup_{i=1}^{r}\sigma_e(D_F^2(\varepsilon_i)).$$

Proposition 4.30 *Assume the hypothesis of 4.26. Suppose $\lambda \in \sigma_e(D)$. Then there exists a Weyl sequence $(\varphi_\nu)_\nu$ for λ such that for any compact subset $K \subset M$*

$$\lim_{\nu\to\infty}|\varphi_\nu|_{L_2(E|_K)} = 0. \tag{4.1}$$

For every $\lambda \in \sigma_e(D^2)$ there exists a Weyl sequence $(\varphi_\nu)_\nu$ satisfying (4.1) and

$$\lim_{\nu\to\infty}|D\varphi_\nu|_{L_2(E|_K)} = 0.$$

4.30 means that w. l. o. g. Weyl sequences should "leave" (in the sense of the L_2-norm) any compact subset, i. e. there must be "place enough at infinity".

Proposition 4.31 *Let $(E, h, \nabla, \cdot) \longrightarrow (M^n, g)$ be a Clifford bundle with (I), $(B_{r-3}(M, g))$, $(B_{r-3}(E, \nabla))$, $r > \frac{n}{2} + 1$ and ∇' a second Clifford connection satisfying $|\nabla' - \nabla|_{\nabla,2,r-1} < \infty$. Then for $D = D(\nabla)$ and $D' = D(\nabla')$ there holds*

$$\mathcal{D}_D = \mathcal{D}_{D'} \tag{4.2}$$

and

$$\sigma_e(D) = \sigma_e(D'). \tag{4.3}$$

Corollary 4.32 *Assume the hypothesis of 4.31. Then for all* $i \in \mathbb{N}$

$$\mathcal{D}_{D^i} = \mathcal{D}_{D'^i} \tag{4.4}$$

and

$$\sigma_e(D^i) = \sigma_e({D'}^i). \tag{4.5}$$

Remark 4.33 For $\Omega^{2,1}(E,\nabla) = \Omega^{2,1}(E,D)$ (B_0) would be sufficient which easily follows from the Weitzenboeck formula, but we need moreover the Sobolev embedding theorem for $\Omega^{2,r-1}$, in particular $\overset{\circ}{\Omega}{}^{2,r-1} = \Omega^{2,r-1}$, hence (B_{r-3}). One could try to estimate $||\eta|_x \cdot |\varphi|_x|_{L_2}$ instead of $\sup\limits_x |\eta|_x \cdot |\varphi|_{L_2}$, i. e. one could try to work without the Sobolev embedding theorem, but in this case we would need some module structure. Hence an assumption (B_i), $i > 0$, seems to be inavoidable in any case. □

For the Laplace operator on forms, 4.31 is not immediately applicable since if we replace for (M^n, g) the Levi-Civita connection ∇^g by another metric connection then we loose many of the standard formulas, i. e. we should consider a change $g \longrightarrow g'$.

Proposition 4.34 *Let* (M^n, g) *be open, complete, with* (I) *and* $(B_r(M^n, g))$, $r > \frac{n}{2} + 1$, g' *another metric satisfying the same conditions and suppose* g, g' *quasi isometric and* $|g' - g|_{g,2,r} = (\int (|g' - g|^2_{g,x} + \sum\limits_{i=0}^{r-1} |(\nabla^g)^i - \nabla^g|^2_{g,x}) \, \mathrm{dvol}_x(g))^{\frac{1}{2}} < \infty$. *Then*

$$\mathcal{D}_{\Delta_q(g)} = \mathcal{D}_{\Delta_q(g')} \text{ as equivalent Hilbert spaces}$$

and

$$\sigma_e(\Delta_q(g)) = \sigma_e(\Delta_q(g')), \quad q = 0, 1, \ldots, n.$$

We state the generalization to forms with values in a vector bundle.

Proposition 4.35 *Let* $(E, h, \nabla^h) \longrightarrow (M^n, g)$ *be a Riemannian vector bundle satisfying* (I), $(B_k(M^n, g))$, $(B_k(E, \nabla))$, $k \geq r > \frac{n}{2} + 1$, *and let* g' *be a second metric,* h' *a second fibre metric with metric connection* $\nabla^{h'}$, g, g' *and* h, h' *quasi isometric, respectively,*

$$|g - g'|_{g,2,r} < \infty, \quad |h - h'|_{h,\nabla^h,g,2,r} < \infty, \quad |\nabla^h - \nabla^{h'}|_{h,\nabla^h,g,2,r-1} < \infty.$$

$(E, h', \nabla^{h'}) \longrightarrow (M^n, g')$ *also satisfying* (I) *and* (B_k). *Then there holds for the Laplace operators* $\Delta = \Delta_q(g, h, \nabla^h)$, $\Delta' = \Delta_q(g', h', \nabla^{h'})$ *acting on forms with values in* E

$$\mathcal{D}_\Delta = \mathcal{D}_{\Delta'} \text{ as equivalent Hilbert spaces} \tag{4.6}$$

and

$$\sigma_e(\Delta) = \sigma_e(\Delta'). \tag{4.7}$$

Proposition 4.36 *Let* (M^n, g) *be open, complete with noncompact isometry group* $\mathfrak{I}(M)$. *Then* $\sigma(\Delta_q)$ *contains no eigenvalues of finite multiplicity,* $0 \leq q \leq n$.

Proof From $\mathfrak{I}(M)$ noncompact one gets the following. For every $x_0 \in M$ and for every $r > 0$ and $B_r(x_0)$ there exist a sequence $(x_i)_{i\geq 0}$ and a sequence $(f_i)_{i\geq 0}$ of isometries of M such that f_i maps $B_i = B_r(x_i)$ isometrically onto $B_0 = B_r(x_0)$ and the $B_r(x_i)$ are pairwise disjoint. Assume $\lambda \in \sigma_p(\Delta_q)$ and $\text{mult}(\lambda) = m$. Let u be a normalized eigenform for λ, $\Delta u = \lambda u$, $||u|| = 1$. Then also $u_j = f_j^* u$ is an eigenform, $\Delta f_j^* u = f_j^* \Delta u = \lambda f_j^* u$. Thus one gets a sequence $(u_j)_j$ at the unit sphere in the eigenspace which has an accumulation point u^*. Without loss of generality we assume that $(u_j)_j$ converges to u^*. We denote $||u^*||_i^2 = ||u^*|_{B_i}||^2 = \int\limits_{B_i} u^* \wedge *u$. Then there holds $1 = ||u^*||^2 \geq \sum\limits_i ||u^*||_i^2$, in particular $||u^*||_i \underset{i\to\infty}{\longrightarrow} 0$. For every $\varepsilon > 0$ there exists an i_0 such that $||u^*||_i < \frac{\varepsilon}{2}$ for all $i \geq i_0$. This implies $||u_j||_i < \frac{\varepsilon}{2}$ for all $j > j_0$ and $i > i_0$, in particular $||u_k||_k < \varepsilon$, $k > \max\{i_0, j_0\}$. But $||u_k||_k = ||u|_{B_0}|| < \varepsilon$. ε was arbitrary, $u|_{B_0}$ has to be zero. According to the unique continuation theorem of Aronszajn (cf. [5]) u has to be zero which contradicts to our assumption. The multiplicity of λ has to be infinite. □

Corollary 4.37 *For the above manifolds there holds*

$$\sigma(\Delta_q) = \sigma_e(\Delta_q), \quad 0 \leq q \leq n.$$

In particualar, for symmetric spaces of noncompact type the Δ_qs don't have eigenvalues of finite multiplicity. □

These results suggest that if we have "space enough" then σ_e should be $\neq \emptyset$. But this philosophy is still to rough as the following example shows. There exists a Dirac operator D on ($\mathbb{R}^n$, metric) such that $\sigma_e(D) = \emptyset$.

Corollary 4.38 *Let (M^n, g) be open, complete, spin, $S \longrightarrow M$ a spinor bundle with spinor connection coming from lifting the Levi-Civita connection s. t. D is symmetric. Assume for the scalar curvature τ of (M^n, g)*

$$\lim_{r\to\infty} \inf\{\tau(x)|d(x, x_0) \geq r\} = \infty.$$

Then the Dirac operator D is an operator with a purely discrete point spectrum.

Let (M^n, g) be closed. Then the spectrum of any self-adjoint elliptic operator is purely discrete, as we pointed out in section 3. Now the task of spectral geometry consists in the task to calculate/estimate the spectrum of Δ, D, D^2 and to draw conclusions from this. We refer to the contribution of Peter Gilkey in this volume.

If (M^n, g) is open, then the spectral theory is more difficult. First one wants to determine the spectral type of the Laplace operators Δ_q, where the spectral type – roughly speaking – is the statement which components in the decompositions

$$\begin{aligned} \sigma(\Delta_q) &= \sigma_{pd}(\Delta_q) \cup \sigma_e(\Delta_q) \\ &= \sigma_{pd}(\Delta_q) \cup \sigma_c(\Delta_q) \cup \sigma_p(\Delta_q)^1 \cup \{\lambda \in \sigma_p(\Delta_q) \,|\, \text{mult}(\lambda) = \infty\} \\ &= \sigma_{pd}(\Delta_q) \cup \sigma_{ac}(\Delta_q) \cup \sigma_{sc}(\Delta_q) \cup \sigma_p(\Delta_q)^1 \cup \{\lambda \in \sigma_p(\Delta_q) \,|\, \text{mult}(\lambda) = \infty\} \end{aligned}$$

are empty or non-empty, respectively. More desirable it would be even to calculate the spectrum or some bounds or some of the spectral values. As well known, already for Euclidean domains this is a rather difficult task. Nevertheless, for some special manifolds as e. g. hyperbolic space and other symmetric spaces there are complete exact calculations.

Of particular interest are the bottom of the spectrum, the point spectrum, the essential spectrum and the spectral value 0. Concerning the spectral type, we refer to many papers of Donnelly, Dodziuk (e.g. [32], [34], [35]), the author (e.g. [45]), other colleagues and to [43].

5 The spectral value zero

In many investigations, assumptions and applications, the spectral value zero plays a particular role. We will discuss this carefully. The background for this is that the spectral value zero is strongly connected with (reduced and non-reduced) L_2-cohomology which reflects also some topological features. The first natural question would be: Does zero always belong to at least one $\sigma(\Delta_q)$, $0 \le q \le n$? The zero in the spectrum conjecture (cf. [168]) says that the answer will be yes for complete Riemannian manifolds (M^n, g). We come back to this question after introducing L_2-cohomology.

For our calculations in the sequel we need a special class of Sobolev spaces. Consider $\Omega^q(E)$, the operators $d = d_q : \Omega^q(E) \longrightarrow \Omega^{q+1}(E)$ and $\delta = \delta_q : \Omega^q(E) \longrightarrow \Omega^{q-1}(E)$, $q = 0, \ldots, n$. Write $d_{q,0} = d_q|_{C_c^\infty}$, $\delta_{q,0} = \delta_q|_{C_c^\infty}$. For a finite set S of polynomials in d, δ with constant coefficients we define

$$\Omega_S^{q,p}(E) = \left\{ \varphi \in \Omega^q(E) \,|\, |D\varphi|_p = \left(\int |D\varphi|_x^p \,\mathrm{dvol}_x(g) \right)^{\frac{1}{p}} < \infty \text{ for all } D \in S \right\}$$

and

$$\begin{aligned} \overline{\Omega}^{q,p,S}(E) &= \text{completion of } \Omega_S^{q,p}(E) \text{ with respect to } |\;|_{p,S} \\ |\varphi|_{p,S} &= |\varphi|_p + \sum_{D\in S} |D\varphi|_p . \end{aligned}$$

By $\mathring{\Omega}^{q,p,S}(E)$ we denote the completion of $C_c^\infty(\Lambda^q(E))$ with respect to $|\;|_{p,S}$ and by $\Omega^{q,p,S}(E)$ the space of all regular distributions φ such that $|\varphi|_{p,S} < \infty$.

Examples 5.1

1) $S = \{d_{q,0}\}$, $\mathring{\Omega}^{q,2,S}(E) = \mathring{\Omega}^{q,2,\{d_{q,0}\}}(E) = D_{\overline{d}_{q,0}}$,

$S = \{\delta_{q,0}\}$, $\mathring{\Omega}^{q,2,\{\delta_{q,0}\}}(E) = D_{\overline{\delta}_{q,0}}$.

2) $S = \{d_q|_{\Omega^{q,2}_{\{d\}}}\}$, $\Omega^{q,2,S}(E) = D_{\overline{d}_q}$,

$S = \{\delta|_{\Omega^{q,2}_{\{\delta\}}}\}$, $\overline{\Omega}^{q,2,\{S\}}(E) = D_{\overline{\delta}_q}$.

3) $S = \{\Delta_q\} = \{d_{q-1}\delta_q + \delta_{q+1} d_q|_{\Omega^{q,2}_\Delta}\}$, $\Omega^{q,2,\{\Delta\}}(E) = D_{\overline{\Delta}}$,

$S = \{\Delta|_{C_c^\infty(\Lambda^q(E))}\}$, $\overline{\Omega}^{q,2,S}(E) = D_{\overline{\Delta}_{q,0}}$.

4) $S = \{(1+\Delta)^k|_{C_c^\infty(\Lambda^q(E))}\}$, $\Omega^{q,2,S}(E) = \mathring{\Omega}^{q,2,2k}(E, \Delta)$.

If $S = \{d\}$ then we sometimes simply write $\Omega^{q,p,d}$ instead of $\Omega^{q,p,\{d\}}$. □

This section is built up as follows. First we introduce L_p- and several versions of L_2-cohomology and connect this with the spectral value zero. Thereafter we discuss the zero in the spectrum conjecture and its solution. Finally we present classes of manifolds from real and complex differential geometry and algebraic geometry for which L_2-cohomology can be calculated or estimated. The methods doing this are really very different. The first part of this paragraph is devoted to L_p-cohomology. We start with the definition of L_p-cohomology and remind for $1 \le p < \infty$ and S a set of polynomials in d and δ with constant coefficients the Sobolev spaces

$$\Omega_S^{q,p}(E), \overline{\Omega}^{q,p,S}(E), \mathring{\Omega}^{q,p,S}(E), \Omega^{q,p,S}(E). \tag{5.1}$$

For $p = \infty$ we define ${}^b\Omega^q_{\{d\}}(E) = \{\varphi \mid \varphi \in \Omega^q(E) \text{ and } \sup\limits_x |\varphi|_x, \sup\limits_x |d\varphi|_x < \infty\}$ which is a normed space w. r. t. ${}^b|\varphi|_d = \sup\limits_x |\varphi|_x + \sup\limits_x |d\varphi|_x$. For $1 \le p < \infty$ the regularized L_p-cohomology $H^{*,p}_{\{d\}}(M^n, g)$ is the cohomology of the complex

$$0 \longrightarrow \Omega^{0,p}_{\{d\}} \longrightarrow \Omega^{1,p}_{\{d\}} \longrightarrow \cdots \longrightarrow \Omega^{q,p}_{\{d\}} \longrightarrow \Omega^{q+1,p}_{\{d\}} \longrightarrow \cdots \longrightarrow \Omega^{n,p}_{\{d\}} \longrightarrow 0, \tag{5.2}$$

i. e.

$$H^{q,p}_{\{d\}}(M^n, g) = \ker(d_q : \Omega^{q,p}_{\{d\}} \longrightarrow \Omega^{q+1,p}_{\{d\}}) / \mathrm{Im}\,(d_{q-1} : \Omega^{q-1,p}_{\{d\}} \longrightarrow \Omega^{q,p}_{\{d\}}) = Z^{q,p}_{\{d\}} / B^{q,p}_{\{d\}}.$$

$H^{*,p}_{\{d\}}$ is the cohomology of the complex $(\Omega^{*,p}_{\{d\}}, d)$. Further we define the reduced regularized L_p-cohomology $\overline{H}^{q,p}_{\{d\}}$ as

$$\overline{H}^{q,p}_{\{d\}} := Z^{q,p}_{\{d\}} / \overline{B}^{q,p}_{\{d\}}$$

and the bounded cohomology ${}^bH^*_{\{d\}}$ as the cohomology of the complex ${}^b\Omega^*_{\{d\}}$. All this can also be done for forms with values in a flat vector bundle E since then at smooth level $d^2 = 0$. Then we get $H^{*,p}_{\{d\}}(M, E)$ etc.

In the sequel we restrict to the case $p = 2$ mainly. Then we have still other canonical complexes, e. g.

$$(\Omega^{*,2}_{c,\{d\}}, d) = (\Omega^*_c, d), \tag{5.3}$$

$$(\mathring{\Omega}^{*,2,\{d\}}, \overline{d}_{*,0}) = (\mathcal{D}_{\overline{d}_{q,0}}, \overline{d}_{q,0})_q, \tag{5.4}$$

$$(\overline{\Omega}^{*,2,\{d\}}, \overline{d}_*) = (\mathcal{D}_{\overline{d}_q}, \overline{d}_q)_q, \tag{5.5}$$

$$(\Omega^{*,2,\{d\}}, d_{*,\max}) = (\mathcal{D}_{d_q,\max}, d_{q,\max}), \tag{5.6}$$

where $d_{q,\max}$ is the maximal extension of d_q, $\mathcal{D}_{d_q,\max} = \{\varphi \in L_2(\Lambda^q T^*) = \Omega^{*,2} \mid d_q\varphi \in L_2,$ where $d_q\varphi$ is taken in the distributional sense.$\}$ First we must see that these are in fact complexes. For (5.3) this is absolutely clear. It yields the (smooth) cohomology with compact support. The complex property of (5.4), (5.5) and (5.6) in fact deserves a proof. We do this for $(\mathcal{D}_{\overline{d}_q}, \overline{d}_q)$. The only point to show is that $\overline{d}_q \overline{d}_{q-1}$ is defined and equals zero. Let $\varphi \in \mathcal{D}_{\overline{d}_{q-1}}$, i. e. there exists a sequence $(\varphi_\nu)_\nu$ in $\Omega^{q-1,2}_{\{d\}}$, such that $(\varphi_\nu)_\nu$ is a Cauchy sequence w. r. t. $|\varphi_\nu|^2_{d_{q-1}} = |\varphi_\nu|^2_{L_2} + |d\varphi_\nu|^2_{L_2}$ and $\varphi_\nu \underset{L_2}{\Longrightarrow} \varphi$, $d\varphi_\nu \underset{L_2}{\Longrightarrow}$

$d\varphi$. $(d\varphi_\nu)_\nu$ is a Cauchy sequence w. r. t. $|d\varphi_\nu|^2_{d_q} = |d_{q-1}\varphi_\nu|^2_{L_2} + |d_q d_{q-1}\varphi_\nu|^2_{L_2} = |d_{q-1}\varphi_\nu|^2_{L_2}$ and $\overline{d}_q d_{q-1}\varphi = \lim_{\nu\to\infty} d_q d_{q-1}\varphi_\nu = 0$. Similarly for the complex $(\mathcal{D}_{\overline{d}_{q,0}}, \overline{d}_{q,0})_q$ and $(\mathcal{D}_{d_{q,\max}}, d_{q,\max})_q$. One would expect that the cohomology of the complexes (5.4) – (5.6) would be distinct. For general elliptic complexes and their distinct closures to Hilbert complexes this will be in fact the case. But fortunately for the de Rham complex the situation is better. Cheeger proved in [24] that the inclusion $\Omega^{*,2}_{\{d\}} \subset \Omega^{*,2,\{d\}}$ is a chain homotopy equivalence between the complexes $(\mathcal{D}_{\overline{d}_q}, \overline{d}_q)_q = (\Omega^{*,2,\{d\}}, \overline{d}_*)$ and $(\Omega^{*,2}_{\{d\}}, d_*)$, i.e. this inclusion induces an isomorphism

$$H^{*,2}_{\{d\}} = H^{*,2,\{d\}} \equiv H^*(\overline{\Omega}^{*,2,\{d\}}, \overline{d_*}) \tag{5.7}$$

Moreover he shows that $d_q\varphi = \eta$ in the distributional sense if and only if $\varphi \in \mathcal{D}_{\overline{d}_q}$ and $\overline{d}_q = \eta$. This has as a consequence

$$H^{*,2,d_{\max}} := H^*(\mathcal{D}_{d_{*,\max}}, d_{*,\max}) = H^*(\mathcal{D}_{\overline{d}_*}, \overline{d}_*) \equiv H^{*,2,\{d\}}. \tag{5.8}$$

We draw the conclusion that $H^{*,2,\{d\}}$ represents the L_2-cohomology of 3 complexes and hence should be of particular importance.

To explore the relations with the complex $(\overset{\circ}{\Omega}{}^{*,2,\{d\}}, \overline{d}_{*,0}) = (D_{\overline{d}_{q,0}}, \overline{d}_{q,0})$ we start with the following

Lemma 5.2 *There holds for all* q

$$\overline{d}_q = \delta^*_{q+1,0}, \tag{5.9}$$
$$\overline{\delta}_{q+1} = d^*_{q,0}. \tag{5.10}$$

Proof Suppose that $\varphi \in \Omega^{q,2}_{\{d\}}$. Then for all $\psi \in \Omega^{q+1}_c$,

$$\langle \delta_{q+1,0}\psi, \varphi\rangle = \langle \psi, d_q\varphi\rangle,$$

which implies $d_q \subset \delta^*_{q+1,0}$, $\overline{d}_q \subseteq \overline{\delta^*}_{q+1,0} = \delta^*_{q+1,0}$, and, in the same manner, $\overline{\delta}_{q+1} \subseteq d^*_{q,0}$. There remains to show $\delta^*_{q+1,0} \subseteq \overline{d}_q$, i.e. if $\varphi \in \mathcal{D}_{\delta^*_{q+1,0}}$ then $\varphi \in \mathcal{D}_{\overline{d}_q}$ and $\overline{d}_q\varphi = \delta^*_{q+1,0}\varphi$. We start with a $\varphi \in \mathcal{D}_{\delta^*_{q+1,0}}$ of class C^1 and show $\varphi \in \mathcal{D}_d$: Assume $\theta \in \Omega^{q+1}_c$, then

$$\langle \varphi, \delta_{q+1,0}, \theta\rangle = \langle \delta^*_{q+1,0}\varphi, \theta\rangle,$$

moreover $\langle d_q\varphi, \theta\rangle$ is well defined and

$$\langle d_q\varphi, \theta\rangle = \langle \varphi, \delta_{q+1,0}\theta\rangle.$$

We obtain $d_q\varphi - \delta^*_{q+1,0}\varphi = 0$ on all compact subsets, which means $d_q\varphi - \delta^*_{q+1,0}\varphi \equiv 0$, $||\delta^*_{q+1,0}\varphi|| < \infty$, $||d_q\varphi|| < \infty$, $\varphi \in \mathcal{D}_{d_q}$. Consider now an arbitrary $\varphi \in \mathcal{D}_{\delta^*_{q+1,0}}$. We have to show that $\varphi \in \mathcal{D}_{\overline{d}_q}$. This would be done if we could show the existence of a sequence $\varphi_\nu \longrightarrow \varphi$, $\varphi_\nu \in \mathcal{D}_{d_q}$, with $d_q\varphi_\nu \longrightarrow \delta^*_{q+1,0}\varphi$. For this we use Friedrich's mollifiers (= smoothing operators) $\mathcal{J}_\varepsilon$ (cf. [123]). If φ is square integrable then $\mathcal{J}_\varepsilon\varphi$ is smooth and $\mathcal{J}_\varepsilon\varphi \underset{\varepsilon\to 0}{\longrightarrow} \varphi$. Furthermore, if $\varphi \in \mathcal{D}_{\delta^*}$ then $d\mathcal{J}_\varepsilon\varphi = \mathcal{J}_\varepsilon\delta^*\varphi$. At the first instance, this holds for ordinary q-forms α. A very simple computation shows for a local form $\alpha \times s$

with values in E again $d\mathcal{J}_\varepsilon(\alpha \times s) = \mathcal{J}_\varepsilon \delta^*(\alpha \times s)$ (the $*$-operator acts only on the usual forms). Setting $\varepsilon = \frac{1}{\nu}$, $\varphi_\nu = \mathcal{J}_{\frac{1}{\nu}}\varphi$, we have

$$\varphi_\nu \in D_{d_q}, \quad \varphi_\nu \longrightarrow \varphi, \quad d\varphi_\nu = d\mathcal{J}_{\frac{1}{\nu}}\varphi = \mathcal{J}_{\frac{1}{\nu}}\delta^*\varphi \longrightarrow \delta^*\varphi.$$

Altogether, we obtained $\overline{d}_q = \delta^*_{q+1,0}$. In a similar manner one shows $d^*_{q,0} = \overline{\delta}_{q+1}$. □

Lemma 5.3 *The following conditions are equivalent.*

a) $\langle \varphi, \overline{d}_q \psi\rangle = \langle \overline{\delta}_{q+1}\varphi, \psi\rangle$ *for all* $\varphi \in \mathcal{D}_{\overline{\delta}_{q+1}}$, $\psi \in \mathcal{D}_{\overline{d}_q}$.

b) $\overline{d}_q = \delta^*_{q+1}$.

c) $\overline{d}^*_q = \overline{\delta}_{q+1}$.

d) $\overline{\delta}_{q+1} = \overline{\delta}_{q+1,0}$.

e) $\overline{d}_q = \overline{d}_{q,0}$.

Proof a) means $\overline{d}_q \subseteq \delta^*_{q+1}$, $\overline{\delta}_{q+1} \subseteq d^*_q$. Further $\delta_{q+1} \supseteq \delta_{q+1,0}$, $\delta^*_{q+1} \subseteq \delta^*_{q+1,0}$ and $\delta^*_{q+1,0} = \overline{d}_q$ according to (5.9). This yields $\overline{d}_q = \delta^*_{q+1}$, i. e. b). c) follows from b) taking adjoints. Assume c), then $\overline{d}^*_q = \overline{\delta}_{q+1} \supseteq \overline{\delta}_{q+1,0}$ and from (5.9) $\overline{\delta}_{q+1,0} = \delta^{**}_{q+1,0} = \overline{d}^*_q$, hence $\overline{\delta}_{q+1} = \overline{\delta}_{q+1,0}$. Assume d). We obtain from (5.9), (5.10) $\delta^*_{q+1,0} = \overline{d}_q \supseteq \overline{d}_{q,0} = d^{**}_{q,0} = \overline{\delta}^*_{q+1}$, from d) $\overline{\delta}^*_{q+1} = \overline{\delta}^*_{q+1,0} = \delta^*_{q+1,0}$, i. e. $\overline{d}_q = \overline{d}_{q,0}$. Finally $d_{q,0} \subseteq \delta^t_{q+1} \subseteq \delta^*_{q+1}$, i. e. with $\overline{d}_q = \overline{d}_{q,0}$, $\overline{d}_q \subseteq \overline{\delta}^*_{q+1} = \delta^*_{q+1}$, quite similar the second inequality. □

As well known, completeness of (M^n, g) is sufficient for each of the conditions in lemma 5.3, in particular for Stokes' theorem a).

As partial examples we obtain the following self-adjoint extensions of the Laplace operator, the Dirichlet Laplace operator

$$\Delta_D = \overline{\delta}\ \overline{d}_0 + \overline{d}_0\overline{\delta}$$

and the Neumann Laplace operator

$$\Delta_N = \overline{\delta}_0\overline{d} + \overline{d}\ \overline{\delta}_0.$$

Here $x \in \mathcal{D}_{AB}$ means $x \in \mathcal{D}_B$ and $Bx \in \mathcal{D}_A$ and $\mathcal{D}_{A+B} = \mathcal{D}_A \cap \mathcal{D}_B$. The proof of the self-adjointness for Δ_D, Δ_N is quite simple. We use the orthogonal (weak) Hodge decomposition of Kodaira

$$\Omega^{q,2} = \mathcal{H}^{q,2}_w \oplus \overline{d_0\Omega^{q-1}_c} \oplus \overline{\delta_0\Omega^{q+1}_c} \tag{5.11}$$

where

$$\mathcal{H}^{q,2}_w = \{\varphi \in \Omega^{q,2} \,|\, d\varphi = 0, \delta\varphi = 0 \text{ in the distributional sense}\}$$

is the space of weakly harmonic L_2-forms.

Since $\overline{\delta}\ \overline{d}_0$, $\overline{d}_0\overline{\delta}$ are self-adjoint according to (5.11) and von Neumann and $\overline{\delta}\ \overline{d}_0|_{\overline{d_0\Omega^{q-1}_c}} \equiv 0$, $\overline{d}_0\overline{\delta}|_{\overline{\delta_0\Omega^{q+1}_c} \oplus \mathcal{H}^{q,2}} \equiv 0$, Δ_D is the orthogonal sum of two self-adjoint operators, i. e. self-adjoint. In a similar manner we argue for Δ_N using (5.9).

From the definition of Δ_D, Δ_N it is clear that in the case $\overline{d}_0 = \overline{d}$, $\overline{\delta}_0 = \overline{\delta}$ the operators Δ_D, Δ_N coincide with the closure $\overline{\Delta}_0$,

$$\overline{\Delta}_0 = \overline{\delta}_0\delta^*_0 + \overline{d}_0 d^*_0 \tag{5.12}$$

of $\Delta_0 = \Delta|_{\Omega_c^q}$.

$(\mathcal{D}_{\overline{d}_{*,0}}, \overline{d}_{*,0})$, $(\mathcal{D}_{\overline{d}_*}, \overline{d}_*)$ and $(\mathcal{D}_{d_{\max}}, d_{\max})$ belong to the class of so called ideal boundary conditions for the de Rham complex $(\Omega_c^*, d_{*,0})$ and are Hilbert complexes. For completeness, we recall the corresponding notions from [16].

Consider an elliptic complex

$$0 \longrightarrow C_c^\infty(E_0) \longrightarrow \cdots \longrightarrow C_c^\infty(E_q) \longrightarrow C_c^\infty(E_{q+1}) \longrightarrow \cdots \longrightarrow C_c^\infty(E_N) \longrightarrow 0 \tag{5.13}$$

of Hermitean vector bundles $(E_q, h_q) \longrightarrow (M^n, g)$ over (M^n, g) with differentials d_q, $d_{q+1}d_q = 0$. For each q there are Hilbert spaces $H_q = L_2(E_q)$. Each operator d_q has a formal adjoint $d_q^t : C_c^\infty(E_{q+1}) \longrightarrow C_c^\infty(E_q)$ which is also a differential operator. Hence each d_q has closed extensions which lie between the minimal extension $d_{q,\min} = \overline{d}_q$ and the maximal extension $d_{q,\max} = (d^t_{q,\min})^*$. Any choice of closed extensions $d_{q,\min} \subset D_q \subset d_{q,\max}$ that produces a Hilbert complex

$$(\mathcal{D}, D) : \cdots \longrightarrow \mathcal{D}_{D_q} \overset{D_q}{\longrightarrow} \mathcal{D}_{D_{q+1}} \overset{D_{q+1}}{\longrightarrow} \mathcal{D}_{D_{q+2}} \longrightarrow \cdots \tag{5.14}$$

will be called an ideal boundary condition.

We recall from [16]

Proposition 5.4 *Let $(C_c^\infty(E), d)$ be an elliptic complex. Then ideal boundary conditions exist. For example, if we put $H_q = L_2(E_q)$ and $\mathcal{D}_q = \mathcal{D}_{d_{q,\min}}$ or $\mathcal{D}_q = \mathcal{D}_{d_{q,\max}}$, then $(\mathcal{D}, D)$ becomes a Hilbert complex.* □

In our case $(C_c^\infty(E_q), d_q)_q = (\Omega_c^q, d_{q,0})_q$ is $d^t_{q,0} = \delta_{q+1,0}$, and, according to (5.9), $d_{q,\max} = \delta^*_{q+1,0} = \overline{d}_q$ and $d_{q,\min} = \overline{d}_{q,0}$. But this special case $d_{q,\max} = \overline{d}_q$ holds only for the de Rham complex. Its proof uses special features of this complex.

If $\{\tilde{d}_q\}_q$, $\overline{d}_{q,0} \subseteq \tilde{d}_q \subseteq d_{q,\max}$, $0 \le q \le n$, $\tilde{d}_{q+1}\tilde{d}_q = 0$, is an ideal boundary condition for the de Rham complex then we define the corresponding Laplace operator $\tilde{\Delta}_q := \tilde{d}_{q-1}\tilde{d}^*_{q-1} + \tilde{d}^*_q\tilde{d}_q = \tilde{d}_{q-1}\tilde{\delta}_q + \tilde{\delta}_{q+1}\tilde{d}_q$ where we have set $\tilde{\delta}_q := \tilde{d}^*_{q-1}$. We define analogously to above

$$\begin{aligned} H^{q,2,\{\tilde{d}\}} &:= \ker \tilde{d}_q / \operatorname{Im} \tilde{d}_{q-1} = Z^{q,2,\{\tilde{d}\}}/B^{q,2,\{\tilde{d}\}}, \\ \overline{H}^{q,2,\{\tilde{d}\}} &:= \ker \tilde{d}_q / \overline{\operatorname{Im} \tilde{d}_{q-1}} \\ \text{and } \mathcal{H}^{q,2,\{\tilde{d}\}} &:= \ker \tilde{d}_q \cap \ker \tilde{\delta}_q. \end{aligned}$$

Proposition 5.5 **a)** *There is an orthogonal decomposition*

$$L_2(\Lambda^q T^*) \equiv \Omega^{q,2} = \mathcal{H}^{q,2,\{\tilde{d}\}} \oplus \overline{\operatorname{Im} \tilde{d}_{q-1}} \oplus \overline{\operatorname{Im} \tilde{\delta}_{q+1}},$$

b) *$\tilde{\Delta}_q$ is self-adjoint,*
c) $Z^{q,2,\{\tilde{d}\}} = \mathcal{H}^{q,2,\{\tilde{d}\}} \oplus \overline{\operatorname{Im} \tilde{d}_{q-1}}$,
d) $H^{q,2,\{\tilde{d}\}} = \mathcal{H}^{q,2,\{\tilde{d}\}} \oplus \overline{\operatorname{Im} \tilde{d}_{q-1}}/\operatorname{Im} \tilde{d}_{q-1}$,
e) $\overline{H}^{q,2,\{\tilde{d}\}} = \mathcal{H}^{q,2,\{\tilde{d}\}}$,
f) $\mathcal{H}^{q,2,\{\tilde{d}\}} = \ker \tilde{\Delta}_q$.

Proof b) – e) follow immediately from a), hence we start with a). $\tilde{d}_q$ closed implies $\ker\tilde{d}_q$ closed. Then

$$\begin{aligned}\Omega^{q,2} &= (\ker\tilde{d}_q)^\perp \oplus \ker\tilde{d}_q == (\ker\tilde{d}_q)^\perp \oplus \overline{\operatorname{Im}\tilde{d}_{q-1}} \oplus \ker\tilde{d}_q \cap (\operatorname{Im}\tilde{d}_{q-1})^\perp = \\ &= \overline{\operatorname{Im}\tilde{\delta}_{q+1}} \oplus \overline{\operatorname{Im}\tilde{d}_{q-1}} \oplus \ker\tilde{d}_q \cap \ker\tilde{\delta}_q == \overline{\operatorname{Im}\tilde{\delta}_{q+1}} \oplus \overline{\operatorname{Im}\tilde{d}_{q-1}} \oplus \mathcal{H}^{q,2,\{\tilde{d}\}}.\end{aligned}$$

Clearly $\mathcal{H}^{q,2,\{\tilde{d}\}} \subseteq \ker\tilde{\Delta}_q$. If $\varphi \in \ker\tilde{\Delta}_q$ then $0 = \langle\tilde{\Delta}_q\varphi,\varphi\rangle = |\tilde{d}_q\varphi|^2_{L_2} + |\tilde{\delta}_q\varphi|^2_{L_2}$, $\tilde{d}_q\varphi = 0$, $\tilde{\delta}_q\varphi = 0$, $\varphi \in \mathcal{H}^{q,2,\{\tilde{d}\}}$. □

The connection between L_2-cohomology $H^{q,2,\{\tilde{d}\}}$ and the spectrum $\sigma(\tilde{\Delta}_q)$ is given by

Theorem 5.6 *Let (M^n, g) be an open Riemannian manifold $\{\tilde{d}_q\}_q$ an ideal boundary condition.*

a) *If $0 \in \sigma_p(\tilde{\Delta}_q)$ then $H^{q,2,\{\tilde{d}\}} \neq \{0\}$.*

b) *If $0 \in \sigma_{c,R}(\tilde{\Delta}_q)$ then $H^{q,2,\{\tilde{d}\}} \neq \{0\}$ or $H^{q+1,2,\{\tilde{d}\}} \neq \{0\}$.*

c) *If $H^{q,2,\{\tilde{d}\}} \neq \{0\}$ then $0 \in \sigma(\tilde{\Delta}_q)$.*

Proof If $0 \in \sigma_p(\tilde{\Delta}_q)$ then there exists $0 \neq \varphi \in \mathcal{D}_{\tilde{\Delta}_q}$ such that $\tilde{\Delta}_q\varphi = 0$. From proposition 5.5 f) follows $\tilde{d}_q\varphi = 0$, φ is closed. Moreover $\varphi \notin \overline{\tilde{d}_{q-1}\mathcal{D}_{\tilde{d}_{q-1}}}$ since $\ker\tilde{\Delta}_q$ and $\overline{\tilde{d}_{q-1}\mathcal{D}_{\tilde{d}_{q-1}}}$ are orthogonal. Hence φ generates a nontrivial L_2-cohomology class $[\varphi] \in H^{q,2,\{\tilde{d}\}}$, a) is done.

If $0 \in \sigma_{c,R}(\tilde{\Delta}_q)$ then $\ker\tilde{\Delta}_q = \{0\}$, $\operatorname{Im}\tilde{\Delta}_q$ is dense but not closed. $\operatorname{Im}\tilde{\Delta}_q$ is closed if and only if $\operatorname{Im}\tilde{d}_{q-1}$ and $\operatorname{Im}\tilde{d}_q$ are closed. Hence under our assumption $\operatorname{Im}\tilde{d}_{q-1}$ or $\operatorname{Im}\tilde{d}_q$ is not closed, $H^{q,2,\{\tilde{d}\}} = \ker\tilde{d}_q/\operatorname{Im}\tilde{d}_{q-1} = \overline{\operatorname{Im}\tilde{d}_{q-1}}/\operatorname{Im}\tilde{d}_{q-1} \neq \{0\}$ or $H^{q+1,2,\{\tilde{d}\}} = \ker\tilde{\Delta}_{q+1} \oplus \overline{\operatorname{Im}\tilde{d}_q}/\operatorname{Im}\tilde{d}_q \neq \{0\}$. b) is done.

Assume conversely $H^{q,2,\{\tilde{d}\}} \neq \{0\}$, $0 \neq [\varphi] \in H^{q,2,\{\tilde{d}\}}$. Then φ is closed and either $\varphi \in \ker\tilde{\Delta}_q$ or $\varphi \in \overline{\operatorname{Im}\tilde{d}_{q-1}} \setminus \operatorname{Im}\tilde{d}_{q-1}$ (after decomposition). If $0 \neq \varphi \in \ker\tilde{\Delta}_q$ then $0 \in \sigma_p(\tilde{\Delta}_q) \subseteq \sigma(\tilde{\Delta}_q)$. Assume for all $0 \neq [\varphi] \in H^{2,p,\{\tilde{d}\}}$ and all $\psi \in [\varphi]$ that $\psi \notin \ker\tilde{\Delta}_q$. Then $\ker\tilde{\Delta}_q = \{0\}$ and $0 \notin \sigma_p(\tilde{\Delta}_q)$. Now we must prove $0 \in \sigma_{c,R}(\tilde{\Delta}_g)$. From $\ker\tilde{\Delta}_q = 0$ follows $\overline{\operatorname{Im}\tilde{\Delta}_q} = \Omega^{q,2,0} = L_2(\Lambda^q)$. Suppose $\operatorname{Im}\tilde{\Delta}_q = L_2(\Lambda^q)$, i.e. $\tilde{\Delta}_q$ is surjective. Let $0 \neq [\varphi] \in H^{q,2,\{\tilde{d}\}}$, $\varphi \in [\varphi]$. There exists ψ such that $\tilde{\Delta}_q\psi = \varphi$. $\tilde{d}_q\varphi = 0$ implies $0 = \tilde{d}_q\tilde{\Delta}_q\psi = \tilde{d}_q\tilde{\delta}_{q+1}\tilde{d}_q\psi$, $0 = |\tilde{\delta}_{q+1}\tilde{d}_q\psi|^2_{L_2}$, $\tilde{\delta}_{q+1}\tilde{d}_q\psi = 0$, $\varphi = \tilde{\Delta}_q\psi = \tilde{d}_{q-1}\tilde{\delta}_q\psi + \tilde{\delta}_{q+1}\tilde{d}_q\psi = \tilde{d}_{q-1}(\tilde{\delta}_q\psi)$. i. e. φ is a coboundary in contradiction to $[\varphi] \neq 0$. Hence $\operatorname{Im}\tilde{\Delta}_q \subset L_2$, $\overline{\operatorname{Im}\tilde{\Delta}_q} = L_2$, i.e. $0 \in \sigma_{c,R}(\tilde{\Delta}_q)$. □

Special cases of proposition 5.5 and theorem 5.6 are $\tilde{d}_q = \overline{d}_{q,0}$, $\tilde{\Delta}_q = \overline{\Delta}_{q,D}$ or $\tilde{d}_q = \overline{d}_q$, $\tilde{\Delta}_q = \Delta_{q,N}$. If $\overline{d}_{q,0} = \overline{d}_q$, or equivalently if Stokes' theorem holds, then for $\tilde{d}_q = \overline{d}_q = \overline{d}_{q,0}$, $\tilde{\Delta}_q = \overline{\Delta}_{q,0}$ the closure of $\Delta_q|_{\Omega^q_c}$ and the assertions above hold for $\overline{\Delta}_{q,0}$. In particular this holds for complete (M^n, g).

We draw as conclusion that the question $0 \in \sigma(\Delta_q)$ or not essentially amounts to the investigation of L_2-cohomology. Moreover, L_2-cohomology is of additionally independent interest since in many interesting cases it is related to combinatorial L_2-cohomology or intersection homology, i. e. to topological features. It is clear that L_2-cohomology is an invariant of the quasi isometry class of g since transition from g to a quasi isometric g'

produces equivalent L_2-norms. We return now to the zero in the spectrum conjecture for complete manifolds.

Conjecture. For all complete Riemannian manifolds (M^n, g), there holds $0 \in \sigma(\Delta_q)$ for some q (between 0 and n).

This conjecture is supported by several facts and examples.

1) For manifolds with finite volume this is true: $0 \in \sigma(\Delta_0)$, $\Delta_0 1 = 0$, $1 \in C^\infty \cap L_2$; in particular this holds for closed manifolds.

2) $\sigma(\Delta_q(\mathbb{R}^n, g_{\text{standard}})) = [0, \infty[$, $0 \leq q \leq n$.

3)

$$\sigma(\Delta_q(H^n_{-1})) = \begin{cases} \left[\frac{(n-1)^2}{4}, \infty\right[= \sigma_e & \text{for } q = 0, n, \\ [\lambda_1, \infty[= \sigma_e, \lambda_1 = \min\left\{\frac{(n-2q+1)^2}{4}, \frac{(n-2q-1)^2}{4}\right\} & \text{for } q \neq 0, n, \frac{n}{2}, \\ \{0\} \cup \left[\frac{n}{2}, \infty\right[= \sigma_e & \text{for } q = \frac{n}{2}. \end{cases}$$

4) $0 \in \sigma(\Delta_i(H^{2n+1}_{-1}))$, $i = n, n+1$.

If $q = \frac{n}{2}$ then 0 is an eigenvalue of infinite multiplicity.

There are many further classes of examples which support the conjecture and for a couple of years nobody has been able to find a counterexample against the conjecture.

It were Michael Farber and Shmuel Weinberger which constructed in 1999 a class of counterexamples (cf. [56]). In the same year Nigel Higson, John Roe and Thomas Schick established a simplified class of counterexamples (cf. [76]). We sketch their main steps.

1) According to theorem 5.6, zero will not be in the spectrum of some Δ_q if and only if $H^{q,2,\{d\}}(M^n, g) = 0$, $0 \leq q \leq n$.

2) Let (M^n, g) be closed and $(\tilde{M}^n, \tilde{g})$ be an infinite covering. Then $H^{q,2,\{d\}}(\tilde{M}) = H^q(M; l_2(\pi))$, where $l_2(\pi)$ is the completion of $\mathbb{C}\pi$, $\pi =$ Deck $(\tilde{M} \longrightarrow M)$. In particular, if $\pi = \pi_1(M)$, $\tilde{M} \longrightarrow M$ the universal covering, then $0 \notin \sigma(\Delta_q(\tilde{M}))$ for all q if $H^q(M; l_2(\pi)) = 0$ for all q.

3) Denote by $C^*_r(\pi)$ the reduced C^*-algebra of π and $\mathcal{N}(\pi)$ the von Neumann algebra. Higson, Roe and Schick made the following fundamental observation. Suppose Z is a CW-complex with countably many cells overall and finitely many cells in dimension 0 through n. The following are equivalent: The homology groups $H_*(Z; C^*_r(\pi))$ are zero in degrees zero through n. The homology groups $H_*(Z; \mathcal{N}(\pi))$ are zero in degrees zero through n. The homology groups $H_*(Z; l_2(\pi))$ are zero in degrees zero through n.

4) Let G be a finitely presented group such that $H_k(G; C^*_r(G))) = 0$ for $k = 0, 1, 2$ and W a finite 2-dimensional CW-complex with $\pi_1(W) = G$. Then $H_0(W; C^*_r(G)) = 0$ and $H_1(W; C^*_r(G)) = 0$ since the classifying map $W \longrightarrow BG$ is 2-connected and because the connectedness assumption on G. $H_2(W; C^*_r(G))$ is finitely generated and stably free.

5) By wedging with finitely many 2-spheres one obtains a finite CW-complex Y with $\pi_1(Y) = G$ for which $H_2(Y; C^*_r(G))$ is a free $C^*_r(G)$ module.

6) The Hurewicz map $h : \pi_2(Y) \otimes C^*_r(G) \longrightarrow H_2(Y; C^*_r(G))$ is surjective since its cokernel $H_2(G; C^*_r(G)) = 0$. The inclusion $Q[G] \longrightarrow C^*_r(G)$ induces a map

$$i : \pi_2(Y) \otimes Q[G] \longrightarrow \pi_2 \otimes C^*_r(G)$$

and the inclusion $\mathbb{Z}[G] \longrightarrow C^*_r(G)$ induces

$$j : \pi_2(Y) = \pi_2(Y) \otimes \mathbb{Z}[G] \longrightarrow \pi_2(Y) \otimes C^*_r(G).$$

7) Main lemma. The image of the composition

$$h \circ j : \pi_2(Y) \longrightarrow H_2(Y; C_r^*(G))$$

contains a basis for the free $C_r^*(G)$-module $H_2(G; C_r^*(G))$.

8) Choose elements $v_1, \ldots, v_d \in \pi_2(Y)$ such that $h \circ j$ sends them to a basis for $H_2(Y; C_r^*(G))$. Each v_k is represented by a map $S^2 \longrightarrow Y$. Attach be means of this maps 3-cells to Y. Denote the result by X. Then

$$H_*(X; C_r^*(G)) = 0.$$

9) Embed X into Euclidean space and take the boundary M^n of a regular neighborhood. Then

$$H_k(M^n; C_r^*(G)) = 0$$

for all k, $0 \notin \sigma(\Delta_q(\tilde{M}^n))$, $0 \le q \le n$.

We summarize these steps as

Theorem 5.7 *Let G be a finitely presented group and suppose that the homology groups $H_k(G; C_r^*(G))$ are zero for $k = 0, 1, 2$. This is equivalent to the vanishing of the (unreduced) L^2-homology of G in dimensions* $0, 1, 2$.

Then there is a 3-dimensional finite CW-complex X with $\pi_1(X) = G$ such that $H_k(X; C_r^*(G)) = 0$ for all $k \in \mathbb{N}$- Moreover, for every dimension $n \ge 6$ there is a closed manifold M_n of dimension n and with $\pi_1(M_n) = G$ such that $H_k(M_n; C_r^*(G)) = 0$ for all $k \in \mathbb{N}$.

For the universal covering $(\tilde{M}^n, \tilde{g})$ then holds $0 \notin \sigma(\Delta_q(\tilde{M}, \tilde{g}))$, $0 \le q \le n$. i.e. $(\tilde{M}^n, \tilde{g})$ is a counterexample to the zero in the spectrum conjecture. □

It would be interesting to produce other complete counterexamples which are not universal coverings of closed manifolds and nontrivial noncomplete counterexamples. The following is clear. If $H^{q,2,\{d\}}(M^n, g) = 0$, $0 \le q \le n$, then in particular $\inf \sigma_e(\Delta_q) = \lambda_{q,e} > 0$, $0 \notin \sigma_e(\Delta_q)$. $\sigma_e(\Delta_q)$ remains unchanged under compact perturbations. Hence, if after a compact perturbation $0 \in \sigma(\Delta_q)$ then it must be an eigenvalue of finite multiplicity, i.e. $0 \in \sigma_{pd}(\Delta_q)$. Ingolf Buttig generated in [19] eigenvalues $\lambda > 0$ below the bottom of the essential spectrum by controled perturbations. The question is, will this be possible also for $\lambda = 0$? The answer is yes.

This finishes our preliminary discussion of the zero in the spectrum conjecture.

The most easy case for L_2-cohomology would be the case which admits only one ideal boundary condition, i.e. the case $d_{q,\min} = \overline{d}_{q,0} = d_{q,\max} = \overline{d}_q$, $0 \le q \le n$, where the latter equality only holds for the de Rham complex, not for arbitrary elliptic complexes (cf. [24] , p. 135). Sufficient for this was the completeness of (M^n, g). Consequently, the study of L_2-cohomology splits into basically distinct cases, the complete and incomplete case. The complete case belongs more to the area of differential geometry and global analysis, the incomplete case more to algebraic geometry. The generic example for the latter are projective algebraic varieties V^n with singularity set Σ and are manifolds $V^n \setminus \Sigma$ with the Kähler metric induced from the Fubini-Study metric. Fortunately in this situation, there is another description of L_2-cohomology, given by intersection homology. We come to this later.

To submit a certain geometrical/analytical feeling for L_2-cohomology, we present some examples. For the complete case, where we have only one ideal b. c., we write $H^{q,2}$ for $H^{q,2,\{d\}}$, $\mathcal{H}^{q,2}$ etc..

Examples 5.8 a) If (M^n, g) is complete and vol $(M^n, g) = \infty$ then always $\overline{H}^{0,2} = H^{0,2} = \{0\}$. In particular, $H^{0,2}(N^{n-1} \times \mathbb{R}) = \{0\}$ if N^{n-1} is closed and we endow $N^{n-1} \times \mathbb{R}$ with the product metric. Moreover, $\overline{H}^{0,2}(N^{n-1} \times \mathbb{R}) = 0$ since for any open manifold $H^{0,2} = \overline{H}^{0,2}$.

b) Consider $N^{n-1} \times \mathbb{R}$ with a metric of infinite volume and Ricci curvature Ric ≥ 0. Then from Ric ≥ 0 we immediately obtain $\{0\} = \overline{H}^{1,2} = \mathcal{H}^{1,2}$. Consider a C^∞ function $h(u, r) = f(r)$ on $M = N \times \mathbb{R}$ such that $f(r) = \frac{1}{r}$ for $|r| \geq 1$. Then $\omega = f(r)dr \in Z^{1,2}$. The most general function φ such that $d\varphi = f(r)dr$ satisfies for $|r| \geq 1$ $\varphi(r) = \log r + C$, hence $\varphi \notin \Omega^{0,2} = L_2$ and $f dr \notin B^{1,2}$, $H^{1,2} \neq \{0\}$. We infer from c) $\dim H^{1,2} = \infty$.

c) If $\mathcal{H}^{q,2} = \{0\}$ then either $\dim H^{q,2} = 0$ or $\dim H^{q,2} = \infty$. Hence $\dim H^{q,2} < \infty$ implies automatically $H^{q,2} = \mathcal{H}^{q,2}$.

d) Let (M^2, g) be the Poincare disc ($\{|z| < 1\}$, $ds^2 = \frac{|dz|^2}{(1-|z|^2)^2}$). Then $\mathcal{H}^{0,2}(M^2) = \mathcal{H}^{2,2}(M) = \{0\}$ and $\mathcal{H}^{1,2}(M^2) = \{\omega = adx + bdy | b_x - a_y = 0, a_x + b_y = 0, \int\limits_{M^2} (a^2 + b^2)dxdy < \infty\}$, which can be identified with the space of square integrable harmonic functions $f(z) = a - ib$. □

We start with calculations in the complete case. Then

$$H^{q,2}(M^n, g) = \mathcal{H}^{q,2}(M^n, g) \oplus \overline{\operatorname{Im} \overline{d}_{q-1}} / \operatorname{Im} \overline{d}_{q-1} = \ker \Delta_q \oplus \overline{\operatorname{Im} \overline{d}_{q-1}} / \operatorname{Im} \overline{d}_{q-1}.$$

One approach to calculate $H^{q,2}$ would be to calculate each single summand in the direct sum. Let us start with $\mathcal{H}^{q,2}$. The classical method to prove the vanishing of $\mathcal{H}^{q,2}$ is Bochner's method which can be extended to open manifolds, and this has been performed by Dodziuk. Write the formula (2.2) in section 2 as $\Delta\omega = \nabla^*\nabla\omega + \mathcal{R}_q\omega$. Denote by (,) the pointwise scalar product.

Theorem 5.9 *Suppose the form $\omega_x \longrightarrow (\mathcal{R}_q\omega, \omega)_x$ is positive semidefinite for all $x \in M$. Then every L_2-harmonic q-form is parallel. If* vol $(M^n, g) = \infty$ *or $(\mathcal{R}_q\omega, \omega)_x$ is positive definite then every L_2-harmonic q-form is identically zero.*

Proof Let $\Phi_{r,s}$ be an almost differentiable function on M such that

a) $0 \leq \Phi_{r,s}(x) \leq 1$,

b) supp $\Phi_{r,s} = B_s(x_0)$,

c) $\Phi_{r,s}(x) = 1$ on $B_r(x_0)$,

d) $\lim\limits_{r,s\to\infty} \Phi_{r,s} = 1$,

e) $|d\Phi_{r,s}(x)| = |\nabla\Phi_{r,s}(x)| \leq \frac{c}{s-r}$ almost everywhere,

and set $\lambda_R(x) = \Phi_{R,2R}$. Then $\lambda_R(x)$ is Lipschitz, $0 \leq \lambda_R(x) \leq 1$ for every $x \in M$, supp $\lambda_R \subset B_{2R}(x_0)$, $\lambda_R(x) = 1$ on $B_R(x_0)$, $\lim\limits_{R\to\infty} \lambda_R = 1$ and $|d\lambda_R(x)| \leq C_1/R$ almost everywhere. Then for any $\omega \in \Omega^{q,2} \cap C^2$

$$\langle \Delta\omega, \lambda_R^2\omega \rangle = \langle \lambda_R\nabla\omega, \lambda_R\nabla\omega \rangle + \langle \lambda_R\nabla\omega, 2d\lambda_R \otimes \omega \rangle + \langle \mathcal{R}_q\omega, \lambda_R^2\omega \rangle. \tag{5.15}$$

We have the simple estimates

$$|d\lambda_R \otimes \omega|_{L_2}^2 \leq \frac{C_1^2}{R^2} |\omega|_{L_2}^2$$

and

$$|(\lambda_R \nabla\omega, 2d\lambda_R \otimes \omega)|_{L_1} \le \frac{1}{2}|\lambda_R \nabla\omega|^2_{L_2} + \frac{2C_1^2}{R^2}|\omega|^2_{L_2}. \tag{5.16}$$

If $\Delta\omega = 0$ then we infer from (5.15), (5.16)

$$\frac{1}{2}|\lambda_R \nabla\omega|^2_{L_2} + \langle \mathcal{R}_q \lambda_R \omega, \lambda_R \omega\rangle \le \frac{2C_1^2}{R^2}|\omega|^2_{L_2}. \tag{5.17}$$

Taking $\lim\limits_{R\to\infty}$ in (5.17) yields $|\nabla\omega|_{L_2} = 0$, ω parallel, $|\omega|$ constant, $\omega = 0$ if vol $(M^n, g) = 0$. Moreover, we conclude from (5.17) that $\langle \mathcal{R}_q\omega, \omega\rangle = \int\limits_M (\mathcal{R}_q\omega, \omega)_x \, \mathrm{dvol}_x(g)$ exists and = 0. Hence $\omega_x = 0$ at every x where $(\mathcal{R}_q\cdot, \cdot)$ is positive definite. If such a point x exists then $\omega_x = 0$, $\omega \equiv 0$. □

Now the main task is to calculate $H^{*,2}$, $\overline{H}^{*,2}$ for as much as possible manifolds. We refer to [22], [32], [35] and present here for the reasons of space only some classes.

Proposition 5.10 *Let (M^n, g) be open, complete, with finite volume and pinched sectional curvature K, $-1 \le K \le -1 + \varepsilon$, $0 \le \varepsilon 1$.*

a) *If $\varepsilon < \frac{1-4q^2}{(n-1)^2}$ and $q < \frac{n-1}{2}$ then $\dim \mathcal{H}^{q,2}(M) < \infty$ and $\dim \mathcal{H}^{n-q,2}(M) < \infty$.*

b) *If $n = 2m$ and $\varepsilon < \frac{2}{(2m-1)^2 + 2(m-1)^2\beta(n)}$ then $\dim \mathcal{H}^{m,2}(M^{2m}) < \infty$.*

c) *If $n = 2$, $0 \le \varepsilon < 1$ arbitrary, then $\dim \mathcal{H}^{1,2}(M^2) < \infty$.* □

For complex manifolds, we mention some very deep results established by Donnelly, Fefferman and Gromov.

Theorem 5.11 *Let (M^n, g) be strictly pseudoconvex in $\mathbb{C}^n$ endowed with its Bergman metric. Then (M^n, g) is a complete Kähler metric of real dimension $2n$. Denote by $\mathcal{H}^{(p,q),2}$ the space of L_2-harmonic (p, q)-forms. Then*

$$\dim \mathcal{H}^{(p,q),2} = \begin{cases} 0, & p+1 \ne n \\ \infty, & p+q = n \end{cases}.$$

The proof essentially uses and sharpens for Hermitean manifolds the estimates of [35], discussed here in the preceding section. Moreover it uses the fact that the holomorphic sectional curvatures approach a negative constant near the boundary. We refer to [36] for the proof. □

Gromov proves in [68] that for a compact Kähler hyperbolic manifold M^n the L_2-Hodge numbers $h^{(p,q),2} = \dim \mathcal{H}^{(p,q),2}$ vanish if and only if $p + q < n$. The vanishing of $\dim_{\mathbb{C}} \mathcal{H}^{(p,q),2}$ has been independently proved by Stern, if (M^n, g) is a complete simply connected Kähler manifold with negatively pinched sectional curvature, $-b \le K \le -a < 0$. Gromov's result immediately implies

$$\operatorname{sign} \chi(M) = (-1)^n$$

for a compact Kähler manifold M^n of negative curvature and of real dimension $2n$, which is a special case of the Hopf conjecture. Other vanishing theorems for L_2-cohomology in the Kähler case are established by Jost/Kang Zuo in [79].

The vanishing properties of $\mathcal{H}^{q,2}$ in the case of negative curvature remain valid for L_2-cohomology with coefficients if M is a symmetric space of noncompact type.

Let G be a connected linear semi-simple Lie group, K a maximal subgroup of G. Then $M = G/K$ has a G-invariant Riemannian metric such that it is a Riemannian symmetric space of negative curvature and homeomorphic to Euclidean space. Let (ϱ, E) be an irreducible finite dimensional representation of G. This yields a complex of differential forms with values in the associated vector bundle which we also denote by $\mathbf{E}$. Then L_2-cohomology $H^{q,2}(M;\mathbf{E})$ is at least well defined.

Theorem 5.12 *Let* $m = \frac{\dim M}{2}$ *and assume* $\operatorname{rk} G = \operatorname{rk} K$. *Then there holds.*

a) $\operatorname{Im} \overline{d}$ *is closed and*

$$H^{q,2}(M;\mathbf{E}) = \overline{H}^{q,2}(M;\mathbf{E}) \cong \mathcal{H}^{q,2}(M;\mathbf{E}) = (0)$$

if $q \neq m$.

b) *The G-space* $H^{m,2}(M;\mathbf{E}) = H^{m,2}(M;\mathbf{E}) \cong \mathcal{H}^{m,2}(M;\mathbf{E})$ *is the direct sum of the discrete series representations of G having the same infinitesimal character as* (ϱ, E).

We refer to [15] for the proof. □

A complete description of $\mathcal{H}^{*,2}$ for a hyperbolic locally symmetric space $M^n = H^n_{-1}/\Gamma$ is given by Mazzeo and Philipps in [86]. Vanishing theorem for L_2-harmonic forms in the case of positive curvature are established by Escobar/Freire/Min-Oo in [55].

As an application of L_2-cohomology in algebraic geometry we present the isomorphism between the L_2-cohomology $H^{*,2}(X \setminus \Sigma)$ and the dual $(IH^{\overline{p}}_*(X))^*$ of the intersection homology for certain classes of stratified spaces with singularity Σ and a suitable metric. To formulate a precise theorem, the introduction of some concepts is unavoidable. Let X^n be a pseudomanifold, i.e. a polyhedron such that there exists a closed subspace Σ with $\dim \Sigma \leq n-2$ and $X \setminus \Sigma$ being a dense oriented manifold in X. Let X be a pseudomanifold with triangulation T. A stratification of X is a filtration by closed subspaces

$$X^n = X_n \supset X_{n-1} \supset X_{n-2} \supset X_{n-3} \supset \cdots \supset X_1 \supset X_0$$

such that for each point $p \in X_i \setminus X_{i-1}$ there is a filtered space

$$V = V_n \supset V_{n-1} \supset \cdots \supset V_i = \text{ a point},$$

and a mapping $V \times B^i \longrightarrow X$ which maps $V_j \times B^i$, for each j, PL-homeomorphically onto a neighborhood of p in X_j, where B^i denotes the PL i-ball. In particular, $X_{(i)} = X_i/X_{i-1}$ is an i-manifold or is empty. Denote by $C^T_*(X;\mathbb{R}) = C^T_*(X)$ the corresponding (to T) chain complex of all simplicial chains with real coefficients and $C_*(X) = \lim\limits_{\vec{T}} C^T_*(X)$ the group of all PL geometric chains. Each $\xi \in C_i(X)$ has a well defined support $|\xi|$. For a perversity $\overline{p}$, i.e. a sequence of integers $\overline{p} = (p_2, p_3, \ldots, p_n)$ with $p_2 = 0$ and $p_k \leq p_{k+1} \leq p_k + 1$, we define

$$\begin{aligned} IC^{\overline{p}}_i(X) = \{\xi \in C_i(X) | \dim(|\xi|) \cap X_{n-k}) \leq i - k + p_k, \\ \dim(|\partial\xi| \cap X_{n-k}) \leq i - 1 - k + p_k \text{ for all } k\}. \end{aligned}$$

The i-th intersection homology group $IH^{\overline{p}}_i(X)$ of X with perversity $\overline{p}$ and with a fixed stratification is defined to be the i-th homology group of the chain complex $IC^{\overline{p}_*(X)}$.

Remark 5.13 Cheeger considers in his paper [24] only the so-called middle perversity $\overline{m} = (0, 0, 1, 1, 2, 2, \ldots, \frac{n}{2-1})$. □

Theorem 5.14 *Let X be a compact pseudomanifold without boundary. Then $IH_*^{\overline{p}}(X)$ is finitely generated and independent of stratification.* □

For a sequence $\overline{c} = (c_2, \dots, c_n)$ of nonnegative real numbers, a metric g on $X_{(n)}$ is said to be associated with $\overline{c}$ if, at a local product representation for a tubular neighbourhood, for every stratum $X_{(j)}$ the metric g is locally of the kind

$$g_U + dr \otimes dr + r^{2c} n - j \cdot g_s,$$

where g_U is a metric in the base U, $dr \otimes dr$ is the metric in radial direction and g_s is the induced metric in some fixed sphere of radius r ([92], p. 345).

Lemma 5.15 *For any $\overline{c}$ there exists a metric at $X_{(n)}$ which is associated with $\overline{c}$.* □

For any perversity $\overline{p} \le \overline{m}$ (i.e. such that $p_i \le m_i$), a metric g is said to be associated with $\overline{p}$, if g is associated with $\overline{c} = (c_2, \dots, c_n)$ and

$$\begin{aligned} &\frac{1}{k-1-2p_k} \le c_k < \frac{1}{k-3-2p_k} \quad \text{if } 2p_k \le k-3, \\ &1 \le c_k < \infty \quad \text{if } 2p_k = k-2. \end{aligned}$$

Theorem 5.16 *Let X^n be an n-dimensional compact stratified space with a fixed PL structure and a stratification $X = X_n \supset X_{n-1} \supset \cdots \supset X_0$ such that $X_{n-1} = X_{n-2}$ and each stratum $X_{(j)}$ of dimension $j \le n-2$ is diffeomorphic to the disjoint union of $]0,1[^j$, $X_{(j)} = \bigcup_\alpha (]0,1[^j)_\alpha$. Let $\overline{p} \le \overline{m}$ be a perversity and g a metric on $X_{(n)}$ associated with $\overline{p}$. Then*

$$H^{*,2}(X_{(n)}, d) \cong (IH_q^p(X))^*$$

(see [92]*).* □

The interesting examples are projective varieties with singularities. There are many very deep and beautiful results established e.g. by Hsiang, Saper, Stern, Zucker and others (cf. [22], [101], [102], [123], [124], [125], [126]).

We mention here as an explicit example related to all the cases above the beautiful result of Saper and Stern in [101]. Let G be the real points of a semi-simple algebraic group defined over $\mathbb{Q}$, K a maximal compact subgroup, and $D = G/K$ the associated symmetric space. Assume that D is Hermitean, that is, a bounded symmetric domain. Let $\Gamma \subset G$ be an arithmetic subgroup which acts freely on D. The quotient $\Gamma \setminus D$ has a natural complete metric induced from the G-invariant metric (or Bergmann metric) on D.

In general D/Γ is not compact, but there is the well known Satake compactification $\Gamma \setminus D^*$ which has the structure of a normal projective algebraic variety. Let $\mathbf{E}$ be a metrized coefficient system on D/Γ arising from a finite dimensional complex representation E of G with admissible metric.

Theorem 5.17 *There is a natural isomorphism*

$$(\Gamma \setminus D; E) \cong IH^*(\Gamma \setminus D^*; E)$$

between the L_2-cohomology of $\Gamma \setminus D$ and the (middle perversity) intersection cohomology of $\Gamma \setminus D^*$. □

6 The heat equation, the heat kernel and the heat flow

Let $(E, h, \nabla, \cdot) \longrightarrow (M^n, g)$ be a Clifford bundle, (M^n, g) complete, $D = D(E, h, \nabla, \cdot)$ the associated generalized Dirac operator and consider the initial value problem

$$\frac{\partial \varphi}{\partial t} + D^2 \varphi = 0, \tag{6.1}$$

$$\varphi(\cdot, 0) = \varphi_0 \tag{6.2}$$

for sections $\varphi = \varphi(x,t)$. Here we suppose $\varphi(\cdot, t) \in \Omega^{0,2,r}(E, D) \subset L_2(E)$, $r > \frac{n}{2} + 2$, $\varphi \in C^1$ with respect to $t > 0$. According to proposition 4.24, all powers D^i are in $L_2(E) = H^0(E)$ essentially self-adjoint, $D^i = \int\limits_{-\infty}^{+\infty} \lambda^i dE_\lambda(D)$.

Denote in the sequel $W^{p,r} \equiv \Omega^{p,r}(E, \nabla)$, $H^{p,r} \equiv \Omega^{p,r}(E, D)$, $W^{2,r} \equiv W^r$, $H^{2,r} \equiv H^r$. For $k < 0$, we define the Sobolev space $H^k = \Omega^{0,2,k}(E) = \Omega^{2,k}(E, D)$ by duality, i.e. $H^k(E) := (H^{-k}(E))^*$.

Lemma 6.1 *e^{-tD^2} maps $L_2(E) \equiv H^0(E) \to H^r(E)$ for any $r > 0$ and*

$$|e^{-tD^2}|_{L_2 \to H^r} \le C \cdot t^{-\frac{r}{2}}, \quad t \in]0, \infty[, \quad C = C(r).$$

Proof Insert into $e^{-tD^2} = \int e^{-t\lambda^2} \, dE_\lambda$ the equation

$$e^{-t\lambda^2} = \frac{1}{\sqrt{4\pi t}} \int\limits_{-\infty}^{+\infty} e^{i\lambda s} e^{-\frac{s^2}{4t}} \, ds$$

and use

$$\sup |\lambda^r e^{-t\lambda^2}| \le C \cdot t^{-\frac{r}{2}}.$$

□

Corollary 6.2 *Let $r, s \in \mathbb{Z}$ be arbitrary. Then $e^{-tD^2} : H^r(E) \to H^s(E)$ is continuous.*

Proof This follows from 6.1, duality and the semi group property of $\{e^{-tD^2}\}_{t \ge 0}$. □

Proposition 6.3 *The initial value problem (6.1), (6.2) has the unique solution*

$$\varphi(\cdot, t) = e^{-tD^2} \varphi(\cdot, 0).$$

Proof Inserting (6.1) into $\frac{\partial}{\partial t} |\varphi(\cdot, t)|^2_{L_2}$, we get $\frac{\partial}{\partial t} |\varphi(\cdot, t)|^2_{L_2} = \frac{\partial}{\partial t} \langle \varphi(\cdot, t), \varphi(\cdot, t) \rangle_{L_2} = -2 |D\varphi(\cdot, t)|^2_{L_2}$, hence $|\varphi(\cdot, t)|^2_{L_2} \le |\varphi(\cdot, 0)|^2_{L_2}$. This proves the uniqueness of the solution. Set

$$\varphi(\cdot, t) = e^{-tD^2} \varphi(\cdot, 0).$$

Then, according to corollary 6.2, $\varphi(\cdot, t) \in \Omega^{0,2,r}(E, D) = H^r(E)$ for all r, $\varphi(\cdot, t) \in C^\infty$. □

From the spectral representation we obtain immediately

$$\frac{\partial \varphi}{\partial t} - D^2 \varphi = -D^2 \varphi + D^2 \varphi = 0,$$

which provides existence. $\square$

Nevertheless, the representation for $\varphi(\cdot,t)$ is in a certain sence of purely theoretical character since the integral in question practically is not available. Hence one would be interested in a more explicit representation of e^{-tD^2} as integral operator.

e^{-tD^2} has a Schwartz kernel $W \in \Gamma(\mathbb{R} \times M \times M, E \boxtimes E)$,

$$W(x,y,t) = \langle \delta(x), e^{-tD^2}\delta(y)\rangle,$$

where $\delta(x) \in H^{-r}(E) \otimes E_x$ is the map $\Psi \in H^r(E) \longrightarrow \langle \delta(x), \Psi\rangle = \Psi(x)$, $r > \frac{n}{2}$. The main result of this section is the fact that for $t > 0, W(x,y,t)$ is a smooth integral kernel in L_2 with good decay properties if we assume bounded geometry.

Denote by $C(x)$ the best local Sobolev constant of the map $\Psi \longrightarrow \Psi(x), r > \frac{n}{2}$, and by $\sigma(D^2)$ the spectrum.

Lemma 6.4 **a)** *$W(x,y,t)$ is smooth for $t > 0$ in all variables.*

b) *For any $T > 0$ and sufficiently small $\varepsilon > 0$, there exists $C > 0$ such that*

$$|W(x,y,t)| \le e^{-(t-\varepsilon)\inf \sigma(D^2)} \cdot C \cdot C(x) \cdot C(y) \ \text{ for all } \ t \in]T, \infty[.$$

c) *Similar estimates hold for $(D^i_x D^j_y W)(x,y,t)$.*

Proof a) First one shows that W is continuous, which follows from the fact that $\langle \delta(x), \cdot\rangle$ is continuous in x and that $e^{-tD^2}\delta(y)$ is continuous in t and y. Then one applies elliptic regularity.

b) Write $|\langle \delta(x), e^{-tD^2}\delta(y)\rangle| = |\langle (1+D^2)^{-\frac{r}{2}}\delta(x), (1+D^2)^r e^{-tD^2}(1+D^2)^{\frac{r}{2}}\delta(y)\rangle|$ and estimate.

c) Follows similar by b). $\square$

Lemma 6.5 *For any $\varepsilon > 0, T > 0, \delta > 0$ there exists $C > 0$ such that for $r > 0$, $x \in M$, $T > t > 0$:*

$$\int\limits_{M \setminus B_r(x)} |W(x,y,t)|^2 \, dp \le C \cdot C(x) \cdot e^{-\frac{(r-\varepsilon)^2}{(4+\delta)t}}.$$

A similar estimate holds for $D^i_x D^j_y W(x,y,t)$.

We refer to [17] for the proof. $\square$

Lemma 6.6 *For any $\varepsilon > 0, T > 0, \delta > 0$ there exists $C > 0$ such that for all $x, y \in M$ with $dist(x,y) > 2\varepsilon, T > t > 0$ holds*

$$|W(x,y,t)|^2 \le C \cdot C(x) \cdot C(y) \cdot e^{-\frac{(dist(x,y)-\varepsilon)^2}{(4+\delta)t}}.$$

A similar estimate holds for $D^i_x D^j_y W(x,y,t)$.

Again, the proof can be found in [17]. $\square$

Proposition 6.7 *Assume (M^n, g) with (I) and (B_K), (E, ∇) with (B_K), $k \ge r > \frac{n}{2} + 1$. Then all estimates in 6.4, 6.5, 6.6 hold with uniform constants.*

Proof From the assumptions, $H^r(E) \cong W^r(E)$ and $\sup_x C(x) = C$ is the global Sobolev constant for $W^r(E)$, according to theorem 2.9 b). $\square$

Let $U \subset M$ be precompact, open, (M^+, g^+) closed with $U \subset M^+$ isometrically and $E^+ \to M^+$ a Clifford bundle with $E^+|_U \cong E|_U$ isometrically. Denote by $W^+(x, y, t)$ the heat kernel of $e^{-tD^{+2}}$.

Lemma 6.8 *Assume $\varepsilon > 0, T > 0, \delta > 0$. Then there exists $C > 0$ such that for all $T > t > 0$, $x, y \in U$ with $B_{2\varepsilon}(x), B_{2\varepsilon}(y) \subset U$, one has*

$$|W(x, y, t) - W^+(x, y, t)| \leq C \cdot e^{-\frac{\varepsilon^2}{(4+\delta)t}}$$

We refer to [17] for the simple proof. □

Corollary 6.9 *For $t \to 0^+$, $\operatorname{tr} W(x, x, t)$ has the same asymptotic expansion as $trW^+(x, x, t)$.*

□

In many applications, the estimates 6.4 – 6.9 are very helpful to handle e^{-tD^2}, as we will see e.g. in section 8.

As we have seen already, the graded Laplace operator $(\Delta_0, \ldots, \Delta_n)$ is a special case of a generalized Dirac operator. Hence for all $e^{-t\Delta_q}$, $0 \leq q \leq n$, hold the estimates 6.4 – 6.9. Denote by $H(x, y, t) = H_{\Delta_0}(x, y, t)$ the (heat) kernel of $e^{t\Delta_0}$ and let $D = D(E, h, \nabla, \cdot)$ be the generalized Dirac operator associated to $(E, h, \nabla, \cdot) \longrightarrow (M^n, g)$. Then, according to (2.9)

$$D^2 = \Delta^E + \mathcal{R}.$$

Let $b := \inf_x b(x)$, $b(x) = \min_{\substack{\psi \in E_x \\ |\psi|=1}} \langle \psi, \mathcal{R}\psi \rangle$.

Theorem 6.10 *There holds*

$$|W(x, y, t)| \leq e^{-bt} H(x, y, t), \quad t > 0, x, y \in M. \tag{6.3}$$

□

(6.3) is called the semigroup domination principle as proved by Donnelly/Li in [34]. It asserts that estimates for the heat kernel of the Laplace operator acting on functions imply estimates for the heat kernel of any generalized Dirac operator over (M^n, g). For this reason, estimates for $H_{\Delta_0}(x, y, t)$ are of particular importance. In the sequel, we mainly concentrate on $\Delta = \Delta_0$. Moreover, the restriction to the case $p = 2$ seems to be artifical. We admit $1 < p < \infty$ and ask for the existence and properties of the heat semigroup $\{e^{-t\Delta}\}$ in $\Omega^{0,p,0}(E) \equiv \Omega^{0,p}(E) \equiv L_p(E)$.

The discussion of these problems and their solutions essentially depend on the functional spaces under consideration. We attack the problems for functions, tensors and differential forms with values in a Riemannian vector bundle E. Let us recall, for the existence and uniqueness questions, some simple facts from the theory of semigroups. If X is a Banach space, $x \in X$, $x \neq 0$, then by the Hahn-Banach theorem there exists an element $x^* \in X^*$ such that $||x^*|| = ||x||$ and $\langle x^*, x \rangle = ||x||^2$. We call such an element a normalized tangent functional. Taking $X = \Omega^{0,p}(T^q_r \otimes E)$, $X^* = \Omega^{0,p'}(T^q_r \otimes E)$ with $\frac{1}{p} + \frac{1}{p'} = 1$, $1 < p < \infty$, $0 \neq u \in X$, such a tangent functional can easily explicitely written down as $c|u|^{p-2} \cdot u$ with $c = |u|_p^{-\frac{p}{p'}}$. Consider $A : \mathcal{D}_A \longrightarrow X$, $\mathcal{D}_A \subset X$ being dense. The operator A is called dissipative if for every $x \in \mathcal{D}$ there exists a normalized tangent functional such that $\langle x^*, Ax\rangle \leq 0$. The closure of a dissipative operator is dissipative again. If X is a Hilbert space and $A : \mathcal{D}_A \longrightarrow X$ is symmetric and $\langle Ax, x\rangle \leq 0$ for all $x \in \mathcal{D}$ then A is dissipative.

A C^0-semigroup $\{T_t\}_{t\geq 0}$ of bounded linear operators $T_t \in L(X, X)$, X being a Banach space, is called a contraction semigroup if $|T_t|_{op} \leq 1$, $0 \leq t < \infty$. The infinitesimal generator of a semigroup $\{T_t\}_{t\geq 0}$ is defined as

$$A := s - \lim_{t\to 0} \frac{T_t - I}{t}.$$

Lemma 6.11 *A closed, densely defined operator $A : \mathcal{D}_A \longrightarrow X$ is the infinitesimal generator of a contraction semigroup if and only if A is dissipative and* $\mathrm{Im}\,(\mu - A) = X$ *for some $\mu > 0$. (cf.* [93]*).* □

The key for the existence of the heat semigroup $\{e^{-t\Delta}\}_{t\geq 0}$ is to establish the conditions of the preceeding lemma. For $p = 2$, $\{e^{-t\Delta}\}_{t\geq 0}$ exists by the spectral theorem as contraction semigroup. The interesting case is the case $p \neq 2$. We start with the Laplace operator acting on functions.

Lemma 6.12 *The operator $-\Delta$ with domain $\Omega^0_c = C^\infty_c$ resp. $\Omega^0_c(T^r_s)$ is dissipative on $\Omega^{0,p}$ resp. $\Omega^{0,p}(T^r_s \times E)$ for $1 < p < \infty$.*

The situation becomes much more difficult if we consider the Laplace operator acting on q-forms with values in E. At a first glance the curvature endomorphism destroys the estimates of the preceding lemmas. But they still remain valid if we assume $\mathcal{R} \geq 0$, i.e. $(\mathcal{R}u, u)_x \geq 0$ for all x.

Lemma 6.13 *Assume in $\Delta = \nabla^*\nabla + \mathcal{R}$, $\mathcal{R} \geq 0$, Δ acting in $\Omega^q(E)$. Then $-\Delta$ with domain $\Omega^q_c(E)$ is dissipative on $\Omega^{q,p}(E) = L_p(\Lambda(E))$ for $1 < p < 3$.* □

Lemma 6.14 *Suppose $1 < p \leq r < 3$, $\mathcal{R} \geq 0$ in $\Delta = \nabla^*\nabla + \mathcal{R}$, $u \in \Omega^{q,p}(E) + \Omega^{q,r}(E)$ and $\Delta u = \mu u$ for $\mu < 0$. Then u identically vanishes.* □

We give some examples for $\mathcal{R} \geq 0$.

1) If $q = 1$, the Ricci curvature on M^n is nonnegative and $R^E = 0$, then according to the definition of $\mathcal{R}$ we have $\mathcal{R} \geq 0$. For ordinary 1-forms the conditions Ricci curvature ≥ 0 and $\mathcal{R} \geq 0$ are equivalent.

2) A sufficient condition for $q \geq 1$ and ordinary forms (i.e. E the trivial line bundle) is given by the nonnegativity of the curvature operator R^{op}. This we will shortly indicate. $R = R^M$ induces a symmetric linear operator $R^{op} : \Lambda^2 TM \longrightarrow \Lambda^2 TM$ in the space of bivectors, called the curvature operator R^{op} and characterized by $(R^{op}(X\wedge Y), Z\wedge W)_x = (R(X, Y)W, Z)_x$. If $R^{op} \geq \lambda$ then $(\mathcal{R}u, u)X \geq \lambda q(n-q)|u|^2_x$ ([63], p. 264), in particular $R^{op} \geq 0$ implies $\mathcal{R} \geq 0$.

3) Of particular interest are those cases where sectional curvature $K \geq 0$ implies $R^{op} \geq 0$ and hence $\mathcal{R} \geq 0$.

If $f : M \longrightarrow \mathbb{R}^{n+2}$ is an isometric immersion, $n = 2k$, and M^n is open, complete, oriented, with sectional curvature $K \geq 0$ and at some point $x \in M$ $K > 0$, then there holds $\mathcal{R} \geq 0$ (cf. [10]). A second class is given by manifolds with pure curvature operator. M^n has pure curvature operator if for each $x \in M$ there exists an orthonormal frame $(e_1, \ldots, e_n)$ in T_xM such that $R^{op}(e_i \wedge e_j) = K_{ij}(e_i \wedge e_j)$, K_{ij} = sectional curvature of the plane spanned by e_i, e_j. For a manifold with sectional curvature ≥ 0 and pure curvature operator there holds $\mathcal{R} \geq 0$ (cf.[10]).

An open manifold which belongs to all three classes is the rotating parabola in $\mathbb{R}^3$.

From 6.14 we immediately obtain

Corollary 6.15 *If the curvature operator for an open complete manifold (M^n, g) is nonnegative and $l < p$, $r < 3$, then $\Delta u = \mu u$, $u \in \Omega^{q,p} + \Omega^{q,r}$, and $\mu < 0$ imply $u = 0$. In particular, this holds for the examples 1), 2), 3).* □

Now we come to the heat semigroup $\{e^{-t\Delta}\}_{t\geq 0}$. For $p = 2$ the operator $e^{-t\Delta}$ is well defined through the spectral theorem,

$$e^{-t\Delta} = \int_0^\infty e^{-t\lambda} dE_\lambda,$$

with Δ acting on functions or tensors with values in E or q-forms with values in E, respectively.

It is, a priori, far from being clear that this semigroup can be defined on L_p and there has to be strongly continuous or additionally contractive. In general, this is not true, as an example presented in [114] by Strichartz shows. We start with the simplest case of functions and follow here [114].

Theorem 6.16 *Let $\{e^{-t\Delta}\}_{t\geq 0}$ be the heat semigroup acting in $\Omega^{0,2} = L_2$ defined by the spectral theorem. Then there exists a heat kernel $H(x, y, t)$ for $e^{-t\Delta}$ satisfying the following conditions.*

1) $H(x, y, t) \in C^\infty(M \times M \times \mathbb{R}_+)$.

2) $H(x, y, t) = H(y, x, t)$.

3) $\int H(x, y, t)\,\mathrm{dvol}_y \leq 1$ *for every x and $t > 0$, such that*

$$e^{-t\Delta}u(x) = \int H(x, y, t)u(y)\,\mathrm{dvol}_y, \quad u \in L_2. \tag{6.4}$$

4) $|e^{-t\Delta}u|_p \leq |u|_p$ *for all $t > 0$ and $u \in \Omega^{0,2} \cap \Omega^{0,p}$, $1 \leq p \leq \infty$.*

5) $|e^{-t\Delta}u - u|_p \underset{t\to 0}{\longrightarrow} 0$ *if $1 \leq p < \infty$.*

6) $\frac{\partial}{\partial t}e^{-t\Delta}u = -\Delta e^{-t\Delta}u$ *for all $u \in \Omega^{0,2} = L_2$.* (6.5)

7) *If one defines $e^{-t\Delta}u$ for $u \in \Omega^{0,p}$, $1 \leq p \leq \infty$, by (6.4), then (6.5) keeps its validity.*

8) *The semigroup is uniquely determined for $1 < p < \infty$ in the following sense: If $\{Q_t\}_{t\geq 0}$ is any strongly continuous contractive semigroup on $\Omega^{0,p}$, $1 < p < \infty$ such that $\{Q_t\}_{t\geq 0}$ satisfies the heat equation $\frac{\partial}{\partial t}Q_t u = -\Delta Q_t u$, then $Q_t = e^{-t\Delta}$.* □

A fundamental solution of the heat operator $\frac{\partial}{\partial t} + \Delta$ acting on functions is a continuous function $h : M \times M \times]0, \infty[\longrightarrow \mathbb{R}$ which is C^2 with respect to x, C^1 with respect to t, and which satisfies the heat equation

$$\left(\frac{\partial}{\partial t} + \Delta_x\right) h = 0, \quad \lim_{t\to 0} h(\cdot, y, t) = \delta_y. \tag{6.6}$$

Suppose that $u : M \times [0, \infty[\longrightarrow \mathbb{R}$ is continous, C^2 in x, C^1 in t for $t > 0$. The function u is called a solution of the Cauchy problem with initial data u_0 if $\left(\frac{\partial}{\partial t} + \Delta\right) u = 0$ on $M \times]0, \infty[$, $u(x, 0) = u_0(x)$ on M.

A smooth parametrix for $\Delta + \frac{\partial}{\partial t}$ is a real valued function $P = P(x, y, t) \in C^\infty(M \times M \times \mathbb{R}_+)$ such that $\lim_{t\to 0} P(x, y, t) = \delta_x$ and for $t \to 0$ $\left(\Delta_x + \frac{\partial}{\partial t}\right) P = O(t^k)$ for all $k > 0$

(often one requires this for $k > \frac{n}{2} + 2$). At this stage, one would ask for the existence of a fundamental solution, the unique solvability of the Cauchy problem, the existence of a parametrix and how these are related to the heat kernel for functions above. Essential results towards this direction were obtained by Dodziuk [23], [33].

Theorem 6.17 *Assume that (M^n, g) is complete with Ricci curvature bounded from below. Then the bounded solutions of the initial-value problem are uniquely determined by their initial data ([33], p. 183–185).* □

Theorem 6.18 *For an arbitrary Riemannian manifold there exists a smooth fundamental solution h to the heat operator $\frac{\partial}{\partial t} + \Delta$ which satisfies*

1) $h(x, y, t) \geq 0$, $h(x, y, t) = h(y, x, t)$ *for* $t > 0$, $x, y \in M$,

2) $\left(\Delta_x + \frac{\partial}{\partial t}\right) h = 0$,

3) $\int h(x, z, t) h(z, y, s) \, \mathrm{dvol}_z = h(x, y, t + s)$ *for* $t, s > 0$, $x, y \in M$,

4) $h(x, y, t) = \sup\limits_{K \subset M} h_K(x, y, t)$.

For the proof, one considers an exhaustion $K_1 \subset K_2 \subset \cdots$, $\bigcup K_i = M$, of M by compact submanifolds with smooth boundary and the corresponding heat kernels $h_i = h_{D,i}(x, y, t)$ for the heat operator on K_i with Dirichlet boundary value conditions. Then one can show that $\lim\limits_{i \to \infty} h_i = h$ exists in an appropriately strong sense and that $h(x, y, t)$ has the desired properties. The proof of the convergence essentially uses a Harnack type inequality for parabolic equations. Furthermore, $h \geq 0$ is proven by a maximum principle. For details, we refer to [33].

The following theorem is a corollary to 6.17 and 6.18.

Theorem 6.19 *Assume that (M^n, g) is complete with Ricci curvature bounded from below. Then the heat operator acting on functions has a unique fundamental solution $h(x, y, t)$, and $h(x, y, t)$ satisfies the conservation law*

$$\int\limits_M h(x, y, t) \, \mathrm{dvol}_y \leq 1.$$

□

We come now to the relation between the fundamental solution $h(x, y, t)$ and the heat kernel $H(x, y, t)$ for complete manifolds.

Theorem 6.20 *If (M^n, g) is complete, then $h(x, y, t) = H(x, y, t)$.* □

We see from the theorems 6.16 and 6.18 that in the complete case any solution $u(x, t)$ of the Cauchy problem $\left(\frac{\partial}{\partial t} + \Delta\right) u(x, t) = 0$, $u(x, 0) = u_0(x)$ is given by

$$u(x, t) = \int\limits_M h(x, y, t) u_0(y) \, \mathrm{dvol}_y(g), \quad t > 0.$$

In [26] the concept of a compactly supported parametrix P for $\frac{\partial}{\partial t} + \Delta$ enters into the proof of existence of the heat kernel, i.e. if x runs through a relatively compact neighborhood $U \subset M$, then the supports of $P(x, y, t)$ are contained in some compact set $C \subset M$. The existence of such a parametrix can be assured as in [26]. Theorem 6.17 expresses the uniqueness for the initial-value problem in the class of bounded functions under certain curvature assumptions. Without these assumptions, the uniqueness fails to hold. There exist complete simply connected manifolds of negative sectional curvature tending to minus infinity sufficiently rapidly such that $e^{-t\Delta}1 \neq 1$ (cf.[9]).

Nevertheless, passing over to $\Omega^{0,p}$, $1 < p < \infty$, the uniqueness statement holds without any curvature assumption if one bounds the growth with respect to t.

Theorem 6.21 *Assume that $u(x,t)$ satisfies the heat equation $\frac{\partial}{\partial t} + \Delta u = 0$ in $M \times \mathbb{R}_+$, $u(\cdot,t) \in \Omega^{0,p}$ for each $t > 0$ and $|u(\cdot,t)|_p \leq a \cdot e^{bt}$ for some a, b and some p, $1 < p < \infty$. Then there exists a uniquely determined $u_0 \in \Omega^{0,p}$ such that $u = e^{-t\Delta}u_0$. More generally, if $1 < p \leq r < \infty$ and $u(\cdot,t) \in \Omega^{0,p} + \Omega^{0,r}$ with $|u(\cdot,t)| \leq a \cdot e^{bt}$ ($|\,|$ in the sense of $\Omega^{0,p} + \Omega^{0,r}$), then there exists a uniquely determined $u_0 \in \Omega^{0,p} + \Omega^{0,r}$ such that $u = e^{-t\Delta}u_0$.*

For manifolds with infinite volume, there is a simple consequence concerning the behavior of heat equation solutions for $t \to \infty$.

Theorem 6.22 *Suppose that* vol $(M) = \infty$, $1 < p < \infty$, $u_0 \in \Omega^{0,p} = L_p$. *Then*

$$\lim_{t\to\infty} |e^{-t\Delta}u_0|_p = 0.$$

Proof The premise vol $(M) = \infty$ immediately implies that there are no L_2-harmonic functions. Thus for $p = 2$, the result follows by the spectral theorem and the fact that $e^{-tD^2} \underset{t\to\infty}{\longrightarrow} P_{\ker D}$ in the strong operator topology. If $p \neq 2$, $1 < p < \infty$, it suffices to show the theorem for a dense subset $\subset \Omega^{0,p}$ since the $e^{-t\Delta}$ are uniformly bounded. Consider $1 < p < 2$, $u_0 \in \Omega^{0,2} \cap \Omega^{0,1}$, s with $\frac{1}{p} = \frac{s}{2} + \frac{1-s}{1}$ and use $|e^{-t\Delta}u_0|_p \leq |e^{-t\Delta}u_0|_2^s \cdot |e^{-t\Delta}u_0|_1^{l-s}$ (Riesz' convexity). The uniform boundedness of $|e^{-t\Delta}u_0|_1$ implies the result. If $2 < p < \infty$, we use an analogous argument replacing $|\,|_1$ by ${}^{\infty}|\,|$. □

Let us add without proof the following result of Yau [120].

Theorem 6.23 *Assume that (M^n, g) is (as always here) complete with Ricci curvature bounded from below. If u_0 is continuous, bounded and $\lim_{x\to\infty} u_0(x) = 0$, then for every $t > 0$*

$$\lim_{x\to\infty} e^{-t\Delta}u_0 = 0.$$

□

We continue with a few remarks concerning comparison theorems and estimates for the heat kernel. Denote by M_K^n the simply connected space form of constant curvature K, i.e. for $K > 0$ $M_K^n = S_{1,K}^n$, for $K = 0$ $M_0^n = \mathbb{R}^n$, for $K < 0$ $M_K^n = H_K^n$. Then for the injectivity radius r_{inj} there holds

$$r_{\text{inj}}(M_K^n) = \begin{cases} \frac{\pi}{\sqrt{K}} & \text{for } K > 0 \\ \infty & \text{for } K \leq 0. \end{cases}$$

If $H_{K,\delta}(x,y,t)$ denotes the Dirichlet heat kernel for a ball $B_\delta(x,K) \subset M_K^n$ with $\delta < r_{inj}$ then there exists a function $\mathcal{E}_K : [0,\delta[\times]0,\infty[\longrightarrow \mathbb{R}$ such that $H_{K,\delta}(x,y,t) = \mathcal{E}_{K,\delta}(d(x,y),t)$ (cf. [23], [26]).

Now we consider an arbitrary noncompact Riemannian manifold (M^n, g), not necessarily complete, a relative compact geodesic ball $B_\delta(x) \subset M^n$ with $\delta < r_{\text{inj}}(M_K^n)$ and the

Dirichlet heat kernel $H_\delta(x,y,t)$ for $B_\delta(x)$. Since $\delta < r_{inj}(M), r_{inj}(M^n_K)$, it is possible to transplant the geodesic ball from M^n to M^n_K and inverse.

Theorem 6.24 *If the sectional curvature of $B_\delta(x)$ is $\leq K$, then*

$$H_\delta(x,y,t) \leq \mathcal{E}_{K,\delta}(d(x,y),t).$$

If, on the other hand, all Ricci curvatures of $B_\delta(x)$ are $\geq (n-1)K$, then

$$H_\delta(x,y,t) \geq \mathcal{E}_{K,\delta}.$$

For the proof we refer to [23], [26]. □

As a corollary to 6.24, one gets

Theorem 6.25 *Suppose that (M^n,g) is complete with Ricci curvature bounded from below by $(n-1)K$, $H_K(x,y,t) = E_K(d(x,y),t)$ being the heat kernel of M^n_K. Then the heat kernel $H(x,y,t)$ of M obeys*

$$H(x,y,t) \geq E_K(d(x,y),t)$$

for every $(x,y,t) \in M \times M \times]0,\infty[$. □

Theorems 6.24 and 6.25 can be generalized as follows. Denote for $x_0 \in (M^n,g)$ and geodesic polar coordinates (u,r) by $(u,r) \longrightarrow m(u,r)$ the mean curvature function at the point (u,r) of $\partial B_r(x_0)$ with $\partial B_r(x_0) \cap C_{x_0}$ deleted, C_{x_0} the cut locus of x_0. A Riemannian manifold $\mathcal{M}$ is called an open model if

a) for some point $z_0 \in \mathcal{M}$ and $0 < R \leq \infty$, $\mathcal{M}$ is diffeomorphic to $B_R(z_0)$ by means of the exponential map $\exp_{z_0} : B_R(0) \longrightarrow B_R(z_0)$ and

b) for all $r < R$, the mean curvature of the distance sphere $\partial B_r(z_0)$ is constant on $\partial B_r(z_0)$, denoted by $m(r)$.

Examples are e.g. rotationally symmetric metrics.

Theorem 6.26 *Let $\mathcal{M}$ be an open model. Then its heat kernel $H(\tilde{x},\tilde{y},t)$, $\tilde{x},\tilde{y} \in \mathcal{M}$, depends only on $r = d(\tilde{x},\tilde{y})$, i.e. $H(\tilde{x},\tilde{y},t) = H(d(\tilde{x},\tilde{y}),t)$.* □

Theorem 6.27 *Let (M^n,g) be open, complete, $\mathcal{M}^n$ an open model and suppose $m(u,r) \leq m(r)$, $0 < r \leq R$. Then there holds*

$$H_{\mathcal{M}}(d(x,y),t) \leq H_M(x,y,t), \quad x,y \in M, \quad t > 0.$$

Equality holds if and only if (M^n,g) is isometric to $\mathcal{M}^n$ and $m(u,r) = m(r)$ for each r.

We refer to [26] for the proof. □

Another estimate from below is given by the following theorem of Li/Yau.

Theorem 6.28 *Let (M^n,g) be complete with* $\operatorname{Ric}_M \geq 0$*. Then for all $\varepsilon > 0$, there exists a constant $C(\varepsilon)$ such that*

$$\begin{aligned} H(x,y,t) &\geq C(\varepsilon)^{-1}\mathrm{vol}\,(B_{\sqrt{t}}(x))^{-1}e^{-d^2(x,y)/(4+\varepsilon)t} \\ H(x,y,t) &\geq C(\varepsilon)^{-1}\mathrm{vol}\,(B_{\sqrt{t}}(x))^{-\frac{1}{2}}\mathrm{vol}\,(B_{\sqrt{t}}(y))^{-\frac{1}{2}}e^{-d^2(x,y)/(4+\varepsilon)t}, \end{aligned}$$

where $C(\varepsilon) \underset{\varepsilon\to 0}{\longrightarrow} \infty$. □

A general upper bound has been established by Cheng, Li and Yau (cf. [27]).

Theorem 6.29 *Let (M^n, g) be complete. Then, for all $\beta > 1$, $T > 0$ and $x \in M$, there exists a constant $C = C(\beta, T, x)$, s.t.*

$$\int\limits_{M \setminus B_R(x)} H(x,y,t)^2 \,\mathrm{dvol}_y(g) \leq C \cdot t^{-\frac{n}{2}} e^{-\frac{R^2}{2\beta t}}$$

for all $t \in [0,T]$ and $R > 0$, where $C \underset{\beta \to 0}{\longrightarrow} \infty$. □

Theorem 6.30 *Let (M^n, g) be complete with bounded sectional curvature. Then for all $\alpha > 4$, $T > 0$ and $x \in M$, there exists a constant $C' = C'(\alpha, T, x)$ s.t.*

$$H(x,y,t) \leq C' t^{-\frac{n}{2}} e^{-\frac{d^2(x,y)}{\alpha t}}, \quad t \in]0,T], \quad y \in M.$$

This is contained in [27]. □

In the case of order zero bounded geometry Varopoulos established an upper bound independent of x and y (cf. [119]).

Theorem 6.31 *Suppose (M^n, g) satisfies (B_0) and (I). Then for all $0 < \varepsilon < 0,1$, there exist constants $C_1, C_2 > 0$, s.t.*

$$\sup_{x,y \in M} H(x,y,t) \leq \min\{C_1 t^{-\frac{1}{2}+\varepsilon}, C_2 t^{-\frac{1}{2}} (\log t)^{1+\varepsilon}\}, \quad t > 1.$$

□

In the case of nonnegative Ricci curvature an upper bound is given by the following theorem of Li/Yau.

Theorem 6.32 *Suppose (M^n, g) is complete and has nonnegative Ricci curvature. Then, for all $0 < \varepsilon < 1$, there exists a constant $C(\varepsilon)$ such that*

$$H(x,y,t) \leq C(\varepsilon) \mathrm{vol}\, (B_{\sqrt{t}}(x))^{-1} e^{-\frac{d^2(x,y)}{(4+\varepsilon)t}}, \quad x,y \in M, \quad t > 0,$$

where $C(\varepsilon) \underset{\varepsilon \to 0}{\longrightarrow} \infty$. □

An upper bound independent of x, y implies automatically exponentially decay, as it has been proven by Davies in [30].

Theorem 6.33 *If the heat kernel of (M^n, g) satifies*

$$H(x,y,t) \leq a t^{-\frac{n}{2}}, \quad x,y \in M, \quad t > 0,$$

for some positive constant a then for all $\delta > 0$ there exists a constant $C(\delta) > 0$, s.t.

$$H(x,y,t) \leq C(\delta) t^{-\frac{n}{2}} e^{-\frac{d^2(x,y)}{4(1+\delta)t}}, \quad x,y \in M, \quad t > 0.$$

□

Remark 6.34 The assumption of 6.33 is equivalent to

$$|f|_{\frac{2n}{n-2}} \leq a \langle \Delta f, f \rangle$$

for all $f \geq 0$, $f \in C_c^\infty$. This latter condition is satisfied in the case Ric $\geq -c$, $c > 0$ and (I).

We refer to [67], p. 179 for a proof. □

A strong mean to get off-diagonal estimates from on-diagonal estimates is given by

Proposition 6.35 *Let (M^n, g) be an arbitrary Riemannian manifold. Suppose $h(x,x,t) \leq f(t)$. Then*

$$h(x,y,t) \leq f(t), \quad x, y \in M.$$

Proof We have from theorem 6.18

$$h(x,x,t) = \int_M h^2(x,z,\frac{t}{2}) \,\mathrm{dvol}_z(g).$$

Then the semigroup identity and Schwarz inequality yield

$$\begin{aligned} h(x,y,t) &= \int_M h(x,z,\frac{t}{2}) h(z,y,\frac{t}{2}) \,\mathrm{dvol}_y(g) \\ &\leq \left(\int_M h^2(x,z,\frac{t}{2}) \,\mathrm{dvol}_z(g) \right)^{\frac{1}{2}} \left(\int_M h^2(y,z,\frac{t}{2}) \,\mathrm{dvol}_y(g) \right)^{\frac{1}{2}} \end{aligned}$$

from where $h(x,y,t) \leq \sqrt{h(x,x,t)h(y,y,t)} \leq f(t)$. □

Define for $D > 0$ the weighted integral of the heat kernel

$$E_D(x,t) = \int_M h^2(x,z,t) e^{\frac{d^2(x,z)}{Dt}} \,\mathrm{dvol}_z(g).$$

We state without proof (cf. [67])

Theorem 6.36 *For any manifold (M^n, g), $E_D(x,t)$ is finite for all $D > 2$, $t > 0$, $x \in M$. Moreover, $E_D(x,t)$ is non-increasing in t.* □

Lemma 6.37 *For any $D > 0$, all $x, y \in M$, $t > 0$*

$$h(x,y,t) \leq \sqrt{E_D(x,\frac{t}{2}) E_D(y,\frac{t}{2})} e^{-\frac{d^2(x,y)}{2Dt}}.$$

Proof Denote for $x, y, z \in M$, $\alpha = d(y,z)$, $\beta = d(x,z)$, $\gamma = d(x,y)$. Then $\alpha^2 + \beta^2 \geq \frac{1}{2}\gamma^2$.

$$\begin{aligned} h(x,y,t) &= \int_M h(x,z,\frac{t}{2}) h(y,z,\frac{t}{2}) \,\mathrm{dvol}_z(g) \\ &\leq \int_M h(x,z,\frac{t}{2}) e^{\frac{\beta^2}{Dt}} h(y,z,\frac{t}{2}) e^{\frac{\alpha^2}{Dt}} e^{-\frac{\gamma^2}{2Dt}} \,\mathrm{dvol}_z(g) \\ &\leq \left(\int_M h^2(x,z,\frac{t}{2}) e^{\frac{2\beta^2}{Dt}} \,\mathrm{dvol}_z(g) \right)^{\frac{1}{2}} \left(\int_M h^2(y,z,\frac{t}{2}) e^{\frac{2\alpha^2}{Dt}} \,\mathrm{dvol}_y(g) \right)^{\frac{1}{2}} \cdot e^{-\frac{\gamma^2}{2Dt}} \end{aligned}$$

$$= \sqrt{E_D(x, \frac{t}{2}) E_D(y, \frac{t}{2})} \cdot e^{-\frac{d^2(x,y)}{2Dt}}.$$

□

Theorem 6.38 *Suppose for some $x \in M$ and all $t > 0$*

$$h(x, x, t) \leq \frac{C}{f(t)},$$

where $f(t)$ is increasing and $\frac{f(as)}{f(s)} \leq A \frac{f(at)}{f(t)}$ for all $0 < s < t$ and some $A \geq 1$, $a > 1$. Then for all $D > 2$ and $t > 0$

$$E_D(x, t) \leq \frac{C'}{f(\varepsilon t)}$$

for some $\varepsilon > 0$ and $C' > 0$.

We refer to [67] for the proof. □

Corollary 6.39 *Assume for some $x, y \in M$ and all $t > 0$*

$$h(x, x, t) \leq \frac{C}{f(t)}, \quad h(y, y, t) \leq \frac{C}{g(t)},$$

where f and g satisfy the hypotheses of theorem 6.38. Then, for all $t > 0$, $D > 2$ and some $\varepsilon > 0$

$$h(x, y, t) \leq \frac{C'}{\sqrt{f(\varepsilon t) g(\varepsilon t)}} e^{-\frac{d^2(x,y)}{2Dt}}.$$

This follows immediately from lemma 6.37 and theorem 6.38. □

The volume growth and on-diagonal estimates are strongly related as the following theorem shows.

Theorem 6.40 *Suppose (M^n, g) is complete, and that for some $x \in M$ and all $r > 0$*

$$\text{vol}\ (B_{2r}(x)) \leq C \text{vol}\ (B_r(x))$$

and for all $t > 0$

$$H(x, x, t) = h(x, x, t) \leq \frac{C}{\text{vol}\ (B_{\sqrt{t}}(x))}.$$

Then for all $t > 0$,

$$h(x, x, t) \geq \frac{c}{\text{vol}\ (B_{\sqrt{t}}(x))},$$

where $c = c(C)$. □

The decay of $h(x, y, t)$ is strongly related with the bottom of the spectrum.

Let $\Omega \underset{\neq}{\subset} (M^n, g)$ be a bounded connected domain, $\Delta_D(\Omega) = \Delta_{0,D}(\Omega)$ the Dirichlet Laplacian of Ω. According to theorem 4.18, $\Delta_D(\Omega)$ has purely discrete point spectrum

$0 \le \lambda_1 \le \lambda_2 \le \cdots \to \infty$, the sequence $\{\lambda_k\}_k$ obeys Weyl's asymptotic formula $\lambda_k(\Omega) \sim c_n \left(\frac{k}{\mathrm{vol}\ (\Omega)}\right)^{\frac{2}{n}}$ for $k \longrightarrow \infty$ and $h_{D,\Omega}(x,y,t)$ has the expression

$$h_{D,\Omega}(x,y,t) = \sum_{k=1}^{\infty} e^{-t\lambda_k(\Omega)} \varphi_k(x)\varphi_k(y). \tag{6.7}$$

Moreover, under our assumptions, $\lambda_1(\Omega) > 0$ with mulitplicity 1, i.e. $\lambda_2(\Omega) > \lambda_1(\Omega)$. This yields together with (6.7)

$$h_{D,\Omega}(x,y,t) \sim e^{-\lambda_1(\Omega)} \varphi_1(x)\varphi_1(y) \quad \text{for } t \longrightarrow \infty.$$

If $\lambda_0(M) = \inf \sigma(\Delta_0(M))$ is the bottom of the spectrum of $\Delta_0(M)$, it follows immediately from the Rayleigh-Ritz characterization of $\lambda_0(M)$

$$\lambda_0(M) = \inf_{\Omega \subset\subset M} \lambda_1(\Omega).$$

We recall the following estimate from [67].

Theorem 6.41 *If $D > 2$ and $\lambda_0(M) > 0$ then for $t > t_0$*

$$h(x,y,t) \le \sqrt{E_D(x, \frac{t_0}{2}) E_D(y, \frac{t_0}{2})} e^{\lambda_0(M)t_0} e^{-\lambda_0(M)t - \frac{d^2(x,y)}{2Dt}}. \tag{6.8}$$

□

Proposition 6.42 *Given $x \in (M^n, g)$, $\varepsilon > 0$, there exists $c = c_x > 0$ s.t.*

$$h(x,x,t) \ge c_x e^{-(\lambda_0(M)+\varepsilon)t}, \quad t > 0. \tag{6.9}$$

Proof Choose $\Omega \subset\subset M$, $x \in \Omega$, s.t. $\lambda_1(\Omega) \le \lambda_0(M) + \varepsilon$. According to the proof of theorem 6.18, $h(x,x,t) \ge h_{D,\Omega}(x,x,t)$, and we have

$$h_{D,\Omega}(x,x,t) = \sum_{k=1}^{\infty} e^{-\lambda_k(\Omega)t} \varphi_k^2(x) \ge e^{-\lambda_1(\Omega)t} \varphi_1^2(x) \ge e^{-(\lambda_0(M)+\varepsilon)t} \varphi_1^2(x). \tag{6.10}$$

□

Corollary 6.43 *For all $x \in M$*

$$\lim_{t \to \infty} \frac{\log h(x,x,t)}{t} = -\lambda_0(M).$$

This follows immediately from (6.8) and (6.9). □

Another asymptotic for $t \to 0$ or $t \to \infty$, respectively, is given by

Theorem 6.44 *Let (M^n, g) be complete.*

a) *There holds $\lim_{t \to 0} -4t \log h(x,y,t) = d^2(x,y)$, $x, y \in M$.*

b) *Assume that (M^n, g) has nonnegative Ricci curvature* Ric ≥ 0, *and there exists $x_0 \in M$ and $c > 0$ s.t. $\lim_{r \to \infty} (\mathrm{vol}\ (B_r(x_0))/r^n) = c$. Then $\lim_{t \to \infty} \mathrm{vol}\ (B_{\sqrt{t}}(x_0)) h(x,y,t) = \mathrm{vol}_{\mathbb{R}^n}(B_1(0))(4\pi)^{-\frac{n}{2}}$, where $B_1(c) \subset \mathbb{R}^n$.*

We refer to [27] for the proof of a). □

We see from the expansion (6.10) that

$$h_{D,\Omega}(x,x,t) \sim e^{-\lambda_1(\Omega)t}\varphi_1^2(x) \quad \text{for } t \to \infty.$$

If one performs a reasonable limit $\Omega \longrightarrow M$ (e.g. by a controled exhaustion) then $\lambda_1(\Omega) \longrightarrow \lambda_0(M)$ and one would expect

$$h(x,x,t) \sim e^{-\lambda_0(M)t} \quad \text{for } t \to \infty.$$

In fact, one can easily deduce from (6.4)

$$h(x,x,t) \leq e^{-\lambda_0(M)(t-t_0)}h(x,x,t_0). \tag{6.11}$$

This is a good estimate if $\lambda_0(M) > 0$. If $\lambda_0(M) = 0$ then it is valueless. But the case $\lambda_0(M) = 0$ often appears as we will see soon. Therefore one has to sharpen considerations and to consider the rate of convergence $\lambda_1(\Omega) \underset{\Omega\to M}{\longrightarrow} \lambda_0(M)$. This rate will affect the rate of convergence.

The key for this will be Faber-Krahn type inequalities

$$\lambda_1(\Omega) \geq \Lambda(\text{vol } (\Omega)),$$

where Λ is a positive decreasing function on $]0,\infty[$, called a Faber-Krahn function of M. For $M^n = \mathbb{R}^n$, $\Lambda(v) = cv^{-\frac{2}{n}}$ is a Faber-Krahn function.

Theorem 6.45 *Assume that (M^n, g) admits Faber-Krahn function Λ and define $f(t)$ by*

$$t = \int_0^{f(t)} \frac{dv}{v\Lambda(v)}, \tag{6.12}$$

where we assume that the integral (6.12) converges at 0. Then for all $t > 0$, $x \in M$, $\varepsilon > 0$

$$h(x,x,t) \leq \frac{2\varepsilon^{-1}}{f((1-\varepsilon)t)}.$$

We refer to [67] for the proof. □

There is a localized version of theorem 6.45 with far-reaching consequences.

Theorem 6.46 *Suppose that for some $x \in M$ and $r > 0$ the following Faber-Krahn inequality holds. For any precompact open $\Omega \subset B_r(x)$ there holds with $a > 0$, $n > 0$*

$$\lambda_1(\Omega) \geq a\text{vol } (\Omega)^{-\frac{2}{n}}. \tag{6.13}$$

Then for any $t > 0$

$$h(x,x,t) \leq \frac{Ca^{-\frac{n}{2}}}{(\min\{t,r^2\})^{\frac{n}{2}}}, \quad C = C(n). \tag{6.14}$$

We refer to [67] for the proof. □

Corollary 6.47 *Suppose for all $x \in M$ and some $r > 0$ the Faber-Krahn inequality (6.13), $\Omega \subset B_r(x)$. Then for any $D > 2$ and all $x, y \in M$, $t > 0$,*

$$h(x,y,t) \leq \frac{Ca^{-\frac{n}{2}}}{(\min\{t, r^2\})^{\frac{n}{2}}} e^{-\frac{d^2(x,y)}{2Dt}}, \quad C = C(n, D).$$

This is a combination of (6.14) and corollary 6.39. □

We finish at this point the general estimates for the heat kernel on functions. Let us now consider the heat operator acting on differential forms or tensors with values in a vector bundle E.

Some assertions of theorem 6.16 can be carried over to this case. Additional curvature assumptions provide better and better properties for $e^{-t\Delta}$ and $\{e^{-t\Delta}\}$. We refer to [47] and [115].

Theorem 6.48 *Denote by $\{e^{-t\Delta}\}$ the heat semigroup acting on $\Omega^{0,2}(T^r_s \otimes E)$. Then $|e^{-t\Delta}u|_p \leq |u|_p$ for all $u \in \Omega^{0,p}(T^r_s \otimes E) \cap \Omega^{0,2}(T^r_s \otimes E)$ and $1 < p < \infty$. Therefore $\{e^{-t\Delta}\}$ extends to a contraction semigroup on $\Omega^{0,p}(T^r_s \otimes E)$, $1 \leq p \leq \infty$. Further, $e^{-t\Delta}u$ satisfies the heat equation $\frac{\partial}{\partial t}e^{-t\Delta}u = -\Delta e^{-t\Delta}u$ for $u \in \Omega^{0,p}(T^r_s \otimes E)$, and $\{e^{-t\Delta}\}$ is the unique semigroup exhibiting these properties for $1 < p < \infty$.*

In a similar manner we conclude for q-forms with values in E.

Theorem 6.49 *Assume that (M^n, g) is open, complete and that $\mathcal{R} \geq 0$ for the endomorphism $\mathcal{R}$ in $\Delta = \nabla^*\nabla + \mathcal{R}$, acting on q-forms with values in E. If $\{e^{-t\Delta}\}_{t\geq 0}$ denotes the heat semigroup on $\Omega^{q,p}(E)$, then $|e^{-t\Delta}u|_p \leq |u|_p$ for all $u \in \Omega^{q,p}(E) \cap \Omega^{q,2}(E)$ and $1 \leq p \leq \infty$. Therefore $\{e^{-t\Delta}\}_{t\geq 0}$ extends to a contraction semigroup on these $\Omega^{q,p}(E)$. $e^{-t\Delta}u$ satisfies the heat equation $\frac{\partial}{\partial t}e^{-t\Delta}u = -\Delta e^{-t\Delta}u$ for all $u \in \Omega^{q,p}(E)$, and $\{e^{-t\Delta}_{t\geq 0}\}$ is the unique semigroup with these properties for $1 < p < \infty$.*

We refer to [115] for the proofs of 6.48 and 6.49. □

Corollary 6.50 *The assertions of 6.49 are valid for the classes 1) – 4) after 6.14.* □

A simple example is the heat semigroup acting on 1-forms on the rotating parabola.

Theorem 6.48 and 6.49 immediately imply

Theorem 6.51 *Assume that (M^n, g) is open, complete $(E, h) \longrightarrow M$ a Riemannian vector bundle, $1 < p < \infty$. Then the initial value problem $\left(\frac{\partial}{\partial t} + \Delta\right) u = 0$ on $M \times]0, \infty[$, $u(x, 0) = u_0(x)$ on M is solvable in the following cases:*

a) *$u(\cdot, t) \in \Omega^{0,p}(T^r_s \otimes E)$ and $u_0 \in \Omega^{0,p}(T^r_s \otimes E)$.*

b) *$u(\cdot, t) \in \Omega^{q,p}(E)$ and $u_0 \in \Omega^{q,p}(E)$ and $\mathcal{R} \geq 0$.*

The remaining open question is the uniqueness which is partially answered by

Theorem 6.52 *Assume that (M^n, g) is open, complete, $(E, h) \longrightarrow M$ a Riemannian vector bundle, $1 < p < \infty$, $u(x, t)$ a solution of the (homogeneous) heat equation with $u(\cdot, t) \in \Omega^{0,p}(T^r_s \otimes E)$ or $u(\cdot, t) \in \Omega^{q,p}(E)$ and $\mathcal{R} \geq 0$, respectively. Assume further $|u(\cdot, t)|_p \leq a \cdot e^{bt}$. Then there exists a uniquely determined $u_0 \in \Omega^{0,p}(T^r_s \otimes E)$ or $u_0 \in \Omega^{q,p}(E)$, respectively, such that $u = e^{-t\Delta}u_o$.* □

Another nice uniqueness theorem has been proved by Dodziuk.

Theorem 6.53 *Assume that (M^n, g) is open, complete with Ricci curvature bounded from below and that $\mathcal{R}_q \geq 0$ in $\Delta = \nabla^*\nabla + \mathcal{R}_q$ acting on Ω^q, $q > 0$. Then every bounded solution of the initial value problem of the heat equation in Ω^q is uniquely determined by its* *initial values. Moreover, if the initial value vanishes at infinity, then the solution vanishes at infinity for every $t > 0$ (cf.* [33]*).* □

Remarks 6.54 **1)** In [84] Lohoue has shown that the contraction semigroup property for $\{e^{-t\Delta}\}_{t\geq 0}$ acting on functions implies this property for $\{e^{-t\Delta}\}_{t\geq 0}$ acting on usual tensors or differential forms, if $e^{-t\Delta}$ has a heat kernel and sufficiently bounded geometry.

2) Strichartz has shown in [110] that on the hyperbolic plane H^2_{-1} for $q = 1$, $\{e^{-t\Delta}\}_{t\geq 0}$ is not a contraction semigroup on $\Omega^{1,p}$, $p \neq 2$. This and all the above results support the hypothesis that some kind of nonnegativity of the curvature should be connected with the contraction property. □

We turn now our attentions to sharper uniformly pointwise estimates of the heat kernel $H^q(x, y, t)$ on forms and its derivatives. If we consider the heat kernel $H^q(x, y, t)$ as Schwartz kernel of $e^{-t\Delta_q}$ then it immediately follows from the mapping properties of $e^{-t\Delta_q}$ and the local Sobolev embedding theorem that $H^q(x, y, t)$ is smooth in all variables for $t > 0$. The point are the estimates for the derivatives. We present here the approach and the results of [20].

Let us give precise defintions. A two-point form E^q with values $E^q(x, y, t) \in \Lambda^q T_x M \otimes \Lambda^q T_y M$ is called a good global heat kernel, if it satisfies the following conditions:

(H1) $E^q(x, y, t)$ is smooth for $t > 0$.

(H2) $\left(\frac{\partial}{\partial t} + \Delta\right) E^q(x, y, t) = 0$, where we apply Δ acting on E^q as a section depending on y.

(H3) $\lim\limits_{t\to 0^+} \int\limits_M E^q(x, y, t) \wedge *\omega_0(y) = \omega_0(x)$ for all $x \in M$ and

$$\omega_0 \in \Omega^q_c, \quad \text{i.e. } E^q(x, y, t) \longrightarrow \delta_{x,y}.$$

(H4) There exist constants $C_1, C_2 > 0$, depending on l, m, n, T, such that for all $x, y \in M, 0 < t < T$

$$\left|\left(\frac{\partial}{\partial t}\right)^l \nabla^m \nabla^n E^q(x, y, t)\right| \leq C_1 t^{-\frac{N}{2} - \frac{m+n}{2} - 1} \exp(-C_2 \frac{r^2(x, y)}{t}).$$

(H5) The heat kernels $E^q(x, y, t)$ and $E^{q+1}(x, y, t)$ are related by $\overline{d}_x(E^q(x, y, t) = \overline{\delta}_y(E^{q+1}(x, y, t)$.

Here $N = \dim M$.

Theorem 6.55 *Let (M^N, g) be open, complete, satisfying (I) and (B_k), $k > \frac{N}{2}$. Then there exists a good global heat kernel $E^q(x, y, t)$ satisfying the conditions* (H1) – (H5) *for $m, n \leq k + 2$, and*

(H6) $E^q(x, y, t) = E^q(y, x, t)$ *for all $x, y \in M$ (symmetry),*

(H7) $E^q(x, y, t + s) = \int\limits_M E^q(x, z, t) \wedge *E^q(z, y, s)$ *(semigroup property).*

Moreover, E^q is uniquely determined.

We refer to [20], [43] for the rather long and complicated proof. □

As it is well known, the existence of a good heat kernel has many good consequences in global analysis. We do not intend to present all this here, but restrict ourselves to a special case of applications. For many purposes one is interested to invert the Laplace operator Δ outside the space of L_2-harmonic forms.

Let $H = P_{\mathcal{H}}$ denote the projection onto

$$\mathcal{H}^{q,2} = \{\omega \in \Omega^{q,2} | \overline{d}\omega = \overline{\delta}\omega = 0\} = \ker \overline{\Delta}$$

(since (M^n, g) is complete).

Then one is searching for an operator G satisfying

$$\Delta G\omega = \omega - H\omega$$

and, if possible, for a meaningful integral representation of G. This G is called Green's operator, its kernel is called Green's kernel.

Theorem 6.56 *Let (M^n, g) be open, complete, and of bounded geometry. Assume further that* $\inf \sigma_e(\Delta_q) = \lambda_e > 0$, *and let be λ_0 the smallest spectral value > 0. Then*

$$G\omega(x) = \int_0^t \int_M E^q(x, y, t) \wedge *(\omega - H\omega)(y) \, \mathrm{dvol}_y$$

is Green's operator and has the following properties:

a) $|G\omega|_2 \leq (2\lambda_0)^{-\frac{1}{2}} |\omega|_2$ *for* $\omega \in \Omega^q_c$. *Hence G can be extended to a bounded linear operator* $G : \Omega^{q,2} \longrightarrow \Omega^{q,2} \equiv L_2(\Lambda^q)$.

b) $G\omega \in \Omega^{q,2,r}$ *for arbitrary large r.*

c) $\omega = H\omega + \overline{d\delta}G\omega + \overline{\delta d}G\omega$ *is the Hodge decomposition.*

Proof A complete proof is given in [19] under the assumption of the existence of a good heat kernel. This existence we have just now established. □

This finishes our short review of the heat equation and the heat kernel. We come back to other parabolic equations and maximum principles in section 11 and 15.

7 The wave equation, its Hamiltonian approach and completeness

Let $(E, h, \nabla, \cdot) \longrightarrow (M^n, g)$ be a Clifford bundle, (M^n, g) complete, $D = D(E, h, \nabla, \cdot)$ the associated essentially self-adjoint generalized Dirac operator. We denote also by D the self-adjoint closure.

Then, for $\varphi \in \Omega^{0,2,1}(E) \otimes \Omega^{0,2,1}(\mathbb{R})$,

$$\frac{\partial \varphi}{\partial t} - iD\varphi = 0 \tag{7.1}$$

$$\varphi(\cdot, 0) = \varphi_0 \tag{7.2}$$

is the initial value problem for the wave equation.

If $\varphi \in \Omega^{0,2,2}(E) \otimes \Omega^{0,2,1}(\mathbb{R})$ satisfies (7.1), then it satisfies the equation

$$\frac{\partial^2 \varphi}{\partial t^2} + D^2 \varphi = 0, \tag{7.3}$$

as follows immediately from $\frac{\partial}{\partial t}$(7.1) and inserting (7.1). Conversely, any nontrivial solution φ of (7.3) satisfying $\left(\frac{\partial \varphi}{\partial t} - iD\varphi\right)\Big|_{t=0} = 0$ produces a solution of (7.1). $\{e^{itD}\}_t$ is called the wave group, and there holds

$$|e^{itD}\varphi_0|_{L_2} = |\varphi_0|_{L_2}. \tag{7.4}$$

We see immediately from the spectral theorem that

$$\varphi(\cdot, t) = e^{itD}\varphi(\cdot, 0) \tag{7.5}$$

solves the equations (7.1), (7.2), hence we obtain from (7.4), (7.5) existence and uniqueness for (7.1), (7.2). As well known, D has finite propagation speed which is expressed by

Proposition 7.1 *If* $\varphi \in C_c^\infty(E)$ *then*

$$\text{supp } (e^{itD}\varphi) \subset U_{|t|}(\text{supp } \varphi). \tag{7.6}$$

□

We refer to [99] for the proof.
e^{-tD^2} and $e^{i\tau D}$ are related by the identity

$$e^{-tD^2} = \frac{1}{\sqrt{4\pi t}} \int e^{-\tau^2/4t} e^{i\tau D} d\tau$$

which comes from the spectral representation and

$$\int e^{-\tau^2/4t} e^{i\tau\lambda} d\tau = \sqrt{4\pi t} e^{-t\lambda^2}.$$

As we already mentioned several times, the graded Laplace operator $(\Delta_0, \ldots, \Delta_q, \ldots, \Delta_n)$ is a special case of a generalized Dirac operator D^2, $D = D(\Lambda^* T^* M, g_{\Lambda^*}, \nabla^{\Lambda^*}, \cdot)$. Hence we get wave equations

$$\frac{\partial^2}{\partial t^2}\varphi + \Delta_q \varphi = 0.$$

Again the most important case is the case $q = 0$, $\varphi = u$,

$$\frac{\partial^2}{\partial t^2}u + \Delta u = 0 \tag{7.7}$$

with initial conditions

$$u|_{t=0} = u_0, \tag{7.8}$$

$$\frac{\partial}{\partial t}u|_{t=0} = u_1. \tag{7.9}$$

Denote by $u[u_0, u_1]$ the (unique) solution of (7.7) – (7.8) if it exists. Then we immediately obtain

$$u[u_0, u_1] = u[0, u_1] + \frac{\partial}{\partial t}u[0, u_0].$$

Thus, the general Cauchy problem is reduced to the special one

$$\left.\begin{array}{l} u|_{t=0} = 0 \\ \frac{\partial u}{\partial t}\Big|_{t=0} = u_1 \end{array}\right\} \tag{7.10}$$

The existence and uniqueness are classical theorems. We will formulate it in the language of vector bundles $(E, D = \nabla^E) \longrightarrow (M^n, g)$ over Lorentz manifolds (M^n, g) with signature $(+, -, \cdots, -)$. Local coordinates are denoted by $x^a = x^1, \ldots, x^n$. The time coordinate t in $(\mathbb{R} \times M, dt^2 - g_M)$ is now replaced by a time function $t = t(x)$ such that

$$g^{ab}(\nabla_a t)(\nabla_b t) > 0.$$

Here $\nabla_a = \nabla^g_{\frac{\partial}{\partial x^a}}$ are the Levi-Civita derivatives.

Introduce the world function $\sigma = \sigma(x, y)$ of (M, g) as the solution of the differential equation

$$g^{ab}(\nabla_a \sigma)(\nabla_b \sigma) = 2\sigma$$

together with the initial conditions

$$(\nabla_a \sigma)(x, x) = 0$$

and

$$(\nabla_a \nabla_b \sigma)(x, x) = g_{ab}(x).$$

The solid conoid $D(x)$ is defined by $\sigma(x, y) \geq 0$ and the conoid surface $C(x)$ is defined by $\sigma(x, y) = 0$.

The zero time hypersurface is defined by the spacelike hypersurface

$$H = \{x \in M | t(x) = 0\}.$$

We set $B(x) = D(x) \cap H$ and $S(x) = C(x) \cap H$.

Finally the Cauchy problem for the wave equation above is replaced by

$$\begin{aligned} &L[u] \equiv g^{ab} D_a D_b u + \mathcal{W} u = f \\ &u|_H = u_0 \\ &g^{ab}(\nabla_a t)(D_b u)|_H = u_1. \end{aligned}$$

We call a differential operator L Laplace-like if it has the same principal symbol as

$$\Delta = g^{ab} \nabla_a \nabla_b.$$

Now we can give a mathematical formulation of **Huygens' principle**. A Laplace-like hyperbolic operator L is Huygens' if for every H, x, f, u_0, u_1 the solution $u = u(x)$ of Cauchy's problem taken at x depends only on the data u_0, u_1 and its derivatives taken on $S(x)$, but not on the values of u_0, u_1 in the interior of $B(x)$. That means, if the data differ only in $B(x) \setminus S(x)$, then the solution $u(x)$ at x will be the same. This can be made more precise, using the linearity of L. Huygens' principle becomes the following peculiar property of a differential operator L: If the initial data u_0, u_1 have support in the interior of $B(x)$, i.e.

$$\text{supp } u_0, \text{supp } u_1 \subset B(x) \setminus S(x),$$

then the solution u of the Cauchy problem (with $f = 0$)

$$\left.\begin{aligned} L[u] &= 0 \\ u|_H &= u_0 \\ g^{ab}(\nabla_a t)(D_b u)|_H &= u_1 \end{aligned}\right\}$$

vanishes at the point x.

There are several theories which give existence, uniqueness, and a construction of a solution of the above Cauchy problem. Each of them leads to its own criterion for Huygens' principle. Following the survey article [12], let us mention the following:

• J. Hadamard [69] constructed a solution by means of "finite parts of divergent integrals". He found that Huygens' principle holds for the differential operator L if and only if n is even, $n > 4$, and the formal adjoint L^* to L admits a logarithm-free elementary solution.

• F. G. Friedländer [59] and P. Günther [71] reformulated the Cauchy problem in terms of distributions and constructed distributional solutions. Huygens' principle holds, for even $n \geq 4$, if and only if the fundamental solution of L^* has its support on the characteristic conoid surface (and not in the interior of the solid conoid). This is equivalent to the condition that the "tail term" to L^* vanishes.

It is not surprising that all the necessary and sufficient criteria for Huygens' principle from different authors turn out to be equivalent. They can all be reduced to one more explicit condition, which is accessible to evaluations and calculations. We have, in order to present this condition, to introduce the so-called Hadamard coefficients $H_k = H_k(x, y)$ $(k = 0, 1, 2, \ldots$ to L. These two-point quantities are recursively defined by

$$\begin{aligned} g^{ab}(\nabla_a \sigma) D_b H_0 + \mu H_0 &= 0, \quad H_0(x, x) = I \\ g^{ab}(\nabla_a \sigma) D_b H_k + (\mu + k) H_k &= L[H_{k-1}]0, \quad \text{for } k \geq 1 \end{aligned}$$

where

$$\mu = \frac{1}{2}(\Delta\sigma - n)$$

and where all differentiations refer to the first argument x. Each $H_k = H_k(x, y)$ behaves like a section of E with respect to $x \in M$ and like a section of the dual bundle E^* with respect to $y \in M$. Thus, a diagonal value $H_k(x, x)$ can be interpreted as a section of $\operatorname{End} E$. In particular, $H_0(x, x) = I$ is required to be a the unit matrix. The differential-recursion system for the Hadamard coefficients has a remarkable property: it can be shown that there is a neighbourhood of the diagonal $M \times M$ where solutions $H_k = H_k(x, y)$ $(k = 0, 1, 2, \ldots)$ exist and are unique.

Theorem 7.2 *Huygens' principle never holds for $n = 2$ or for odd $n \geq 3$. So, let $n = 2m + 2 \geq 4$ be even. Huygens' principle holds for the formal adjoint L^* of the Laplace-like hyperbolic operator L if and only if the m-th Hadamard coefficient to L contains the world function $\sigma = \sigma(x, y)$ as a factor, that means there is a regular two-point function $R = R(x, y)$ such that*

$$H_m(x, y) = \sigma(x, y) R(x, y). \tag{7.11}$$

An explicit or implicit proof of this criterion can be found e.g. in the paper [72]. Notice that L and its formal adjoint L^* can easily interchange their roles since $L^{**} = L$. The two-point condition (7.11) implies a sequence of one-point conditions, namely condition for the Taylor coefficients of H_m with respect to the running point x and the origin y. We need, in order to present these, elements of a calculus of symmetric differential forms.

A symmetric p-form

$$u = u_p = u_{a_1 a_2 \dots a_p} dx^{a_1} dx^{a_2} \dots dx^{a_p}$$

is a special notation for a totally symmetric covariant tensor field of valence p. For instance, the Riemannian metric $g = g_{ab} dx^a dx^b$ is a symmetric 2-form. The multiplication of symmetric forms is the tensor multiplication followed by symmetrization. A metric g defines a trace operator tr by

$$\left.\begin{array}{l} \operatorname{tr} u_0 = 0 \\ \operatorname{tr} u_1 = 0 \\ \operatorname{tr} u_2 = g^{ab} u_{ab} \\ \operatorname{tr} u_p = g^{ab} u_{aba_3 \dots a_p} dx^{a_3} \dots dx^{a_p} \text{ for } p \geq 3. \end{array}\right\}$$

Every p-form u_p admits a unique decomposition into a part proportional to g and a trace-free part TSu_p:

$$u_p = g \cdot u_{p-2} + TSu_p, \quad \operatorname{tr}(TSu_p) = 0,$$

where "·" indicates symmetric multiplication. All these facts are naturally generalized to $(\operatorname{End} E)$-valued symmetric forms.

Theorem 7.3 *Let $n = 2m + 2 \geq 4$ be even. If Huygens' principle holds for L^*, then*

$$TS(D_{a_1} D_{a_2} \dots D_{a_p} H_m)(x, x) dx^{a_1} dx^{a_2} \dots dx^{a_p} = 0 \tag{7.12}$$

for $p = 0, 1, 2, \dots$. If, in particular, the objects $\mathcal{M}$, g and L are analytic, then the conditions (7.12) are not only necessary but also sufficient for Huygens' principle.

The proof of theorem 7.3 evaluates (7.11) and is based on

$$\begin{array}{rcl} (\nabla_a \nabla_b \sigma)(x, x) & = & g_{ab}(x) \\ (\nabla_{a_1} \nabla_{a_2} \cdots \nabla_{a_p} \sigma)(x, x) dx^{a_1} dx^{a_2} \cdots dx^{a_p} = 0 \text{ for } p \geq 3. & & \end{array} \tag{7.13}$$

Let us explain the left-hand side of (7.12): Covariant derivatives $D_{a_1}, D_{a_2}, \dots, D_{a_p}$ with respect to x are applied to the m-th Hadamard coefficient $H_m(x, y)$, the result is restricted to the diagonal $x = y$, finally the symmetric and trace-free part of the resulting p-tensor is taken.

Standard arguments of the theory of invariants show that the $(\operatorname{End} E)$-valued tensor components

$$(D_{a_1} D_{a_2} \dots D_{a_p} H_m)(x, x)$$

are polynomials in the variables

$$g_{ab}, g^{ab}, R_{abcd}, \nabla_{c_1} R_{abcd}, \nabla_{c_1} \nabla_{c_2} R_{abcd}, \dots$$

$$F_{ab}, D_{c_1}F_{ab}, D_{c_1}D_{c_2}F_{ab}, \dots \tag{7.14}$$

where the components of the Riemannian curvature R_{abcd} and the components of the gauge field curvature F_{ab} are defined through the Ricci identities

$$\begin{aligned}(\nabla_a\nabla_b - \nabla_b\nabla_a)v_c &= R_{abcd}v^d \\ (D_aD_b - D_bD_a)u &= F_{ab}u\end{aligned}$$

for a 1-form $v = v_a dx^a$, $v^a := g^{ab}v_b$ and for a section u of E, respectively.

For low values of n and p the condition (7.12) have been made explicit.

Theorem 7.4 *Let* $n = 4$. *If a differential operator*

$$L = g^{ab}D_aD_b + W$$

satisfies Huygens' principle, then

$$\begin{aligned}&\text{(i)} && W - \frac{R}{6}I = 0 \\ &\text{(ii)} && D^bF_{ab} = 0 \\ &\text{(iii)} && -\frac{2}{5}B_{ab}I = F_{abc}F_b^c + F_{bc}F_a^c - \frac{1}{2}g_{ab}F_{cd}F^{cd}.\end{aligned} \tag{7.15}$$

Here

$$W_{abcd}$$

are the components of the conformal curvature tensor (the definition of which we omit here) and

$$B_{ab} = \nabla^c\nabla^d W_{abcd} - \frac{1}{2}R^{cd}W_{acbd}$$

are the components of the Bach tensor.

A proof of theorem 7.4 was given for the scalar case in [66] and for the vector-bundle case in [104]. □

Theorem 7.5 *Let* $n = 6$. *If a differential operator*

$$L = g^{ab}D_aD_b + W$$

satisfies Huygens' principle, then

$$30g^{ab}D_aD_bC + 6RC + W_{abcd}W^{abcd}I + 15F_{ab}F^{ab} + 90C^2 = 0,$$

where $C = W - \frac{R}{5}I$ *is the so-called Cotton endomorphism to* L.

A proof of theorem 7.5 was given for the scalar case in [72] and for the general case in [104]. The result also appeared earlier in the context of spectral geometry [64], [65]. □

Note that the formula in theorem 7.4 also admits an interpretation, namely as a nonlinear Higgs equation for $C = C(x)$ with some source terms.

We finish at this point our short review of Huygens' principle and refer for further information to [71] and the excellent survey [12]. The latter was our guide through this topic.

As we already pointed out in the introduction, one effective method to treat PDEs as manifolds is to transform them into ODEs on infinite-dimensional manifolds. Our first example for this will be the wave equation and the Klein-Gordon equation considered as infinite-dimensional Hamiltonian system.

Let (M^n, g) be complete and consider the manifold $(P, \omega) = (\Omega^{0,2,1}(M, g) \times \Omega^{0,2,0}(M, g)) \equiv \Omega^{0,2,1}(M, g) \times L_2(M, g), \omega)$, where $\omega((u_1, v_1), (u_2, v_2)) = \langle v_2, u_1 \rangle - \langle v_1, u_2 \rangle$. Then (P, ω) is an infinite-dimensional symplectic manifold. Consider the Hamiltonian

$$H(u, v) = \frac{1}{2}\left[\int (v^2 + |\nabla u|^2 - m^2 u^2)\, \mathrm{dvol}_x(g)\right]$$

and the vector field X_H with $\mathcal{D}_{X_H} = \Omega^{0,2,2} \times \Omega^{0,2,1}$,

$$X_H(u, v) = (v, \Delta u - m^2 u).$$

Then X_H is densely defined, $\mathcal{D}_{X_H}$ is a submanifold, and we can apply Stokes' theorem since (M^n, g) is complete.

Lemma 7.6 *X_H is Hamiltonian with respect to H.*

Proof We have to show $i_{X_H}\omega = dH$, i.e.

$$\omega((v, \Delta u + m^2 u), (w_1, w_2)) = dH(w_1, w_2) \tag{7.16}$$

for $(w_1, w_2) \in \Omega$. But

$$\begin{aligned} \omega((v, \Delta u - m^2 u), (w_1, w_2)) &= \int (w_2 v + (\Delta u - m^2 u)\cdot w_1)\, \mathrm{dvol}_x(g) \\ &= \int (w_2 v + (\nabla u, \nabla w_1) - m^2 u w_1)\, \mathrm{dvol}_x(g) \\ dH(u, v)(w_1, w_2) &= D_2 H(u, v)(w_2) + D_1 H(u, v)(w_1) \\ &= \int (v w_2 + (\nabla u, \nabla w_1) - m^2 u w_2)\, \mathrm{dvol}_x(g), \end{aligned}$$

hence (7.16) is satisfied. □

We obtain $\frac{\delta H}{\delta v} = v$, $\frac{\delta H}{\delta u} = \Delta u - m^2 u$: The equations of motion are

$$\frac{\partial u}{\partial t} = \frac{\delta H}{\delta v} = v,\ \frac{\partial v}{\partial t} = -\frac{\delta H}{\delta u} = -\Delta u + m^2 u.$$

In second order form, we obtain

$$\frac{\partial^2 u}{\partial t^2} u = -\Delta u + m^2 u.$$

We established

Theorem 7.7 *The Klein-Gordon equation*

$$\frac{\partial^2}{\partial t^2}u + \Delta u = m^2 u \tag{7.17}$$

can be described as an Hamiltonian system which has a flow.

Proof The existence of a flow follows from the hyperbolic version of the Hille-Yosida theorem. □

The Hamiltonian representation immediately generalizes to the case

$$H(u,v) = \int \left(\frac{1}{2}v^2 + \frac{1}{2}|\nabla u|^2 - F(u)\right) \mathrm{dvol}_x(g),$$

$X_H(u,v) = \left(\frac{\delta H}{\delta v}, -\frac{\delta H}{\delta u}\right) = (v, \Delta u - F'(u))$, which yields

$$\frac{\partial^2 u}{\partial t^2} = -\Delta u + F'(u). \tag{7.18}$$

If we choose $F(u) = \frac{1}{2}m^2u^2 + G(u)$, then we obtain back (7.17) in the case $G(u) = 0$. Other choices of $G(u)$ appear in the quantum theory of self-interacting reasons. An important special case is $G(u) = \frac{\alpha}{p+1}u^p$.

The advantage of the description of a PDE as Hamiltonian system consists in the fact that

1) one can often relatively easy get existence und uniqueness and

2) one often gets relatively easy integrals of motion.

We finish at this point our brief discussion of the wave equation. Concerning estimates of the wave kernel and calculation of wave invariants, we refer e.g. to [25], [122].

8 Index theory on open manifolds

As well known, index theory dominated a big part of global analysis from 1962 up to now. Its facinating and striking achievements in geometry, topology and mathematical physics support the meaning of global analysis.

But almost all of its achievements concern underlying closed manifolds. This will be the topic of David Bleeker. We devote this section to the index theory on open manifolds, presenting here a brief survey and start with an outline of the initial situation.

If $X \supset \mathcal{D}_D \xrightarrow{D} Y$ is a Fredholm operator, i.e. $\dim\ker D < \infty$, $\mathrm{Im}\, D$ closed and $\dim\mathrm{coker}\, D = \dim Y/\mathrm{Im}\, D < \infty$, then the analytical index $\mathrm{ind}_a D = \dim\ker D - \dim\mathrm{coker}\, D$ is well defined. From $\mathrm{ind}_a D > 0$ one concludes immediately that the solutions of $Du = \varphi$ are not unique, from $\mathrm{ind}_a D < 0$ one concludes that for certain φ the equation $Du = \varphi$ is not solvable. Moreover, $\mathrm{ind}_a D$ provides in many applications a very important information about the analysis and geometry in question. It is very well known, that an elliptic operator $D : C^\infty(E) \longrightarrow C^\infty(E)$ is (after completion and extension) Fredholm if the underlying manifold is closed. This follows from the existence of a parametrix P,

$$PD - \mathrm{id} = K_1, \quad DP - \mathrm{id} = K_2, \tag{Par}$$

and the fact that integral operators K with C^∞-kernels $\mathcal{K}$ over closed manifolds are compact. It is very well known too, that on open manifolds these statements are wrong in general. (Par) is still solvable but the K_i are not compact in general. Elliptic operators on open manifolds are not Fredholm – or one adds additional restricting conditions. Hence a general index theory valid for all open elliptic situations cannot be established.

There are 3 ways out from this difficult situation.

1) One could ask for special conditions in the open case under which an elliptic D is still Fredholm, then try to establish an index formula, and finally present applications. These conditions could be conditions on D, on M and E or a combination of both. In [4] the author formulates an abstract (and very natural) condition for the Fredholmness of D and assumes nothing for the geometry. But in all substantial applications this condition can be assured by conditions for the geometry. The other extreme case is that discussed in [89], [90], where the authors consider the L_2-index theorem for locally symmetric spaces. Under relatively restricting conditions concerning the geometry and topology at infinity, Fredholmness and an index theorem is proved in [21].

2) One could generalize the notion of being Fredholm (using other operator algebras) and then establish a meaningful index theory with applications. The discussion of these two approaches will be contained in this paragraph.

3) Another approach will be relative index theory which is less restrictive concerning the geometrical situation (compared with the absolute case) but its outcome are only statements on the relative index, i.e. how much the analytical properties of D differ from those of an appropriate perturbation D'. This approach will be discussed briefly.

4) For open coverings $(\tilde{M}, \tilde{g})$ of closed manifolds (M^n, g) and lifted data there is an approach which goes back to Atiyah, (cf. [7]). This has been further elaborated by Cheeger, Gromov and others. The main point is that all considered (Hilbert-) modules are modules over a von Neumann algebra and one replaces the usual trace by a von Neumann trace. We will not dwell on this approach since there is a well established highly elaborated theory. Moreover special features of openess are not discussed. The openess is reflected by the fact that all modules under consideration are modules over the von Neumann algebra $\mathcal{N}(\pi)$, $\pi = \text{Deck}(\tilde{M} \longrightarrow M)$.

We start with the first approach, and with the question which elliptic operators over open manifolds are Fredholm in the classical sense above. Let (M^n, g) be open, oriented, complete, $(E, h) \longrightarrow (M^n, g)$ be a Hermitean vector bundle with involution $\tau \in \text{End}\,(E)$, $E = E^+ \oplus E^-$, $D : C^\infty(E) \longrightarrow C^\infty(E)$ an essentially self-adjoint first order elliptic operator satisfying $D\tau + \tau D = 0$. We denote $D^\pm = D|_{C^\infty(E^\pm)}$. Then we can write as usual

$$D = \begin{pmatrix} 0 & D^- \\ D^+ & 0 \end{pmatrix} : \begin{matrix} C^\infty(E^+) \\ \oplus \\ C^\infty(E^-) \end{matrix} \longrightarrow \begin{matrix} C^\infty(E^+) \\ \oplus \\ C^\infty(E^-) \end{matrix} . \tag{8.1}$$

The (L_2-)index $\text{ind}_a D$ is defined as

$$\text{ind}_a D := \text{ind}_a D^+ := \dim \ker\, D^+ - \dim \text{coker}\, D^+ = \dim \ker\, D^+ - \dim \ker\, D^- \tag{8.2}$$

if these numbers would be defined.

We refer to section 2 for the definition of $\Omega^{2,i}(E, D)$. N. Anghel proved in [4] necessary and sufficient criteria for the Fredholmness of $D : \Omega^{2,1}(E, D) \longrightarrow \Omega^{2,0}(E, D) = L_2(E)$.

Proposition 8.1 *The following statements are equivalent*

a) *D is Fredholm.*

b) $\dim \ker D < \infty$ *and there is a constant $c > 0$ such that*

$$|D\varphi|_{L_2} \geq c \cdot |\varphi|_{L_2}, \quad \varphi \in (\ker D)^{\perp} \cap \Omega^{2,1}(E, D), \tag{8.3}$$

where $(\ker D)^{\perp} \equiv \mathcal{H}^{\perp}$ is the orthogonal complement of $\mathcal{H} = \ker D$ in $L_2(E)$.

c) *There exists a bounded non-negative operator $P : \Omega^{2,2}(E, D) \longrightarrow L_2(E)$ and bundle morphism $R \in C^{\infty}(\mathrm{End}\, E)$, R positive at infinity, i.e. there exist a compact $K \subset M$ and a $k > 0$ s. t. pointwise on $E|_{M \setminus K}$, $R \geq k$, and such that on $\Omega^{2,2}(E, D)$*

$$D^2 = P + R. \tag{8.4}$$

d) *There exist a constant $c > 0$ and compact $K \subset M$ such that*

$$|D\varphi|_{L_2} \geq c \cdot |\varphi|, \quad \varphi \in \Omega^{2,1}(E, D), \quad \mathrm{supp}\ (\varphi) \cap K = \emptyset. \tag{8.5}$$

Example 8.2 A class of examples for proposition 8.1 is given by the generalized Dirac operator D of a Clifford bundle $(E, h, \nabla_i) \longrightarrow (M^n, g)$ over a complete Riemannian manifold such that at infinity $\mathcal{R} \geq c > 0$ in I. In particular the Dirac operator of a spin structure of a complete Riemannian spin manifold (M^n, g) with scalar curvature $\geq c > 0$ at infinity provides a class of examples. □

We see that, if any of the conditions of proposition 8.1 is satisfied, then $\mathrm{ind}_a D \equiv \mathrm{ind}_a D^+$ is well defined. The main task now is to establish a meaningful index theorem. This has been performed in [4].

Theorem 8.3 *Let (M^n, g) be open, complete, oriented, $(E, h, \tau) = (E^+ \oplus E^-, h) \longrightarrow (M^n, g)$ a $\mathbb{Z}_2$-graded Hermitean vector bundle and $D : C_c^{\infty}(E) \longrightarrow C_c^{\infty}(E)$ first order elliptic, essentially self-adjoint, compatible with the $\mathbb{Z}_2$-grading (i.e. supersymmetric), $D\tau + \tau D = 0$. Let $K \subset M$ be a compact subset such that 8.1 a) for K is satisfied, and let $f \in C^{\infty}(M, \mathbb{R})$ be such that $f = 0$ on $U(K)$ and $f = 1$ outside a compact subset. Then there exist a volume density ω and a contribution I_ω such that*

$$\mathrm{ind}_a \overline{D}^+ = \int_M (\omega(1 - f(x))\ \mathrm{dvol}_x(g) + I_\omega, \tag{8.6}$$

where ω has an expression locally depending on D and I_ω depends on D and f restricted to $\Omega = M \setminus K$.

Until now, the differential form $\omega\, \mathrm{dvol}_x(g)$ is a complete mystery. One would like to express it by well known canonical terms coming e. g. from the Atiyah-Singer index form $\mathrm{ch}\ \sigma(D^+) \wedge \tilde{S}(M)$, where $\tilde{S}(M)$ denotes the Todd genus of M. In fact this can be done in two steps. The first step is a generalization of the relative index theorem of Gromov / Lawson (cf. [69]).

Theorem 8.4 *Let (M_j^n, g_j) be open, oriented, complete. $(E_j, \tau_j) \longrightarrow (M_j^n, g_j)$, $j = 1, 2$, two $\mathbb{Z}_2$-graded Hermitean vector bundles, $D_j : C_c^{\infty}(E_j) \longrightarrow C_c^{\infty}(E_j)$, $j = 1, 2$, two supersymmetric essentially self-adjoint first order elliptic differential operators, $D_j\tau_j + \tau_j D_j = 0$, $j = 1, 2$, and assume that the D_j agree outside compact sets K_1, K_2, i. e. there are compact sets $K_j \subset M_j$, an orientation preserving isometry $F : M_1 \setminus K_1 \longrightarrow M_2 \setminus K_2$*

which is covered by a bundle isometry $\tilde{F} : E_1|_{M_1 \setminus K_1} \longrightarrow E_2|_{M_2 \setminus K_2}$ *such that on* $M_1 \setminus K_1$, $\tilde{F}\tau_1 = \tau_2 \tilde{F}$ *and* $D_1 = \tilde{F}^{-1} \circ D_2 \circ \tilde{F}$. *Then*

$$\mathrm{ind}_a D_1^+ - \mathrm{ind}_a D_2^+ = \int\limits_{M_1} \mathrm{ch}\ \sigma(D_1^+) \wedge \tilde{S}(M_1) - \int\limits_{M_2} \mathrm{ch}\ \sigma(D_2^+) \wedge \tilde{S}(M_2), \tag{8.7}$$

where the r. h. s. will be explained below.

Proof Choose K_j, $j = 1, 2$, large enough such that $M_1 \setminus K_1$ and $M_2 \setminus K_2$ are isometric as stated above and moreover (8.5) holds. We identify $M_1 \setminus K_1$ and $M_2 \setminus K_2$ and write $M_1 \setminus K_1 = \Omega = M_2 \setminus K_2$. We obtain from theorem 8.3

$$\mathrm{ind}_a D_j^+ = \int\limits_{M_j} w_j(x)(1 - f_j(x))\,\mathrm{dvol}_x(g) + I_\Omega^j.$$

$D_1|_\Omega = D_2|_\Omega$, $f_1|_\Omega = f_2|_\Omega$ imply $\omega_1|_\Omega = \omega_2|_\Omega$ and $I_\Omega^1 = I_\Omega^2$. Hence

$$\begin{aligned}\mathrm{ind}_a D_1^+ - \mathrm{ind}_a D_2^+ &= \int\limits_{M_1} \omega_1(1 - f_1)\,\mathrm{dvol}_x(g) - \int\limits_{M_2} \omega_2(1 - f_2)\,\mathrm{dvol}_x(g) \\ &= \int\limits_{K_1} \omega_1\,\mathrm{dvol}_x(g_1) - \int\limits_{K_2} \omega_2\,\mathrm{dvol}_x(g_2).\end{aligned}$$

Repeat the glueing procedure from Gromov / Lawson: Let $N \subset \Omega$ be a closed hypersurface which decomposes Ω into a bounded and unbounded part, form a Riemannian, bundle and D-collar along N which finally yields $\tilde{M} := M_1' \cup_N (-M_2')$ and $\tilde{D}$ on $\tilde{M}$. According to the Atiyah-Singer index theorem,

$$\begin{aligned}\mathrm{ind}_a \tilde{D}^+ &= \int\limits_{M_1'} \mathrm{ch}\ \sigma(D_1^+) \wedge \tau(M) - \int\limits_{M_2'} \mathrm{ch}\ \sigma(D_2^+) \wedge \tau(M) = \\ &= \int\limits_{K_1} \mathrm{ch}\ \sigma(D_1^+) \wedge \tau(M) - \int\limits_{K_2} \mathrm{ch}\ \sigma(D_2^+) \wedge \tau(M).\end{aligned}$$

By means of local parametrices as above we get

$$\mathrm{ind}_a \tilde{D}^+ = \int\limits_{M_1'} \omega_1\,\mathrm{dvol}_x(g_1) - \int\limits_{M_2'} \omega_2\,\mathrm{dvol}_x(g_2) = \int\limits_{K_1} \omega_1\,\mathrm{dvol}_x(g_1) - \int\limits_{K_2} \omega_2\,\mathrm{dvol}_x(g_2),$$

i. e.

$$\int\limits_{K_1} \omega_1\,\mathrm{dvol}_x(g_1) - \int\limits_{K_2} \omega_2\,\mathrm{dvol}_x(g_2) = \int\limits_{K_1} \mathrm{ch}\ \sigma(D_1^+) \wedge \tau(M) - \int\limits_{K_2} \mathrm{ch}\ \sigma(D_2^+) \wedge \tau(M).$$

If we now define

$$\int\limits_{M_1} \mathrm{ch}\ \sigma(D_1^+) \wedge \tau(M) - \int\limits_{M_2} \mathrm{ch}\ \sigma(D_2^+) \wedge \tau(M) := \int\limits_{K_1} \mathrm{ch}\ \sigma(D_1^+) \wedge \tau(M) - \int\limits_{K_2} \mathrm{ch}\ \sigma(D_2^+) \wedge \tau(M)$$

(8.8)

then (8.7) is established. (8.8) is well motivated since $D_1|_\Omega = D_2|_\Omega$. □

As a corollary from theorem 8.4 we obtain the following

Index Theorem 8.5 *Let (M^n, g) be open, oriented, complete, $(E, h, \tau) \longrightarrow (M^n, g)$ a $\mathbb{Z}_2$-graded Hermitean vector bundle, $D : C_c^\infty(E) \longrightarrow C_c^\infty(E)$ a first order elliptic essentially self-adjoint supersymmetric differential operator, $D\tau + \tau D = 0$, which shall be assumed to be Fredholm. Let $K \subset M$ compact such that (8.5) is satisfied. Then*

$$\mathrm{ind}_a D^+ = \int_K \mathrm{ch}\ \sigma(D^+) \wedge \tau(M) + I_\Omega, \tag{8.9}$$

where $\mathrm{ch}\ \sigma(D^+) \wedge \tau(M)$ *is the Atiyah-Singer index form and I_Ω is a bounded contribution depending only on $D|_\Omega$, $\Omega = M \setminus K$.*

Proof Set $J_\Omega := \mathrm{ind}_a D^+ - \int_K \mathrm{ch}\ \sigma(D^+) \wedge \tau(M)$. Then we infer from (8.7), (8.8) that J_Ω is independent of the local expression of $D|_K$, i. e. depends only on $D|_\Omega$. □

Remarks 8.6 **a)** As we already mentioned, $\mathbb{Z}_2$-graded Clifford bundles, and associated generalized Dirac operators D such that $D^2 = \Delta^E + \mathcal{R}$, $\mathcal{R} \geq c \cdot \mathrm{id}$, $c > 0$, outside some compact $K \subset M$, yield examples for theorem 8.5. A special case is the Dirac operator over a Riemannian spin manifold with scalar curvature $\geq c > 0$ outside $K \subset M$.

b) Much more general perturbations than compact ones will be considered at the end of this section. □

The practical value of theorem 8.5 depends on the concrete situation. (8.9) contains still the number I_Ω which has no canonical expression.

But on the topology and geometry in the formulation of 8.5 nothing has been explicitely assumed. It is clear, the more we assume on the topology, geometry and the explicit expression of the operator the more concrete becomes a potential index theorem.

The starting point for such very good tractable cases are two papers of Atiyah, the paper on the η-invariant [6] and on the index theorem for covering manifolds. Let (M^n, g) be a compact (oriented) Riemannian manifold with boundary $\partial M = N$ and form the open manifold $X = M \cup_N N \times [0, \infty[$ with product metric outside a compact set. If $D : C^\infty(M, E) \longrightarrow C^\infty(M, F)$ is a first order elliptic operator on M which takes in a collar $[0, \varepsilon[\times N$ the special form $D = \sigma\left(\frac{\partial}{\partial u} + A\right)$, where $\frac{\partial}{\partial u}$ is the inward normal, $\sigma : E|_N \longrightarrow F|_N$ a bundle isomorphism and $A : C^\infty(N) \longrightarrow C^\infty(N)$ is selfadjoint, then E, F and D extend to X, $D : C_c^\infty(E) \longrightarrow C_c^\infty(F)$. It was proved in [6] that the (L_2-) closure $\overline{D}$ of D has well defined finite index and

$$\mathrm{ind}_a \overline{D} = \int_X \omega_D - \frac{\eta(0)}{2} - \frac{1}{2}(h_\infty(E) - h_\infty(F)), \tag{8.10}$$

where ω is the local index form, $\eta(0)$ the η-invariant of A, $h_\infty(E)$ the dimension of the subspace of $\ker A$ consisting of limiting values of extended L_2-sections φ of E satisfying $D\varphi = 0$ and $h_\infty(F)$ is analogously defined by $D^*\psi = 0$. $h_\infty(E) - h_\infty(F)$ can be rewritten as

$$\mathrm{tr}\, S(0) = h_\infty(E) - h_\infty(F),$$

where $S(\lambda)$ is the scattering matrix for the pair $\overline{D^*D|_{C_c^\infty(X,E)}}, \overline{-\frac{\partial^2}{\partial u^2} + A^2|_{C_c^\infty(N\times\mathbb{R}^+,E)}}$ which Dirichlet b. c..

If $\ker A = \{0\}$ then

$$\mathrm{ind}_a \overline{D} = \int \omega_D - \frac{\eta(0)}{2}. \tag{8.11}$$

(8.10), (8.11) are very explicit. For the signature operator e. g. very well known formulas come out.

The next generalization of product ends would be ends with warped product metrics, multiply warped product, asymptotically multiply warped product ends and finally a bounded variation of such ends, all this for Dirac type operators, suitably parametrized.

The other case of a very special class of open manifolds are coverings $(\tilde{M}, \tilde{g})$ of a closed manifold (M^n, g). Let $E, F \longrightarrow (M^n, g)$ be Hermitean vector bundles over the closed manifold (M^n, g). $D : C^\infty(E) \longrightarrow C^\infty(F)$ be an elliptic operator, $(\tilde{M}, \tilde{g}) \longrightarrow (M, g)$ a Riemannian covering, $\tilde{D} : C_c^\infty(\tilde{E}) \longrightarrow C_c^\infty(\tilde{F})$ the corresponding lifting and $\Gamma = \text{Deck } (\tilde{M}^n, \tilde{g}) \longrightarrow (M^n, g)$. The actions of Γ and $\tilde{D}$ commute. If $P : L_2(\tilde{M}, \tilde{E}) \longrightarrow \mathcal{H}$ is the orthogonal projection onto a closed subspace $\mathcal{H} \subset L_2(\tilde{M}, \tilde{E})$ then one defines the Γ-dimension $\dim_\Gamma \mathcal{H}$ of $\mathcal{H}$ as

$$\dim_\Gamma \mathcal{H} := \mathrm{tr}_\Gamma P,$$

where tr_Γ denotes the von Neumann trace and $\mathrm{tr}_\Gamma P$ can be any real number ≥ 0 or $= \infty$.

If one takes $\mathcal{H} = \mathcal{H}(\tilde{D}) = \ker \tilde{D} \subset L_2(\tilde{E})$, $\mathcal{H}^* = \mathcal{H}(\tilde{D}^*) = \ker(\tilde{D}^*) \subset L_2(\tilde{F})$ then one defines the Γ-index $\mathrm{ind}_\Gamma \tilde{D}$ as

$$\mathrm{ind}_\Gamma \tilde{D} := \dim_\Gamma \mathcal{H}(\tilde{D}) - \dim_\Gamma \mathcal{H}(\tilde{D}^*).$$

Atiyah proves in [7] the following main

Theorem 8.7 *Under the assumptions above there holds*

$$\mathrm{ind}_D D = \mathrm{ind}_\Gamma \tilde{D}.$$

□

It was this theorem which was the orign of the von Neumann analysis as a fastly growing area in geometry, topology and analysis. Moreover, the proof of theorem 8.3 is strongly modelled by that of 8.7.

Another very important special case which is related to the case above of coverings are locally symmetric spaces of finite volume. There is a vast number of profound contributions, e. g. [11], [28], [89], [90], [91]. For reasons of space, we cannot give here an overview.

There is another approach by Gilles Caron which we briefly present now (cf. [21]).

Let $(E, h, \nabla, \cdot) \longrightarrow (M^n, g)$ be a Clifford bundle over the complete Riemannian manifold (M^n, g) and $D : C^\infty(E) \longrightarrow C^\infty(E)$ the associated generalized Dirac operator. D is called non-parabolic at infinity if there exists a compact set $K \subset M$ such that for any open and relatively compact $U \subset M \setminus K$ there exists a constant $C(U) > 0$ such that

$$C(U)|\varphi|_{L_2(E|_U)} \leq |D\varphi|_{L_2(E|_{M\setminus K})} \text{ for all } \varphi \in C_c^\infty(E|_{M\setminus K}). \tag{8.12}$$

To exhibit the consequences of (8.12), we establish another characterization of it.

Proposition 8.8 *Let $(E, h, \nabla, \cdot) \longrightarrow (M^n, g)$ and D as above and let $W(E)$ be a Hilbert space of sections such that*

a) *$C_c^\infty(E)$ is dense in $W(E)$ and*

b) *the injection $C_c^\infty(E) \hookrightarrow \Omega_{\text{loc}}^{2,1}(E, D)$ extends continuously to $W(E) \longrightarrow \Omega_{\text{loc}}^{2,1}(E, D)$.*

Then $D : W(E) \longrightarrow L_2(E)$ is Fredholm if and only if there exist a compact $K \subset M$ and a constant $C(K) > 0$ such that

$$C(K) \cdot |\varphi|_W \leq |D\varphi|_{L_2(E|_{M\setminus K})} \text{ for all } \varphi \in C_c^\infty(E|_{M\setminus K}). \tag{8.13}$$

□

Remark 8.9 The norm $\mathcal{N}(\cdot)$ above is equivalent to the norm

$$\mathcal{N}_{\overline{U}(K)}(\cdot), \mathcal{N}_{\overline{U}(K)}(\varphi)^2 = |\varphi|^2_{L_2(E|_{\overline{U}(K)})} + |D\varphi|^2_{L_2(E)}. \tag{8.14}$$

Corollary 8.10 *$D : C^\infty(E) \longrightarrow C^\infty(E)$ is non-parabolic at infinity if and only if there exists a compact $K \subset M$ such that the completion of $C_c^\infty(E)$ w. r. t. $\mathcal{N}_K(\cdot)$,*

$$\mathcal{N}_K(\varphi)^2 = |\varphi|^2_{L_2(E|_K)} + |D\varphi|^2_{L_2} \tag{8.15}$$

yields a space $W(E)$ such that the injection $C_c^\infty(E) \longrightarrow \Omega_{\text{loc}}^{2,1}(E, D)$ continuously extends to $W(E)$.

The point now is that we know if D is non-parabolic at infinity then $D : W(E) \longrightarrow L_2(E)$ is Fredholm. We emphasize, this does not mean $L_2(E) \supset \mathcal{D}_D \stackrel{D}{\longrightarrow} L_2(E)$ is Fredholm. We get a weaker Fredholmness, not the desired one. But in certain cases this can be helpful too.

Suppose again a $\mathbb{Z}_2$-grading of E and D, $D = \begin{pmatrix} 0 & D^- \\ D^+ & 0 \end{pmatrix}$, $L_2(E) = L_2(E^+) \oplus L_2(E^-)$, $W(E) = W(E^+) \oplus W(E^-)$. Following Gilles Carron, we now define the extended index $\text{ind}_e D^+$ as

$$\begin{aligned} \text{ind}_e D^+ &:= \dim \ker {}_W D^+ - \dim \ker {}_{L_2} D^- \\ &= \dim\{\varphi \in W(E^+) \,|\, D^+\varphi = 0\} - \\ &\quad - \dim\{\varphi \in L_2(E^-) \,|\, D^-\varphi = 0\}. \end{aligned} \tag{8.16}$$

If we denote $h_\infty(D^+) := \dim(\ker_W D^+ / \ker_{L_2} D^+)$ then we can (8.16) rewrite as

$$\text{ind}_e D^+ = h_\infty(D^+) + \text{ind}_{L_2} D^+ = h_\infty(D^+) + \dim \ker {}_{L_2} D^+ - \dim \ker {}_{L_2} D^-. \tag{8.17}$$

The most interesting questions now are applications and examples. For $D =$ Gauß-Bonnet operator there are applications and examples in [22]. For the general case it is not definitely clear, is non-parabolicity really a practical sufficient criterion for Fredholmness since in concrete cases it will be very difficult it to establish. In some well known standard cases which have been presented by Carron and which we will discuss now it is of great use.

Proposition 8.11 *Let $D : C^\infty(E) \longrightarrow C^\infty(E)$ be a generalized Dirac operator and assume that outside a compact $K \subset M$ the smallest eigenvalue $\lambda_{\min}(x)$ of $\mathcal{R}_x$ in $D^2 = \nabla^*\nabla + \mathcal{R}$ is ≥ 0. Then D is non-parabolic at infinity.*

Proof Let K be as supposed and $\varphi \in C_c^\infty(E|_{M\setminus K})$. Then

$$|D\varphi|^2_{L_2(E|_{M\setminus K})} = \langle D^2\varphi, \varphi\rangle_{M\setminus K} \geq |\nabla\varphi|^2_{M\setminus K}. \tag{8.18}$$

Moreover $|\nabla\varphi|_{M\setminus K} \geq |d|\varphi||_{L_2(M\setminus K)}$ by Kato's inequality. According to [7], there exists for any bounded $U \subset M$ a constant $C(U)$ such that

$$\langle(\Delta_0 + \chi_K)u, u\rangle_M \geq C(U)|u|^2_{L_2(U)} \text{ for all } u \in C_c^\infty(M).$$

Hence for $u(x) = |\varphi(x)|$ (which is not necessarily C^∞ but the integral estimates remain true), $\varphi \in C_c^\infty(E|_{M\setminus K})$, $||\varphi|_{L_2(M)}| = |\varphi|_{L_2(E)}$, we obtain

$$|d|\varphi||^2_{L_2(M\setminus K)} = |d|\varphi||^2_{L_2(M)} + \langle\chi_k|\varphi|, |\varphi|\rangle \geq C(U)|\varphi|^2_{L_2(E|_U)}, \tag{8.19}$$

i. e. we infer from (8.18), (8.19) for bounded $U \supset M \setminus K$

$$C(U)|\varphi|_{L_2(E|_U)} \leq |D\varphi|_{L_2(E|_{M\setminus K})}, \quad \varphi \in C_c^\infty(E|_{M\setminus K}).$$

□

For the general understanding of this approach we still add a remark concerning the samewhat strange norms on $C_c^\infty(E)$ used here. These are

$$\begin{aligned} \mathcal{N}_\Phi(\varphi) &= |D\varphi|_{L_2}, \\ \mathcal{N}_K(\varphi)^2 &= |\varphi|^2_{L_2(E|_K)} + |D\varphi|^2_{L_2}, \\ \mathcal{N}_{\Phi,\nabla}(\varphi) &= |\nabla\varphi|_{L_2(T^*\otimes E)}, \\ \mathcal{N}_{K,\nabla}(\varphi)^2 &= |\varphi|^2_{L_2(E|_K)} + |\nabla\varphi|^2_{L_2}. \end{aligned}$$

All these are norms on C_c^∞. The only property which is not absolutely evident is $\mathcal{N}(\varphi) = 0 \Longrightarrow \varphi = 0$. Consider $\mathcal{N}_\Phi(\varphi) = |D\varphi|_{L_2}$. $|D\varphi|_{L_2} = 0$ implies $|D^2\varphi|_{L_2} = 0$, $D^2\varphi = 0$. Moreover, outside a compact $L \subset M$, $\varphi = 0$. Then the unique continuation theorem of Aronszajn says $\varphi \equiv 0$. $\mathcal{N}_{\Phi,\nabla}(\varphi) = 0$ implies φ parallel. $\varphi = 0$ outside L, i. e. $\varphi \equiv 0$. Quite analogous for the other two norms. It is very easy to see that for K compact with non-empty interior $W_\mathcal{N}(E)$ is independent of K. Moreover, for D non-parabolic at infinity, $\mathcal{N}_K \sim \mathcal{N}_\Phi$. Always $\mathcal{N}_\Phi \leq \mathcal{N}_K$ is clear. Conversely, (8.15) implies (after decomposition of $\varphi \in C_c^\infty(E)$) $\mathcal{N}_K \leq C \cdot \mathcal{N}_\Phi$. Hence in this case we have $W_{\mathcal{N}_K}(E) = W_{\mathcal{N}_\Phi}(E)$ and write $W_{\mathcal{N}_\Phi}(E) = W_0(E)$. From the definition of the $\mathcal{N}$s it is clear that the elements of the completion $W_\mathcal{N}(E) = \overline{C_c^\infty(E)}^{\mathcal{N}}$ must not be $\in L_2$. This just generates the additional term h_∞ in the extended index and is simultaneously the heart of the whole approach. One admits a "non-L_2 perturbation" of L_2-Fredholmness and tries nevertheless to draw fruitful conclusions.

We obtain from proposition 8.11

Corollary 8.12 *Assume the hypothesis of 8.11. Then $D : W_0(E) \longrightarrow L_2(E)$ is Fredholm.*

□

Under certain additional assumptions the pointwise condition on $\lambda_{\min}(x)$ of $\mathcal{R}_x$ can be replaced by a (weaker) integral condition. Denote $\mathcal{R}_-(x) = \max\{0, -\lambda_{\min}(x)\}$, where $\lambda_{\min}(x)$ is the smallest eigenvalue of $\mathcal{R}_x$.

Theorem 8.13 *Suppose that for a $p > 2$ (M^n, g) satisfies the Sobolev inequality*

$$c_p(M)\left(\int\limits_M |u|^{\frac{2p}{p-2}}(x)\,\mathrm{dvol}_x(g)\right)^{\frac{p-2}{2}} \le \int\limits_M |du|^2(x)\,\mathrm{dvol}_x(g) \text{ for all } u \in C_c^\infty(M) \tag{8.20}$$

and

$$\int\limits_M |\mathcal{R}_-|^{\frac{p}{2}}(x)\,\mathrm{dvol}_x(g) < \infty.$$

Then $D : W_0(E) \longrightarrow L_2(E)$ is Fredholm.

Proof Choose K so large that $|\mathcal{R}_-|_{L_{\frac{p}{2}}(M\setminus K)} \le \frac{c_p}{2}$. Then we obtain for $\varphi \in C_c^\infty(E|_{M\setminus K})$, $|\nabla\varphi|_{M\setminus K} \equiv |\nabla\varphi|_{L_2((T^*\otimes E)|_{M\setminus K})}$

$$\begin{aligned}
\int\limits_M |D\varphi|^2\,\mathrm{dvol}_x g &= \int\limits_M |\nabla\varphi|^2\,\mathrm{dvol}_x(g) + \langle\mathcal{R}\varphi,\varphi\rangle = \\
&= \frac{1}{2}|\nabla\varphi|^2_{M\setminus K} + \frac{1}{2}|\nabla\varphi|^2_{M\setminus K} + \langle\mathcal{R}\varphi,\varphi\rangle \ge \\
&\ge \frac{1}{2}|\nabla\varphi|^2_{M\setminus K} + \frac{c_p}{2}|\varphi|_{L_{\frac{2p}{p-2}}} - \int\limits_{M\setminus K} |\mathcal{R}_-|^{\frac{p}{2}}\cdot(|\varphi|^2)^{\frac{p}{p-2}}\,\mathrm{dvol}_x(g) \ge \\
&\ge \frac{1}{2}|\nabla\varphi|^2_{M\setminus K} + \frac{c_p}{2}|\varphi|^2_{L_{\frac{2p}{p-2}}} - |\mathcal{R}_-|_{L_{\frac{p}{2}}(M\setminus K)}\cdot|\varphi|^2_{L_{\frac{2p}{p-2}}} \ge \\
&\ge \frac{1}{2}|\nabla\varphi|^2_{L_2} \ge \\
&\ge \frac{1}{2}|d|\varphi||^2_{L_2(M\setminus K)}
\end{aligned}$$

and on $\frac{1}{2}|d|\varphi||^2_{L_2(M\setminus K)}$ we apply (8.19). □

Another important example are manifolds with a cylindrical end which we already mentioned. In this case, there is a compact submanifold with boundary $K \subset M$ such that $(M\setminus K, g)$ is isometric to $(]0,\infty[\times\partial K, dr^2 + g_{\partial K})$. One assumes that $(E,h)|_{]0,\infty[\times\partial K}$ also has product structure and $D|_{M\setminus K} = \nu\cdot\left(\frac{\partial}{\partial r} + A\right)$, where $\nu\cdot$ is the Clifford multiplication with the exterior normal at $\{\gamma\}\times\partial K$ and A is first order elliptic and self-adjoint on $E|_{\partial K}$.

Proposition 8.14 *D is non-parabolic at infinity.*

Proof There are two proofs. The first one refers to [6]. According to proposition 2.5 of [6], there exists on $M\setminus K$ a parametrix $Q : L_2(E|_{M\setminus K}) \longrightarrow \Omega^{2,1}_{\mathrm{loc}}E|_{M\setminus K}, D)$ such that $QD\varphi = \varphi$ for all $\varphi \in C_c^\infty(E|_{M\setminus K})$. Hence for $C_c^\infty(E|_{M\setminus K})$, $U \supset M\setminus K$ bounded,

$$|\varphi|_{L_2(E|_U)} = |QD\varphi|_{L_2(E|_U)} \le |Q|_{L_2\to\Omega^{2,1}}\cdot|D\varphi|_{L_2}.$$

The other proof is really elementary calculus. For $\varphi \in C_c^\infty(E|_{M\setminus K})$, $|\varphi(r,y)| = \left|\int_0^r \frac{\partial\varphi}{\partial r}\,dr\right| \le \sqrt{r}\cdot\left|\frac{\partial\varphi}{\partial r}\right|_{L_2}$. Hence $|\varphi|^2_{L_2(E|_{]0,T[\times\partial K})} \le \frac{T^2}{2}\left|\frac{\partial\varphi}{\partial r}\right|^2_{L_2} \le \frac{T^2}{2}\left(\left|\frac{\partial\varphi}{\partial r}\right|^2_{L_2} + |A\varphi|^2_{L_2}\right) = \frac{T^2}{2}|D\varphi|^2_{L_2}$. □

The authors of [6] define extended L_2-sections of $E|_{]0,\infty[\times\partial K}$ as sections $\varphi \in L_{2,\mathrm{loc}}$, $\varphi(r,y) = \varphi_0(r,y) + \varphi_\infty(y)$, $\varphi_0 \in L_2$, $\varphi_\infty \in \ker A$.

Proposition 8.15 *The extended solutions of $D\varphi = 0$ are exactly the solutions of $D\varphi = 0$ in W.*

Proof Let $\{\varphi_\lambda\}_{\lambda\in\sigma(A)}$ be a complete orthonormal system in $L_2(E|_{\partial K})$ consisting of the eigensections of A. Then we can a solution φ of $D\varphi = 0$ on $]0,\infty[\times\partial K$ decompose as

$$\varphi(r,y) = \sum_{\lambda\in\sigma(A)} c_\lambda e^{-\lambda r}\varphi_\lambda(y)$$

and $\varphi \in W$ if and only if $c_\lambda = 0$ for $\lambda < 0$. In this case

$$\varphi_0(r,y) = \sum_{\substack{\lambda\in\sigma(A)\\ \lambda>0}} c_\lambda e^{-\lambda r}\varphi_\lambda(y), \quad \varphi_\infty(y) = \sum_{\lambda\in\sigma(A)} c_{0,i}\varphi_{0,i}(y).$$

□

This proposition can also be reformulated as

Proposition 8.16 *Denote by $P_{\le 0}$ or $P_{<0}$ the spectral projection of A onto the sum of eigenspaces belonging to eigenvalues ≤ 0 or < 0, respectively. Then*

a) *φ is a solution in W of $D\varphi = 0$ if and only if*

$$D\varphi = 0 \text{ on } K$$

and

$$P_{<0}\varphi = 0 \text{ on } \partial K.$$

b) *φ is an L_2-solution of $D\varphi = 0$ if and only if*

$$D\varphi = 0 \text{ on } K$$

and

$$P_{\le 0}\varphi = 0 \text{ on } \partial K.$$

□

We finish the introductory discussion of the extended index at this point, having in mind that at this place we did not discuss its calculation, i.e. we did not present an index formula for the extended index.

As we mentioned at the beginning of this section, elliptic operators on open manifolds are in general not Fredholm, the classical analytical index is not defined, and the same holds for the topological index since the corresponding integral in general diverges. To save a big part of the classical approach, one restricts to very special cases which we pointed out in the preceding discussions, i.e. assuming the condition (8.3), or restricting to the locally symmetric case with finite volume or to manifolds and operators which are non-parabolic at infinity.

Another fundamental approach is to establish a new concept of Fredholmness and of the index by associating it with the K-theory of certain operator algebras and its pairing with cyclic cohomology.

This approach has been elaborated by J. Roe, G. Yu, N. Higson and others (cf. [94], [95], [96], [97], [98], [121]). The initial point is the general scheme of index theory.

Given a Riemannian manifold (M^n, g), Hermitean vector bundles E, $F \longrightarrow M$ and an elliptic operator $D : C_c^\infty(E) \longrightarrow C_c^\infty$, the general scheme of index theory consists of the following. One chooses a suitable extension of D, an operator algebra $\mathfrak{B}$, a functor $K_i(\mathfrak{B})$, constructs an element $\mathrm{Ind}D \subset K_i(\mathfrak{B})$, defines an element $\mathcal{I}_D$ of cohomological nature, mainly a differential form, defines pairings $\langle \mathcal{I}_D, m\rangle$, $\langle \mathrm{Ind}D, \xi\rangle$ and sets $\mathrm{ind}_t D = \langle \mathcal{I}_D, m\rangle$, $\mathrm{ind}_a D = \langle \mathrm{Ind}D, \xi\rangle$ is a pairing with a cyclic cohomology class. This gives a diagram

$$\begin{array}{ccc} D & \longrightarrow & \mathrm{Ind}D \in K_i(\mathfrak{B}) \\ \downarrow & & \downarrow \\ \mathcal{I}_D & \longrightarrow & \langle \mathcal{I}_D, m\rangle = \mathrm{ind}_t D \stackrel{?}{=} \mathrm{ind}_a D = \langle \mathrm{Ind}D, \xi\rangle. \end{array} \tag{8.21}$$

The commutative closure of the diagram by an equality $\mathrm{ind}_t D = \mathrm{ind}_a D$ means the establishing of an index theorem. In the classical example of (M^n, g) closed, oriented $\mathfrak{B} = \mathfrak{K}$ the algebra of compact operators, $\mathcal{I} = \mu_0(D^*D) - \mu_0(DD^*)$, where $\mu_0(\cdots)$ is the coefficient (= differential form) in the asymptotic expansion of the heat kernel on the diagonal without t-power, $\mathrm{Ind}D \in K_0(\mathfrak{K}) = [P_{\ker D}] - [P_{\mathrm{coker}\ D}]$, we obtain the diagram

$$\begin{array}{ccc} D & \longrightarrow & \mathrm{Ind}D \in K_0(\mathfrak{K}) \\ \downarrow & & \downarrow \text{trace} \\ \mathcal{I}_D = \mu_0(D^*D) - \mu_0(DD^*) & \underset{\int}{\longrightarrow} & \int \mathcal{I}_D = \mathrm{ind}_t = \mathrm{ind}_a D. \end{array} \tag{8.22}$$

In the open case, (8.22) does not make sense, and one has to make in (8.21) appropriate choices. This concerns the

1) choice of the D's

2) choice of $\mathfrak{B}$ and i,

3) construction of $\mathrm{Ind}D \in K_i(\mathfrak{B})$,

4) choice of $\xi \in HC^*(\mathfrak{B})$ s.t. $\langle \mathrm{Ind}D, \xi\rangle ='' \mathrm{ind}_a D''$ is well defined,

5) choice of $\mathcal{I}_D$ and the functional $\mathcal{I}_D \longrightarrow'' \mathrm{ind}_t D''$.

Thereafter one has to check the validity of the index theorem, the rigidity of the index and to present meaningful applications. This program has been performed e.g. in the papers above and by the author.

Initial point is the K-theory of operator algebras. Concerning this, we refer to the standard books. Let $0 \longrightarrow J \longrightarrow A \longrightarrow \mathfrak{B} = A/J \longrightarrow 0$ be an exact sequence of Banach algebras, $J \subset A$ a closed ideal. Then there is an exact 6-term sequence of K-groups

$$\begin{array}{ccccc} K_1(J) & \longrightarrow & K_1(A) & \longrightarrow & K_1(A/J) \\ \uparrow & & & & \downarrow \partial \\ K_0(A/J) & \longleftarrow & K_0(A) & \longleftarrow & K_0(J) \end{array} \qquad (8.23)$$

If one defines an operator T acting on A to be Fredholm (= J-Fredholm) if πT is invertible in A/J then πT gives rise to a class $[\pi T]$ in $K_1(A/J)$ and one could define

$$\mathrm{Ind} T := \partial[\pi T] \in K_0(J)$$

in analogy to (8.22).

At this stage, the decisive stage is the choice of A and J. In the classical compact case, one takes e.g.

J = algebra of smoothing operators,

A = algebra of pseudo-differential operators,

$\mathfrak{B} = A/J$ = algebra of complete symbols.

An elliptic operator D is not bounded. For this reason one considers $f(D)$, where $f : \mathbb{R} \longrightarrow [-1, 1]$ is an odd continuous function satisfying $f(t) \underset{t\to\pm\infty}{\longrightarrow} \pm 1$, a so called chopping function. Here $f(D)$ is defined e.g. by the Cauchy integral formula and is a pseudo-differential operator. If the underlying manifold is open then there is no canonical choice for A and J, and this choice is in fact a rather delicate matter. There are two conflicting wishes. J must be sufficiently large to contain the operators $1 - f(D)^2$ so that $\mathrm{Ind} D \in K(J)$ and J should be sufficiently small to permit some knowledge of its K-theory.

These choices and corresponding constructions have been performed by J. Roe, G. Yu in a series of papers. Having in mind the later definition of traces or more general the pairing with cyclic cohomology, J. Roe considers J as an algebra of bounded smoothing operators that are "uniformly nearly local" and generalizes this notion step by step.

In the sequel, we restrict to Clifford bundles $E \longrightarrow (M^n, g)$ satisfying (I), $(B_\infty(M))$, $(B_\infty(E))$ and D = associated generalized Dirac operator. An operator $A : \Omega^{0,2,k}(E) \longrightarrow \Omega^{0,2,l}(E)$ is called uniformly nearly local if

$$\int\limits_{\mathrm{dist}(x,\mathrm{supp}\ u)>R} |(Au)(x)|^2 dx \le F(R)|u|^2_{L_2}$$

where $F(R) \underset{R\to\infty}{\longrightarrow} 0$. An operator $A : C_c^\infty(E) \longrightarrow C^\infty(E)$ is called a uniform operator of order $\le k$ if for each r it has a continuous extension from $\Omega^{0,2,r}(E) \longrightarrow \Omega^{0,2,r-k}$ as a uniformly nearly local operator. Denote the corresponding collection by $U_k(E)$. All $U_k(E)$ form a filtered algebra $U(E)$ with ideal $U_{-\infty}(E) = \bigcap\limits_k U_k(E)$.

Proposition 8.17 *Each $A \in U_{-\infty}(E)$ has a uniformly bounded smooth Schwartz kernel $\mathcal{K}(x, y)$.*

The proof is the same as that of

Proposition 8.18 *Let $f \in \mathcal{S}(\mathbb{R})$ (= Schwartz class). Then*

$$f(D) = \frac{1}{2\pi} \int \hat{f}(t) e^{-itD} dt$$

belongs to $U_{-\infty}(E)$.

We refer to [94] for the proof. □

Consider more generally the space

$$\mathcal{S}^m(\mathbb{R}) = \{f \in C^\infty(R) | \, |f^{(k)}(\lambda)| \le c_k(1+|\lambda|)^{-k}\}.$$

Then $\mathcal{S}(\mathbb{R}) = \bigcap \mathcal{S}^m(\mathbb{R})$, and $\mathcal{S}^m(\mathbb{R})$ is a Frechet space with the best c_k as semi-norms. Proposition 8.18 generalizes to

Proposition 8.19 *For $f \in \mathcal{S}^m(\mathbb{R})$, $f(D)$ is (defined by the spectral theorem) an element of $U_m(E)$.*

Proof One proves this first for $f \in \mathcal{S}(\mathbb{R})$ and $f(D) = \frac{1}{2\pi}\int \hat{f}(t)e^{itD}dt$ and applies thereafter an approximation technique. □

Suppose now E equipped with a grading η and let D be the associated (compatible with η) generalized Dirac operator.

Proposition 8.20 *D is abstractly elliptic between the $U(E)$-modules given by the eigen-projections $(1+\eta)/2$ and $(1-\eta)/2$.*

Proof We must construct a parametrix. Let $\omega \in C_c^\infty(\mathbb{R})$, $\omega(x) = 1$ in a neighborhood of zero and set $\psi(x) = (1-\omega(x^2))/x^2$. Then $\psi \in \mathcal{S}^{-2}(\mathbb{R})$, hence $\psi(D) \in U_{-2}(E)$, $D\psi(D) \in U_{-1}(E)$. According to 8.18, $\omega(D^2) \in U_{-\infty}(E)$ and it commutes with D, whereas $D\psi(D)$ anticommutes with η. Now the equation

$$D(D\psi(D)) = D^2\psi(D) = 1 - \omega(D^2)$$

shows that $D\psi(D)$ is a parametrix for D. □

Corollary 8.21 *D has an index $\mathrm{Ind}D \in K_0(U_{-\infty}(E))$.*

Proof Consider the sequence $0 \longrightarrow U_{-\infty}(E) \longrightarrow U(E) \longrightarrow U(E)/U_{-\infty}(E) \longrightarrow 0$. D is invertible in $U(E)/U_{-\infty}(E)$. ∂ in (6.70) then yields an element $\mathrm{Ind}D \in K_0(U_{-\infty}(E))$. □

Remark 8.22 J. Roe proves in [94] the simple fact that any Clifford bundle E of bounded geometry has a good inclusion $i : E \longrightarrow$ trivial bundle and one can consider $\mathrm{Ind}D$ as an element of $K_0(U_{-\infty}(M))$, too. □

If we consider $0 \longrightarrow J \longrightarrow A \longrightarrow A/J = \mathfrak{B} \longrightarrow 0$, T is A/J invertible (i.e. abstractly elliptic) then $[\pi T] \in K_1(A(/J) = K_1(\mathfrak{B})$ and $\mathrm{Ind} = \partial[\pi T] \in K_0(J)$ and it would be desirable to associate to $\mathrm{Ind}T$ a "honest number". This can be in fact performed (under certain assumptions) in two steps, 1. construction of a trace τ (= element of the 0-dimensional cyclic cohomology of J) and pairing it with $\mathrm{Ind}D$, thus getting a "dimension homomorphism" $\dim_\tau : K_0(J) \longrightarrow$ ground field k. Here a k-linear functional $\tau : J \longrightarrow k$ is called a trace if $\tau(j_1 j_2) = \tau(j_2 j_1)$. If M is a projective $\mathfrak{B}$-module and $\mathrm{End}_J(M)$ the algebra of $\mathfrak{B}$-linear morphisms that map M to $M \otimes J$ then a trace τ on J can easily be extended to a trace on $\mathrm{End}_J(M)$ (representing the latter by matrices) then yielding $\dim_\tau : K_0(J) \longrightarrow k$.

Lemma 8.23 *If P is an abstract elliptic operator with parametrix Q then*

$$\dim_\tau(\mathrm{Ind}P) = \tau(1-QP)^2 - \tau(1-PQ)^2.$$

We refer to [94] for the proof. □

For practical applications we must now find a trace τ. Suppose (M^n, g) open, oriented, of bounded geometry. An exhaustion $\{M_i\}_i$, $M_1 \subset M_2 \subset \cdots$, $\bigcup M_i = M$ by compact subsets will be called regular if for each $r \geq 0$

$$\lim_{i\to\infty} [\text{vol } (\text{Pen}^+(M_i, r))/\text{vol } (\text{Pen}^-(M_i, r))] = 1,$$

where

$$\text{Pen}^+(K, r) = \text{CL}(\bigcup_{x\in K} B_r(x)),$$
$$\text{Pen}^-(K, r) = \text{CL}(M \setminus \text{Pen}^+(M \setminus K, r)).$$

Examples 8.24 **1)** $(\mathbb{R}^n, g_{st})$,

2) $(\tilde{M}, \tilde{g}) \xrightarrow{P} (M, g)$ with M closed and $\text{Deck}(\tilde{M} \longrightarrow M)$ amenable and

3) manifolds with subexponential growth (for all $c > 0$, $e^{-cr}\text{vol } (B_r(x_0)) \underset{r\to\infty}{\longrightarrow} 0$) admit regular exhaustions.

4) Hyperbolic n-space H^n_{-1} does not admit a regular exhaustion. □

We call (M^n, g) closed at infinity if for any $\lambda \in C^0(M)$ s.t. $0 < c^{-1} < \lambda < c$, c a constant, the form $\lambda \cdot$ dvol generates a nontrivial cohomology class $\in {}^bH^n(M)$.

A positive continuous linear functional $\mathfrak{m} : {}^b\Omega^n(M) \longrightarrow \mathbb{R}$ is called a fundamental class if $\langle \mathfrak{m}, \text{dvol}\rangle \neq 0$ and $\mathfrak{m} \circ d = 0$.

Proposition 8.25 *(M^n, g) has a fundamental class if and only if ist is closed at infinity.*

This follows immediately from the Hahn-Banach extension theorem. □

Examples 8.26 **1)** Any closed manifold and the $(\mathbb{R}^n, g_{st})$ are closed at infinity,

2) the hyperbolic n-space H^n_{-1} is not. □

An element $\mathfrak{m} \in ({}^b\Omega^n(M))^*$ is said to be associated to a regular exhaustion $\{M_i\}_i$ of (M^n, g) if for each bounded n-form α

$$\lim_{i\to\infty} \left| \langle \mathfrak{m}, \alpha\rangle - \frac{1}{\text{vol } (M_i)} \int_{M_i} \alpha \right| = 0.$$

Proposition 8.27 *Every regular exhaustion defines such functionals.*

Proof Suppose $\{M_i\}_i$, α be given. Set $\langle \mathfrak{m}_i, \alpha\rangle := \frac{1}{\text{vol } M_i} \int \alpha$. Then $\mathfrak{m}_i \in \overline{B_1(0)} \subset ({}^b\Omega^n)^*$. According the Banach-Alaoglu theorem, $\overline{B_1(0)}$ is compact in the weak-*-topology, i.e. the sequence $(\mathfrak{m})_i$ has a limit point $\mathfrak{m}$. □

Proposition 8.28 *Any $\mathfrak{m}$ given by 8.27 is a fundamental class.* □

Let $\mathfrak{m}$ be a fundamental class associated to regular exhaustion $\{M_i\}_i$ and let $A \in U_{-\infty}(M)$. Then A has a uniquely determined bounded smoothing kernel $\mathcal{K}_A$ such that

$$Au(x) = \int \mathcal{K}_A(x, y)u(y)\, \text{dvol}_y(g).$$

Consider the bounded n-form $\alpha(x) = \mathcal{K}_A(x, x)\, \text{dvol}_x(g)$ and define $\tau(A) := \langle \mathfrak{m}, \alpha\rangle$.

Proposition 8.29 *The functional τ defines a trace on $U_{-\infty}(M)$.*

We refer to [94], [95] for the simple proof. □

Remark 8.30 In the same manner, a trace τ on $U_{-\infty}(E)$ can be defined by

$$\tau(A) = \langle \mathfrak{m}, (\operatorname{tr}_x \mathcal{K}_A(x,x)) \operatorname{dvol} \rangle$$

where tr_x means the fibrewise trace. □

Suppose as above $E \longrightarrow M$ of bounded geometry, η a grading, D the generalized Dirac operator, $\{M_i\}_i$ a regular exhaustion with associated $\mathfrak{m}$ and τ. According to lemma 8.23, we get a dimension homomorphism

$$\dim_\tau : K_0(U_{-\infty}(M)) \longrightarrow \mathbb{R}$$

or

$$\dim_\tau : K_0(U_{-\infty}(E)) \longrightarrow \mathbb{R},$$

respectively. There remains the task to compute $\dim_\tau(\mathrm{Ind} D)$.

Proposition 8.31 *If $f \in \mathcal{S}(\mathbb{R}^+)$ with $f(0) = 1$ then*

$$\dim_\tau(\mathrm{Ind} D) = \tau(\eta f(D^2)). \tag{8.24}$$

We refer to [94] for the proof which is an easy calculation. □

Setting in 8.31 $f(\lambda) = e^{-t\lambda}$, we have the well known asymptotic expansion

$$\mathcal{K}(t,x,x) \underset{t \to 0^+}{\sim} \sum_{k \geq 0} t^{(k-n)/2} b_k(x) \tag{8.25}$$

where the b_k are smooth sections of $\operatorname{End}(E) \otimes \Lambda^n T^* M$. Combining (8.24) and (8.25), we obtain

$$\dim_\tau(\mathrm{Ind} D^+) \sim \sum_{k \geq 0} t^{(k-n)/2} \langle \mathfrak{m}, \operatorname{tr}(\eta b_k) \rangle. \tag{8.26}$$

The left hand side of (8.26) is independent of t. Hence on the right hand side can appear only the term without t-power, i.e. $\langle \mathfrak{m}, \operatorname{tr} \eta b_n \rangle \equiv \langle \mathfrak{m}, \mathcal{I}_D \rangle$. This yields the index theorem.

Theorem 8.32 *Suppose $(E, h, \nabla, \eta) \longrightarrow (M^n, g)$ is a $\mathbb{Z}_2$-graded Clifford bundle satisfying (I), $(B_\infty(M))$, $(B_\infty(E, \nabla)$, admitting a regular exhaustion $\{M_i\}_i$ with associated fundamental class $\mathfrak{m}$. If D is the associated generalized Dirac operator then its index $\mathrm{Ind} D \in K_0(U_{-\infty}(E))$ satisfies the equation*

$$\dim_\tau(\mathrm{Ind} D) = \langle \mathfrak{m}, \mathcal{I}(D) \rangle.$$

□

It is natural to ask for meanigful applications of (8.26). This means in particular assertions concerning $\ker D^+$, $\ker D^-$. We rewrite (8.24),

$$\dim_\tau(\mathrm{Ind} D) = \tau(f(D^- D^+)) - \tau(f(D^+ D^-)), \tag{8.27}$$

where $f \in \mathcal{S}(\mathbb{R}^+)$ with $f(0) = 1$. If $\inf \sigma_e(D^2|_{(\ker D^2)^\perp}) > 0$, i.e. one would have a spectral gap for D above and below zero, then one could choose f with compact support such that $= 1$ and $f(\lambda) = 0$ for all nonzero eigenvalues of D. In this case, $f(D^-D^+)$ is the projection $P_{\ker D^+}$ and $f(D^+D^-)$ is the projection $P_{\ker D^-}$, hence

$$\dim_\tau(\text{Ind}D) = \tau(P_{\ker D^+}) - \tau(P_{\ker D^-}).$$

Unfortunately, the existence of such a spectral gap is not equivalent to the existence of a regular exhaustion, as at least many geometric examples indicate. Hence one has to argue much more careful. Consider again $f_t(D^-D^+) = e^{-tD^-D^+}$. This converges for $t \to \infty$ to $P_{\ker D^+}$ in the strong operator topology on L_2. If this convergence would hold in the uniform topology then (8.27) would hold since $\mathfrak{m}$ is continuous against the uniform topology. So the presentation of substantial applications amounts to find out a "thin" area where one can assure (8.27). Consider the Weitzenboeck formula $D^2 = \nabla^*\nabla + \mathcal{R}$ and $\mathcal{R}^\pm = \mathcal{R}|_{\pm \text{ eigenspace of } \eta}$.

A subset $L \subset M$ is said to have density 0 if for all $\omega \in {}^b\Omega^n(M)$ with support $\omega \subseteq L$ there holds $\langle \mathfrak{m}, \omega \rangle = 0$. It has density 1 if $M \setminus L$ has density zero. $L \subset M$ is called small if for all $r > 0$ $Pen^+(L, r)$ has density 0. L is called large if $M \setminus L$ is small.

Proposition 8.33 *Suppose the hypothesis of 8.32.*

a) *If* $\mathcal{R}^+ \geq 0$ *on a large set, then* $\dim_\tau(\text{Ind}) \leq 0$.

b) *If* $\mathcal{R}^- \geq 0$ *on a large set, then* $\dim_\tau(\text{Ind}) \geq 0$.

We refer to [95] for some special applications to the Gauß-Bonnet, Dirac and Dolbeault operator. Until now, we only considered traces τ as functionals on K_0. In [196], [197] J. Roe considered much more general algebras and ideals (instead of $U(E)$, $U_{-\infty}(E)E$ and cyclic cocycles. We will not repeat here what is written there but we will only indicate here the main lines of this further approach.

Let again $E \longrightarrow M$ be a graded Clifford bundle with grading operator η, $E = E_+ \oplus E_-$, and generalized Dirac operator $D = \begin{pmatrix} 0 & D_- \\ D_+ & 0 \end{pmatrix}$. A positive operator b acting on $L_2(E)$ is called locally traceable if for all $f \in C_c(M)$ the operator fbf is of trace class. A general operator is called locally traceable if it is a finite linear combination of positive locally traceable operators. An operator b is said to have bounded propagation if there exists $r > 0$ s.t. for any $u \in L_2(E))$

$$\text{supp }(bu) \cup \text{supp }(b^*u) \subseteq \{x \in M | \text{dist}(x, \text{supp } u) \leq r\}.$$

The Roe algebra $\mathcal{B} = \mathcal{B}_E$ consists of all locally traceable operators with bounded propagation. Denote by $\overline{\mathcal{B}}_E$ the norm closure of $\mathcal{B}$.

Lemma 8.34 *If f is chopping function then*

$$f(D)^2 - 1 \in \overline{\mathcal{B}}_E.$$

□

f odd implies the decomposition

$$f(D) = \begin{pmatrix} 0 & f(D)_- \\ f(D)_+ & 0 \end{pmatrix}$$

and we see that $f(D)_-$ is a parametrix of $f(D)_+$: $f(D)_+f(D)_- - 1 = C_-$, $f(D)_-f(D)_+ = C_+$, where

$$C = \begin{pmatrix} C_+ & 0 \\ 0 & C_- \end{pmatrix} \in \overline{\mathcal{B}}_E.$$

Set

$$L = \begin{pmatrix} C_+ & f(D)_- - C_+f(D)_- \\ f(D)_+ & C_- \end{pmatrix} \in \overline{\mathcal{B}}_E.$$

Then L is invertible, and

$$L^{-1} = \begin{pmatrix} C_+ & f(D)_-(1 - C_{)} \\ f(D)_+ & -C_- \end{pmatrix} \in \overline{\mathcal{B}}_E.$$

Define the index $\mathrm{Ind}D$ as

$$\begin{aligned} \mathrm{Ind}D &:= \left[L \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} L^{-1} - \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right] \\ &= \left[\begin{pmatrix} C_+^2 & f(D)_-C_-(1 - C_{)} \\ f(D)_+C_+ & -C_-^2 \end{pmatrix} \right] \in K_0(\overline{\mathcal{B}}_E). \end{aligned} \tag{8.28}$$

$\mathrm{Ind}D$ is independent of f. The motivation for the definition (8.28) comes from operator K-theory. Let A_E be the C^*-algebra of bounded operators acting on $L_2(E)$ with bounded propagation. Then $\overline{\mathcal{B}}_E$ is a $*$-ideal of A_E. Operator K-theory yields the exact sequence (6.70), in our case

$$\begin{array}{ccccc} K_1(\overline{\mathcal{B}}_E) & \longrightarrow & K_1(A_E) & \longrightarrow & K_1(A_E/\overline{\mathcal{B}}_E) \\ \uparrow & & & & \downarrow \partial \\ K_0(A_E/\overline{\mathcal{B}}_E) & \longleftarrow & K_0(A_E) & \longleftarrow & K_0(\overline{\mathcal{B}}_E) \end{array}$$

Now $[f(D)] = [L]$ in $A_E/\overline{\mathcal{B}}_E$ and in fact

$$\mathrm{Ind}D = \partial[f(D)].$$

We remark without proof (cf. [121]) the following

Proposition 8.35 $\mathrm{Ind}D$ *is equivalent to the Kasparov module* $(\overline{\mathcal{B}}_E, f(D), \Phi)$ *for* $(\mathbb{C}, \overline{\mathcal{B}}_E)$ *by the natural identification of* $K_0(\overline{\mathcal{B}}_E)$ *with* $KK(\mathbb{C}, \overline{\mathcal{B}}_E)$. □

Now the point is to produce cyclic cocycles. Here comes in a fundamental idea of Roe, namely to produce cyclic cocycles from coarse cohomology $HX^*(M)$ (cf. [98]).

J. Roe defines in [98] a cyclic character map $\chi : HX^{2q}(M) \longrightarrow HC^{2q}(\overline{\mathcal{B}}_E)$. Moreover, one has the Connes' pairing between $K_0(\overline{\mathcal{B}})$ and $HC^{even}(\overline{\mathcal{B}})$. Hence any $[\varphi] \in HX^{2q}(M)$ yields a number

$$\mathrm{Ind}_\varphi D = \langle \chi\varphi, \mathrm{Ind}D \rangle.$$

There is a character map $c : HX^q(M) \longrightarrow H_c^q(M)$ which sends a cocycle φ to its truncation to any penumbra $\mathrm{Pen}(\Delta, R)$ of the multidiagonal $\Delta \subset M \times \cdots \times M$ ($q+1$ times).

Now we can state the even index theorem of J. Roe.

Theorem 8.36 *Let (M^{2m}, g) be complete, $E \longrightarrow M$ a graded Clifford bundle and D the associated graded generalized Dirac operator. If $[\varphi] \in HX^{2q}(M)$ then*

$$\langle \chi[\varphi], \mathrm{Ind}D \rangle = \frac{q!}{(2q)!(2\pi i)^q} \langle \mathcal{I}_D \cup c[\varphi], [M] \rangle.$$

A similar theorem holds in the odd case with $\mathrm{Ind}D \in K_{-1}(\overline{\mathcal{B}}_E)$.

Theorem 8.37 *Let (M^{2m-1}, g) be complete, $E \longrightarrow M$ an (ungraded) Clifford bundle and D the associated generalized Dirac operator. If $[\varphi] \in HX^{2q-1}(M)$ then*

$$\langle \chi[\varphi], \mathrm{Ind}D \rangle = \frac{q!}{(2q-1)!(2\pi i)^q} \langle \mathcal{I}_D \cup c[\varphi], [M] \rangle.$$

We refer to [98] for the proof. □

We now briefly present our approach to relative index theorems on open manifolds.

Let (M^n, g) be closed, oriented $(E, h, \nabla, \cdot, \tau) \longrightarrow (M^n, g)$ a supersymmetric Clifford bundle with involution τ and $D = \begin{pmatrix} 0 & D^- \\ D^+ & 0 \end{pmatrix}$ the associated generalized Dirac operator. Then

$$\mathrm{ind}_a D^+ = \mathrm{tr}(\tau e^{-tD^2}).$$

Starting with this simple fact, one could attempt to define in the open case a relative index for a pair of generalized Dirac operators D, D' by

$$\mathrm{ind}(D, D') := \mathrm{tr}(\tau(e^{-tD^2} - e^{-tD'^2})).$$

With this intention in mind there arise immediately several problems.

1) One has to assure that D, D' are self-adjoint in the same Hilbert-space.

2) One has to assure that $e^{-tD^2} - e^{-tD'^2}$ is of trace class.

3) One has to assure that $\mathrm{tr}(\tau(e^{-tD^2} - e^{-tD'^2}))$ is independent of t.

4) Finally one has to present substantial applications.

The initial data for a fixed vector bundle $E \longrightarrow M$ and different Clifford structures are $(E, h, \nabla = \nabla^h, \cdot) \longrightarrow (M^n, g)$ and $(E, h', \nabla' = \nabla^{h'}, \cdot') \longrightarrow (M^n, g')$, respectively. These yield generalized Dirac operators $D = D(h, \nabla, \cdot, g)$ and $D' = D(h', \nabla', \cdot', g')$. D and D' act in different Hilbert spaces, i.e. $e^{-tD^2} - e^{-tD'^2}$ is not defined. But this can be repared by two unitary transformations. Denote by D' already the result after performing this transformations.

To describe the possibly maximal perturbations $(h', \nabla', \cdot', g')$ of $(h, \nabla, \cdot, g)$, we introduce in [52] uniform structures of Clifford structures and defined generalized components. We indicate here briefly the main definitions. First of all, we restrict to manifolds and bundles $(E, h, \nabla) \longrightarrow (M^n, g)$ of bounded geometry of order k i.e. we assume

$$r_{\mathrm{inj}}(M^n, g) > 0, \tag{I}$$
$$|(\nabla^g)^i R^g| \le C_i, \quad 0 \le i \le k, \tag{$B_k(M,g)$}$$

$$|\nabla^i R^E| \le D_i, \quad 0 \le i \le k. \tag{$B_k(E,\nabla)$}$$

Then we defined a local Lipschitz-Sobolev distance $d^{p,r}_{L,diff,rel}(E_1, E_2) = d^{p,r}_{L,diff,rel}((E_1, h_1, \nabla^{h_1}, \cdot_1, g_1), (E_2, h_2, \nabla^{h_2}, \cdot_2, g_2))$, an associated metrizable uniform structure and considered generalized components $\text{gen comp}^{p,r}_{L,diff,rel}(E)$. Roughly speaking, $E' \in \text{gen comp}^{p,r}_{L,diff,rel}(E)$ means that $E|_{M\setminus K} \cong E'|_{M'\setminus K'}$ as vector bundles by means of a bounded diffeomorphism, on $M \setminus K \cong M' \setminus K'$ h and h' are quasiisometric, g and g' are quasiisometric and the Sobolev distances $|h - h'|_{g,h,\nabla^h,p,r}$, $|\nabla - \nabla'|_{g,h,\nabla^h,p,r}$, $|\cdot - \cdot'|_{g,h,\nabla^h,p,r}$, $|g - g'|_{g,p,r}$ are finite.

Denote by $\text{CL}\mathcal{B}^{N,n}(I, B_k)$ the set of (Clifford-)isometry classes over n-manifolds with fibre dimension N and with bounded geometry of order k. For $E' \in \text{gen comp}(E)$, D and D' are defined in different Hilbert spaces. But there is a common Hilbert space H which contains $\mathcal{D}_D$ and $\mathcal{D}_{\tilde{D}'}$, ($\tilde{D}'$ is D' after some unitary transformation) as subspaces. Denote by P and P' the corresponding orthogonal projections. The key for everything is the following main theorem.

Theorem 8.38 *Let* $E = ((E, h, \nabla^h) \longrightarrow (M^n, g)) \in \text{CL}\mathcal{B}^{N,n}(I, B_k)$, $k \ge r+1 > n+3$, $E' \in \text{gen comp}^{1,r+1}_{L,diff,rel}(E) \cap \text{CL}\mathcal{B}^{N,n}(I, B_k)$. *Then for* $t > 0$

$$e^{-tD^2} P - e^{-t\tilde{D}'^2} P' \tag{8.29}$$

and

$$e^{-tD^2} D - e^{-t\tilde{D}'^2} \tilde{D}' \tag{8.30}$$

are of trace class and their trace norms are uniformly bounded on any t*-intervall* $[a_0, a_1]$, $a_0 > 0$.

We refer to [52], [43] for the proof which occupies 60 pages. □

We conclude that after fixing $E \in \text{CL}\mathcal{B}^{N,n}(I, B_k)$, $k \ge r + 1 > n + 3$, we can attach to any $E' \in \text{gen comp}^{1,r+1}_{L,diff,rel}(E)$ a number-valued invariant, namely

$$E' \longrightarrow \text{tr}(e^{-tD^2} P - e^{-t\tilde{D}'^2} P'.$$

This is a contribution to the classification inside a component but still unsatisfactory insofar as it

1) could depend on t.

2) will depend on the $K \subset M$, $K' \subset M'$ in question,

3) is not yet clear the meaning of this invariant.

We are in a much nore comfortable situation if we additionally assume that the Clifford bundles under consideration are endowed with an involution $\tau : E \longrightarrow E$, s.t.

$$\tau^2 = 1, \quad \tau^* = \tau \tag{8.31}$$

$$[\tau, X]_+ = 0 \text{ for } X \in TM \tag{8.32}$$

$$[\nabla, \tau] = 0 \tag{8.33}$$

Then $L_2((M, E), g, h) = L_2(M, E^+) \oplus L_2(M, E^-)$

$$D = \begin{pmatrix} 0 & D^- \\ D^+ & 0 \end{pmatrix}$$

and $D^- = (D^+)^*$. If M^n is compact then as usual

$$\mathrm{ind}D := \mathrm{ind}D^+ := \dim\ker\ D^+ - \dim\ker\ D^- \equiv \mathrm{tr}(\tau e^{-tD^2}), \tag{8.34}$$

where we understand τ as

$$\tau = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}.$$

For open M^n $\mathrm{ind}D$ in general is not defined since τe^{-tD^2} is not of trace class. The appropriate approach on open manifolds is relative index theory for pairs of operators D, D'. If D, D' are selfadjoint in the same Hilbert space and $e^{tD^2} - e^{-tD'^2}$ would be of trace class then

$$\mathrm{ind}(D, D') := \mathrm{tr}(\tau(e^{-tD^2} - e^{-tD'^2})) \tag{8.35}$$

makes sense, but at the first glance (8.35) should depend on t.

If we restrict to Clifford bundles $E \in \mathrm{CL}\mathcal{B}^{N,n}(I, B_k)$ with involution τ then we assume that the maps entering in the definition of $\mathrm{comp}^{1,r+1}_{L,diff,F}(E)$ or $\mathrm{gen\,comp}^{1,r+1}_{L,diff,rel}(E)$ are τ-compatible, i.e. after identification of $E|_{M\setminus K}$ and $f_E^* E'|_{M'\setminus K}$ holds

$$[f_E^*\nabla^{h'}, \tau] = 0, \quad [f^* \cdot', \tau]_+ = 0. \tag{8.36}$$

We call $E|_{M\setminus K}$ and $E'|_{M'\setminus K'}$ τ-compatible. Then, according to the preceding theorems,

$$\mathrm{tr}(\tau(e^{-tD^2}P - e^{-t\tilde{D}'^2}P')) \tag{8.37}$$

makes sense.

Theorem 8.39 *Let $((E, h, \nabla^h) \longrightarrow (M^n, g), \tau) \in \mathrm{CL}\mathcal{B}^{N,n}(I, B_k)$ be a graded Clifford bundle, $k \geq r > n+2$.*

a) *If $\nabla'^h \in \mathrm{comp}^{1,r}(\nabla) \subset C_E^{1,r}(B_k)$, ∇' τ-compatible, i.e. $[\nabla', \tau] = 0$ then*

$$\mathrm{tr}(\tau(e^{-tD^2} - e^{-tD'^2}))$$

is independent of t.

b) *If $E' \in \mathrm{gen\,comp}^{1,r+1}_{L,diff,rel}(E)$ is τ-compatible with E, i.e. $[\tau, X\cdot']_+ = 0$ for $X \in TM$ and $[\nabla', \tau] = 0$, then*

$$\mathrm{tr}(\tau(e^{-tD^2}P - e^{-t\tilde{D}'^2}P'))$$

is independent of t. □

We denote $Q^\pm = D^\pm$

$$Q = \begin{pmatrix} 0 & Q^+ \\ Q^- & 0 \end{pmatrix}, \quad H = \begin{pmatrix} H^+ & 0 \\ 0 & H^- \end{pmatrix} = \begin{pmatrix} Q^-Q^+ & 0 \\ 0 & Q^+Q^- \end{pmatrix} = Q^2, \tag{8.38}$$

$Q'^\pm = \tilde{D}'^\pm$, Q', H' analogous, assuming (8.31) – (8.33) as before and $\cdot'$, ∇' τ-compatible. H, H' form by definition a supersymmetric scattering system if the wave operators

$$W^\mp(H, H') := \lim_{t\to\mp\infty} e^{itH}e^{-tH'} \cdot P_{ac}(H') \text{ exist and are complete} \tag{8.39}$$

and

$$QW^{\mp}(H,H') = W^{\mp}(H,H')H' \text{ on } \mathcal{D}_{H'} \cap \mathcal{H}'_{ac}(H'). \tag{8.40}$$

Here $P_{ac}(H')$ denotes the projection on the absolutely continuous subspace $\mathcal{H}'_{ac}(H') \subset \mathcal{H}$ of H'.

A well known sufficient criterion for forming a supersymmetric scattering system is given by

Proposition 8.40 *Assume for the graded operators Q, Q' (= supercharges)*

$$e^{-tH} - e^{-tH'}$$

and

$$e^{-tH}Q - e^{-tH'}Q$$

are for $t > 0$ of trace class. Then they form a supersymmetric scattering system. □

Corollary 8.41 *Assume the hypotheses of 8.39. Then D, $\tilde{D}'$ form a supersymmetric scattering system, respectively. In particular, the restriction of D, D' to their absolutely continuous spectral subspaces are unitarily equivalent, respectively.* □

Until now we have seen that under the hypotheses of 8.39

$$\operatorname{ind}(D, \tilde{D}') = \operatorname{tr} \tau(e^{-tD^2}P - e^{-t\tilde{D}'^2}P') \tag{8.41}$$

is a well defined number, independent of $t > 0$ and hence yields an invariant of the pair (E, E'), still depending on K, K'. Hence we should sometimes better write

$$\operatorname{ind}(D, \tilde{D}', K, K'). \tag{8.42}$$

We want to express in some good cases $\operatorname{ind}(D, \tilde{D}', K, K')$ by other relevant numbers. Consider the abstract setting (8.38). If $\inf \sigma_e(H) > 0$ then $\operatorname{ind}D := \operatorname{ind}D^+$ is well defined.

Lemma 8.42 *If $e^{-tH}P - e^{-tH'}P'$ is of trace class for all $t > 0$ and $\inf \sigma_e(H)$, $\inf \sigma_e(H') > 0$ then*

$$\lim_{t\to\infty} \operatorname{tr} \tau(e^{-tH}P - e^{-tH'}P') = \operatorname{ind}Q^+ - \operatorname{ind}Q^-. \tag{8.43}$$

□

We infer from this

Theorem 8.43 *Assume the hypotheses of 8.39 and $\inf \sigma_e(D^2) > 0$. Then $\inf \sigma_e({D'}^2) > 0$ and for each $t > 0$*

$$\operatorname{tr} \tau(e^{-tD^2} - e^{-t\tilde{D}'^2}) = \operatorname{ind}D^+ - \operatorname{ind}D'^+. \tag{8.44}$$

□

It would be desirable to express $\operatorname{ind}(D, \tilde{D}', K, K')$ by geometric topological terms. In particular, this would be nice in the case $\inf \sigma_e(D^2) > 0$. In the compact case, one sets

$\text{ind}_a D := \text{ind}_a D^+ = \dim\ker D^+ - \dim\ker(D^+)^* = \dim\ker D^+ - \dim\ker D^- = \lim_{t\to\infty} \text{tr}\,\tau e^{-tD^2}$. On the other hand, for $t \to 0^+$ there exists the well known asymptotic expansion for the kernel of τe^{-tD^2}. Its integral at the diagonal yields the trace. If $\text{tr}\,\tau e^{-tD^2}$ is independent of t (as in the compact case), we get the index theorem where the integrand appearing in the L_2-trace consists only of the t-free term of the asymptotic expansion. Here one would like to express things in the asymptotic expansion of the heat kernel of $e^{-tD'^2}$ instead of $e^{-t\tilde{D}'^2}$. For this reason we restrict in the definition of the topological index to the case $E' \in \text{comp}^{1,r+1}_{L,diff,F}(E)$ or $E' \in \text{comp}^{1,r+1}_{L,diff,F,rel}(E)$, i.e. we admit Sobolev perturbation of $g, \nabla^h, \cdot$ but the fibre metric h should remain fixed. Then for $D' = D(g', h, \nabla'^h, \cdot')$ in $L_2((M,E), g, h)$ the heat kernel of $e^{-t(U^*D'U)^2} = U^* e^{-tD'^2} U$ equals to $\alpha(q)^{-\frac{1}{2}} W'(t,q,p)\alpha(p)^{\frac{1}{2}}$. At the diagonal this equals to $W'(t,m,m)$, i.e. the asymptotic expansion at the diagonal of the original $e^{-tD'^2}$ and the transformed to $L_2((M,E), g, h)$ coincide.

Consider

$$\text{tr}\,\tau W(t,m,m) \underset{t\to 0^+}{\sim} t^{-\frac{n}{2}} b_{-\frac{n}{2}}(D,m) + \cdots + b_0(D,m) + \cdots \tag{8.45}$$

and

$$\text{tr}\,\tau W'(t,m,m) \underset{t\to 0^+}{\sim} t^{-\frac{n}{2}} b_{-\frac{n}{2}}(D',m) + \cdots + b_0(D',m) + \cdots . \tag{8.46}$$

We state without proof

Lemma 8.44

$$b_i(D,m) - b_i(D',m) \in L_1, \quad -\frac{n}{2} \le i \le 1. \tag{8.47}$$

□

Define for $E' \in \text{gen}\,\text{comp}^{1,r+1}_{L,diff,F}(E)$

$$\text{ind}_{top}(D,D') := \int_M b_0(D,m) - b_0(D',m). \tag{8.48}$$

According to (8.47), $\text{ind}_{top}(D,D')$ is well defined.

Theorem 8.45 *Assume* $E' \in \text{gen}\,\text{comp}^{1,r+1}_{L,diff,F,rel}(E)$

a) *Then*

$$\text{ind}(D,D',K,K') = \int_K b_0(D,m) - \int_{K'} b_0(D',m) + \tag{8.49}$$

$$+ \int_{M\setminus K = M'\setminus K'} b_0(D,m) - b_0(D',m). \tag{8.50}$$

b) *If* $E' \in \text{gen}\,\text{comp}^{1,r+1}_{L,diff,F}(E)$ *then*

$$\text{ind}(D,D') = \text{ind}_{top}(D,D'). \tag{8.51}$$

c) *If* $E' \in \mathrm{gen\,comp}^{1,r+1}_{L,diff,F}(E)$ *and* $\inf \sigma_e(D^2) > 0$ *then*

$$\mathrm{ind}_{top}(D, D') = \mathrm{ind}_a D - \mathrm{ind}_a D'. \tag{8.52}$$

Proof All this follows from 8.39, the asymptotic expansion, (8.47) and the fact that the L_2-trace of a trace class integral operator equals to the integral over the trace of the kernel. □

Remarks 8.46 **1)** If $E' \in \mathrm{gen\,comp}^{1,r+1}_{L,diff,rel}(E)$, g and g', ∇^h and ∇'^h, $\cdot$ and $\cdot'$ coincide in $V = M \setminus L = M' \setminus L'$, $L \supseteq K$, $L' \supseteq K'$, then we conclude from standard heat kernel estimates that

$$\int_V |W(t,m,m) - W'(t,m,m)| \; dm \le C \cdot e^{-\frac{d}{t}}$$

and obtain

$$\mathrm{ind}(D, D', L, L') = \int_L b_o(D, m) - \int_{L'} b_0(D', m).$$

This follows immediately from 8.45 a).

2) The point here is that we admit much more general perturbations than in preceding approaches to prove relative index theorems.

3. $\inf \sigma_e(D^2) > 0$ is an invariant of $\mathrm{gen\,comp}^{1,r+1}_{L,diff,F}(E)$. If we fix E, D as reference point in $\mathrm{gen\,comp}^{1,r+1}_{L,diff,F}(E)$ then 8.45 c) enables us to calculate the analytical index for all other D's in the component from $\mathrm{ind}D$ and a pure integration.

4) $\inf \sigma_e(D^2) > 0$ is satisfied e.g. if in $D^2 = \nabla^*\nabla + \mathcal{R}$ the operator $\mathcal{R}$ satisfies outside a compact K the condition

$$\mathcal{R} \ge \kappa_0 \cdot id, \kappa_0 > 0. \tag{8.53}$$

(8.53) is an invariant of $\mathrm{gen\,comp}^{1,r+1}_{L,diff,F}(E)$ (with possibly different K, κ_0).

□

It is possible that $\mathrm{ind}D$, $\mathrm{ind}D'$ are defined even if $0 \in \sigma_e$. For the corresponding relative index theorem we need the scattering index.

To define the scattering index and in the next section relative ζ-functions, we must introduce the spectral shift function of Birman/Krein/Yafaev. Let A, A' be bounded self adjoint operators, $V = A - A'$ of trace class, $R'(z) = (A' - z)^{-1}$. Then the spectral shift function

$$\xi(\lambda) = \xi(\lambda, A, A') := \pi^{-1} \lim_{\varepsilon \to 0} \arg \det(1 + VR'(\lambda + i\varepsilon))$$

exists for a.e. $\lambda \in \mathbb{R}$. $\xi(\lambda)$ is real valued, $\in L_1(\mathbb{R})$ and

$$\mathrm{tr}(A - A') = \int_{\mathbb{R}} \xi(\lambda) \; d\lambda, \quad |\xi|_{L_1} \le |A - A'|_1.$$

If $I(A, A')$ is the smallest interval containing $\sigma(A) \cup \sigma(A')$ then $\xi(\lambda) = 0$ for $\lambda \notin I(A, A')$.

Let

$$\mathcal{G} = \{f : \mathbb{R} \longrightarrow \mathbb{R} \mid f \in L_1 \quad \text{and} \quad \int_{\mathbb{R}} |\widehat{f}(p)|(1+|p|)\, dp < \infty\}.$$

Then for $\varphi \in \mathcal{G}$, $\varphi(A) - \varphi(A')$ is of trace class and

$$\operatorname{tr}(\varphi(A) - \varphi(A')) = \int_{\mathbb{R}} \varphi'(\lambda)\xi(\lambda)\, d\lambda.$$

We state without proof

Lemma 8.47 *Let $H, H' \geq 0$, selfadjoint in $\mathcal{H}$, $e^{-tH} - e^{-tH'}$ for $t > 0$ of trace class. Then there exist a unique function $\xi = \xi(\lambda) = \xi(\lambda, H, H') \in L_{1,loc}(\mathbb{R})$ such that for > 0, $e^{-t\lambda}\xi(\lambda) \in L_1(\mathbb{R})$ and the following holds.*

a) $\operatorname{tr}(e^{-tH} - e^{-tH'}) = -t \int_0^\infty e^{-t\lambda}\xi(\lambda)\, d\lambda.$

b) *For every $\varphi \in \mathcal{G}$, $\varphi(H) - \varphi(H')$ is of trace class and*

$$\operatorname{tr}(\varphi(H) - \varphi(H')) = \int_{\mathbb{R}} \varphi'(\lambda)\xi(\lambda)\, d\lambda.$$

c) $\xi(\lambda) = 0$ *for* $\lambda < 0$. □

We apply this to our case $E' \in \operatorname{gen\,comp}^{1,r+1}_{L,diff,rel}(E)$. According to 8.38, D and $U^*i^*D'iU$ form a supersymmetric scattering system, $H = D^2$, $H' = \tilde{D'}^2$. In this case

$$e^{2\pi i\xi(\lambda,H,H')} = \det S(\lambda),$$

where $S = (W^+)^*W^- = \int S(\lambda)\, dE'(\lambda)$ and $H'_{ac} = \int \lambda\, dE'(\lambda)$.

Let $P_d(D)$, $P_d(\tilde{D}')$ be the projector on the discrete subspace in $\mathcal{H}$, respectively and $P_c = 1 - P_d$ the projector onto the continuous subspace. Moreover we write

$$D^2 = \begin{pmatrix} H^+ & 0 \\ 0 & H^- \end{pmatrix}, \quad \tilde{D'}^2 = \begin{pmatrix} H'^+ & 0 \\ 0 & H'^- \end{pmatrix}.$$

We make the following additional assumption.

$$e^{-tD^2}P_d(D), e^{-t\tilde{D'}^2}P_d(\tilde{D}') \quad \text{are for } t > 0 \text{ of trace class.} \tag{8.54}$$

Then for $t > 0$

$$e^{-tD^2}P_c(D) - e^{-t(U^*i^*D'iU)^2}P_c(U^*i^*D'iU)$$

is of trace class and we can define

$$\begin{aligned} \xi^c(\lambda, H^\pm, H'^\pm) \quad := \quad & -\pi \lim_{\varepsilon\to 0^+} \arg\det[1 + (e^{-tH^\pm}P_c(H^\pm) - e^{-tH'^\pm}P_c(H'^\pm)) \\ & (e^{-tH'^\pm}P_c(H'^\pm) - e^{-\lambda t} - i\varepsilon)^{-1}] \end{aligned}$$

According to 8.47 a),

$$\operatorname{tr}(e^{-tH^{\pm}}P_c(H^{\pm}) - e^{-tH'^{\pm}}P_c(H'^{\pm})) = -t\int_0^\infty \xi^c(\lambda, H^{\pm}, H'^{\pm})e^{-t\lambda}\, d\lambda.$$

The assumption (8.54) in particular implies that for the restriction of D and $\tilde{D}'$ to their discrete subspace the analytical index is well defined and we write $\operatorname{ind}_{a,d}(D, \tilde{D}') = \operatorname{ind}_{a,d}(D) - \operatorname{ind}_{a,d}(\tilde{D}')$ for it. Set

$$n^c(\lambda, D, \tilde{D}') := -\xi^c(\lambda, H^+, H'^+) + \xi^c(\lambda, H^-, H'^-). \tag{8.55}$$

Theorem 8.48 *Assume the hypotheses of 8.39 and (8.54).*
Then $n^c(\lambda, D, \tilde{D}') = n^c(D, \tilde{D}')$ is constant and

$$\operatorname{ind}(D, \tilde{D}') - \operatorname{ind}_{a,d}(D, \tilde{D}') = n^c(D, \tilde{D}').$$

Proof

$$\begin{aligned}\operatorname{ind}(D, \tilde{D}') &= \operatorname{tr}\tau(e^{-tD^2}P - e^{-t\tilde{D}'^2}P') = \\ &= \operatorname{tr}\tau e^{-tD^2}P_d(D)P - \operatorname{tr}\tau e^{-t\tilde{D}'^2}P_d(\tilde{D}')P' + \\ &\quad + \operatorname{tr}\tau(e^{-tD^2}P_c(D) - e^{-t\tilde{D}'^2}P_c(\tilde{D}')) = \\ &= \operatorname{ind}_{a,d}(D, \tilde{D}') + t\int_0^\infty e^{-t\lambda}n^c(\lambda, D, \tilde{D}')\, d\lambda.\end{aligned}$$

According to 8.39, $\operatorname{ind}(D, \tilde{D}')$ is independent of t. The same holds for $\operatorname{ind}_{a,d}(D, \tilde{D}')$. Hence $t\int_0^\infty e^{-t\lambda}n^c(\lambda, D, \tilde{D}')\, d\lambda$ is independent of t. This is possible only if $\int_0^\infty e^{-t\lambda}n^c(\lambda, D, D')\, d\lambda = \frac{1}{t}$ or $n^c(\lambda, D, \tilde{D}')$ is independent of λ. □

Corollary 8.49 *Assume the hypotheses of 8.45 and additionally* $\inf \sigma_e(D^2|_{(\ker D^2)^\perp}) > 0$. *Then* $n^c(D, \tilde{D}') = 0$.

Proof In this case $\operatorname{ind}_{a,d}(D, \tilde{D}') = \operatorname{ind}D - \operatorname{ind}\tilde{D}' = \operatorname{ind}(D, \tilde{D}')$, hence $n^c = 0$. □

This finishes the outline of our relative index theory.

9 The continuity method for non-linear PDEs on open manifolds

The heart of the continuity method consists in the following. Given a (possibly non-linear) PDE $Au = f$, a parametrized deformation $A_t u_t = f_t$ of this equation, $0 \le t \le 1$, $A_1 = A$, $u_1 = u$ (which has to be found), $f_1 = f$ (assumed to be given) and suppose that a solution of $A_0 u_0 = f_0$ is given or at least exists. Then one tries to construct a curve $\{u_t\}_{0\le t\le 1}$ of solutions, by means of linearization and the implicit function theorem. For this, one tries to show that $\mathcal{S} = \{t \in [0,1] | u_t \text{ exists}\}$ is open and closed in [0,1]. More general, we consider infinite-dimensional Banach manifolds $\mathcal{M}_1$, $\mathcal{M}_2$ and a non-linear map

$$G : \mathcal{M}_1 \longrightarrow \mathcal{M}_2.$$

Let $m_2 \in \mathcal{M}_2$ be given (variation of m_2 within a class admitted), one has to find $m_1 \in \mathcal{M}_1$ such that $G(m_1) = m_2$. To make this more clear, we present two examples.

1) Harmonic maps

Let (M^n, g), $(M^{n'}, h)$ be open with (I) and (B_k), $k \geq r > \frac{n}{2} + 1$. Set $\mathcal{M}_1 = \Omega^r(M, N)$, $\mathcal{M}_2 = T\Omega^r(M, N)$, $G : \mathcal{M}_1 = \Omega^r K(M, N) \longrightarrow T\Omega^r(M, N) = \mathcal{M}_2$ is the section $u \underset{G}{\longrightarrow} \tau(u)$ = tension of $u = \mathrm{tr}_g \nabla du$, m_2 = zero's of G, i.e. $\tau(u) = 0$. This is a non-linear PDE for u.

2) Teichmüller theory

(M^2, g) open with (I), (B_∞), $\mathcal{M}_1 = \Omega^{0,2,r}(M^2, g) \equiv \Omega^r$, $\mathcal{M}_2 = \Omega^{r-2}$, $G(u) = \Delta_g u + K_g + e^u$, $m_2 = 0$, i.e. one has to solve the equation $\Delta_g u + K_g + e^u = 0$ which is equivalent to $K_{e^u \cdot g} = -1$. Assume in the sequel that $\mathcal{M}_2$ locally has a vector space structure. This is satisfied e.g. if it is a vector bundle. Then we consider w.l.o.g. the equation $G(m) = 0$. If the initial task is $G(m) = m_2$, then consider G_{new}, $G_{new}(m) = G(m) - m_2$. The parametrized picture then looks as follows. Given $G_t(\cdot) \equiv F(t, \cdot)$, $0 \leq t \leq 1$ such that

a) there exists a solution m_0, $G_0(m_0) = F(0, m_0) = 0$ and

b) $G_1(\cdot) = F(1, \cdot) = G(\cdot)$, the given G.

Examples 9.1 **1)** Harmonic maps, e.g. $M = N$, $h \in \mathrm{comp}^{2,r}(g)$, i.e. $m_0 = u_0 = \mathrm{id} : (M, g) \longrightarrow (M, g)$, $\{g_t\}_{0 \leq t \leq 1}$, $g_0 = g$, $g_1 = h$, $G_t(u) = F(t, u) = \mathrm{tr}_{g_0}(\nabla_{g_t} du) = \tau_t(u)$; or $v : (M, g) \longrightarrow (N, h_0)$ harmonic, $\{h_t\}_{0 \leq t \leq 1}$ in $\mathrm{comp}^{2,r}(h)$, $h_1 = h$, $\mathcal{M}_1 = \mathrm{comp}^r(v) \subset \Omega^r(M, N)$, $G_t(u) = F(t, u) = \mathrm{tr}_g(\nabla h_t u) = \tau_t(u)$.

2) Teichmüller theory, (M^2, g_0), $\mathrm{comp}^{2,r}(g_0)$, $K_{g_0} \equiv -1$, $g \in \mathrm{comp}^{2,r}(g_0)$. $g_t = (1-t)g_0 + tg$, $G_t(u) = F(t, u) = \Delta_{g_t} u + K_{g_t} + e^u \overset{!}{=} 0$, $F(0, u_0 = 0) = 0 = \Delta_{g_0} 0 + K_{g_0} + e^0 = 0$.

We will now establish a theorem which is in fact convenient in many cases.

Assume in the sequel that $\mathcal{M}_1$ is Riemannian with $\exp_m : T_m\mathcal{M}_1 \longrightarrow \mathcal{M}_1$ and with $r_{\mathrm{inj}} > 0$. Hence for $\mathrm{dist}(m, m')$ sufficiently small there exists a unique $X \in T_m\mathcal{M}_1$ such that $\exp_m X = m'$.

Set now

$$\mathcal{S} = \{t \in [0,1] | \exists m_t \text{ such that } F(t, m_t) = 0\}.$$

Theorem 9.2 *Assume the following hypothesis.*

a) *There exists a $C > 0$ such that $F(t_0, m_{t_0}) = 0$ implies*

$$|F_m(t_0, m_{t_0})^{-1}(F_m(t_0, m_{t_0}) - F_m(t, m_{t_0}))|_{op} \leq C \cdot |t - t_0|,$$

C independent of $t_0 \in \mathcal{S}$.

b) *There exist $\delta_1 > 0$, $\delta_2 > 0$, such that for $|X| < \delta_1$, $|t - t_0| < \delta_2$ $\left|F_m(t_0, m_{t_0})^{-1}(F_m(t, m_{t_0})X - P_{-X}F(t, \exp X))\right| \leq \frac{1}{8}$, δ_i independent of $t_0 \in \mathcal{S}$.*

c) *There exists $\delta_3 > 0$, $\delta_3 \leq \min\{\frac{1}{4}\frac{1}{C}, \delta_1\}$, such that for $|t - t_0| < \delta_3$ $\left|F_m(t_0, m_{t_0})^{-1}F(t, m_{t_0})\right| \leq \frac{1}{2}\delta_1$, δ_3 independent of $t_0 \in \mathcal{S}$.*

Then there exists $\delta_4 > 0$, such that $[t_0 - \delta_4, t_0 + \delta_4] \cap [0,1] \subseteq \mathcal{S}$, δ_4 independent of t_0.

Corollary 9.3 *Assume the hypotheses of the theorem. Then there exists a solution $m_1 \in \mathcal{M}_1$ such that $G(m_1) = F(1, m_1) = 0$.* □

Proof of the theorem. Denote $\frac{\partial F}{\partial m} = F_m$ and by P_{-X} the parallel transport from m to m_0 along the geodesic $\exp_m(-PX)$, $X \in T_{m_0}\mathcal{M}_1$. Set $g(t, X) =$

$F_m(t_0, m_{t_0})X - P_{-X}F(t, \exp X)$. $F(t, \exp X) = 0 \iff P_{-X}F(t, \exp X) = 0 \iff X = F_m(t_0, m_{t_0})^{-1}g(t, x) \iff T_tX = X$, where $T_tX = F_m(t, m_0)^{-1}g(t, X)$. Hence $F(t, \exp X) = 0$ if and only if X is a fixed point of T_t. Set $M_{t_0,\delta_1} = \{X \in T_{m_0}\mathcal{M}_1 | |X| \leq \delta_1\}$, δ_1 from b). We obtain from a) and b) that

$$\begin{aligned} |T_tX - T_tY| &= |F_m(t_0, m_{t_0})^{-1}(g(t, X) - g(t, Y))| \\ &= |F_m(t_0, m_{t_0})^{-1}(F_m(t_0, m_{t_0})X - F_m(t_0, m_{t_0})Y) \\ &\quad - P_{-X}F(t, \exp X) - P_YF(t, \exp y))| \\ &\leq \left(C \cdot |t - t_0| + \frac{1}{4}\right)|X - Y| \end{aligned} \tag{9.1}$$

Moreover

$$\begin{aligned} |T_tX| = |T_tX - 0| &\leq |T_tX - T_t0| + |T_t0 - T_{t_0}0| \\ &\leq \left(C \cdot |t - t_0| + \frac{1}{4}\right)|X| + |F_m(t_0, m_{t_0})^{-1}F(t, m_{t_0})| \\ &\leq \left(C \cdot |t - t_0| + \frac{1}{4}\right)|X| + \frac{1}{2}\delta_1. \end{aligned} \tag{9.2}$$

We obtain from (9.2) that for $C \cdot |t - t_0| + \frac{1}{4} \leq \frac{1}{2}$, T_t maps M_{t_0,δ_1} into itself and from (9.1) that T_t is contracting. Choose δ_4 so that $\delta_4 \leq \delta_2, \delta_3$ and $C \cdot \delta_4 + \frac{1}{4} \leq \frac{1}{2}$. Then for $t \in [t_0 - \delta_4, t_0 + \delta_4] \cap [0, 1]$, T_t has a unique fixed point in M_{t_0,δ_1}. □

Perhaps the assumptions of theorem 9.2 look artificial and even strange. We will present in the next section a straightforward application of it.

Remark 9.4 It is easy to see that the reduction to the Banach fixed point theorem is equivalent to the proof of the closedness of the set $\mathcal{S}$. The required a-priori-estimates are the same. □

10 Teichmüller theory

The next example of non-linear PDEs on open manifolds is the equation of Teichmüller theory. We treat it by help of the continuity method, applying all the analytic tools developed until now and other deep results of Yau. The definition and the study of Teichmüller spaces for closed or compact surfaces with boundary or surfaces with punctures was a long time a frequent topic in geometry and analysis. There are many approaches. First we must mention Ahlfors in [3] and Bers in [13] which rely heavily on the theory of quasiconformal maps. A more geometric fibre bundle approach has been established by Earle and Eells in [37], [38]. Finally, an approach which relies on methods of differential geometry and global analysis has been presented by Fischer and Tromba in [57], [118]. What they are doing is in a certain sense canonical and at the same time very beautiful. Let M^2 be a closed oriented surface of genus $p > 1$, $\mathcal{M}$ its set of Riemannian metrics, $\mathcal{M}^r$ its Sobolev completion, $\mathcal{M}^r_{-1}$ the submanifold of metrics g with scalar curvature $K(g) \equiv -1$, $\mathcal{P}^r$ the completed space of positive conformal factors, $\mathcal{A}^r$ the completed space of almost complex structures, $\mathcal{D}^{r+1}$ as above the completed diffeomorphism group, $\mathcal{D}^{r+1}_0 \subset \mathcal{D}^{r+1}$ the component of the identity. Then Fischer and Tromba define as Teichmüller space

$$\mathcal{T}^r(M^2) := \mathcal{A}^r / \mathcal{D}^{r+1}_0$$

and prove $\mathcal{D}_0^{r+1}$ -equivariant isomorphisms

$$\mathcal{M}^r/\mathcal{P}^r \cong \mathcal{A}^r \quad \text{and} \quad \mathcal{M}^r_{-1} \cong \mathcal{M}^r/\mathcal{P}^r.$$

Hence there are three models for the Teichmüller space:

$$\mathcal{T}^r = \mathcal{A}^r/\mathcal{D}_0^{r+1} \cong (\mathcal{M}^r/\mathcal{P}^r)/\mathcal{D}_0^{r+1} \cong \mathcal{M}^r_{-1}/\mathcal{D}_0^{r+1}.$$

The isomorphism $\mathcal{M}^r_{-1} \cong \mathcal{M}^r/\mathcal{P}^r$ is known as Poincaré's theorem.

We study Teichmüller spaces for certain open oriented surfaces M^2 of infinite genus. At the beginning it is totally unclear how to define completed spaces $\mathcal{M}^r$, $\mathcal{M}^r_{-1}$, $\mathcal{T}^r$, $\mathcal{A}^r$, $\mathcal{D}^{r+1}$. A second difficult obstruction is the fact that the used results, e.g. the properness of the $\mathcal{D}^{r+1}$-action and the used theorems of elliptic theory, are totally wrong for open manifolds.

Nevertheless, the general uniformization theorem tells us that there are many complex = almost complex structures and metrics of curvature -1, i.e. there should be a Teichmüller space which "counts" these structures. The main question is how to count them, how to define a Teichmüller space? In this section, we present a canonical and natural approach but under certain restrictions. We restrict ourselves to open oriented surfaces of the following kind. Start with a closed oriented surface and form the connected sum with a finite number of half ladders $\sharp_1^\infty T^2$, where T^2 is the 2-torus. Now we allow the repeated addition of a finite number of half ladders in such a manner that there arises a surface with at most countably many ends. Surfaces of the admitted topological type can be built up by Y-pieces which guarantees the existence of a metric g_0 satisfying $K(g_0) \equiv -1$ and $r_{inj}(g_0) > 0$. We exclude metric cusps, but we admit additional metric trumpets, i.e. topological punctures. We consider the space $\mathcal{M}^r(I, B_k), r \leq k$. According to section 3, $\mathcal{M}^r(I, B_k)$ has a representation as topological sum

$$\mathcal{M}^r(I, B_k) = \sum_{i \in I} \text{comp}(g_i)$$

and for $k \geq r > \frac{n}{2} + 1$ each component $\text{comp}(g_i)$ is a Hilbert manifold. In section 3 we defined the completed space of positive conformal factors,

$$\mathcal{P}^r(g) = \sum_i \text{comp}(e^{u_i}),$$

and $\text{comp}(1) \subset \mathcal{P}^r(g)$ is an invariant of $\text{comp}(g)$. Return now to M^2 of the above topological type. Denote by $\text{comp}(g)_{-1} \subset \text{comp}(g)$ the subspace of all metrics $g' \in \text{comp}(g)$ such that $K(g') \equiv -1$. Then we would define

$$\mathcal{T}^r(\text{comp}(g)) := \text{comp}(g)_{-1}/\mathcal{D}_0^{r+1}$$

and expect

$$\text{comp}(g)_{-1} \cong \text{comp}(g)/\text{comp}(1). \tag{10.1}$$

But there are simple examples of components $\text{comp}(g)$ with $\text{comp}(g)_{-1} = \emptyset$. Moreover, we don't see any chance to prove (10.1) for arbitrary g. To have $\text{comp}(g)_{-1} \neq \emptyset$, we start with a metric $g_0 \in \mathcal{M}(I, B_\infty)$ with $K(g_0) \equiv -1$. To g_0 we attach an almost complex structure $J_0 = J(g_0) := g_0^{-1}\mu(g_0)$, where $\mu(g_0)$ is the volume form. Then we can summarize our main results in the following

Theorem 10.1 *Suppose* $g_0 \in \mathcal{M}(I, B_\infty), K(g_0) \equiv -1,\ inf\,\sigma_e(\Delta g_0) > 0, r > 3$. *Then* $\mathrm{comp}(g_0)_{-1} \subset \mathrm{comp}(g_0) \subset \mathcal{M}(I, B_\infty)$ *is a submanifold. There is a* $\mathcal{D}_0^{r+1}(g_0)$ *-equivariant isomorphism*

$$\mathrm{comp}(g_0)_{-1} \cong \mathrm{comp}(g_0)/\mathrm{comp}(1) \cong \mathrm{comp}(J_0). \tag{10.2}$$

If we define the Teichmüller space $\mathcal{T}^r(\mathrm{comp}(g_0))$ *of* $\mathrm{comp}(g_0)$ *as*

$$\mathcal{T}^r(\mathrm{comp}(g_0)) := \mathrm{comp}(J_0)/\mathcal{D}_0^{r+1}$$

then

$$\mathcal{T}^r(\mathrm{comp}(g_0)) \cong \mathrm{comp}(g_0)_{-1}/\mathcal{D}_0^{r+1} \cong (\mathrm{comp}(g_0)/\mathrm{comp}(1))/\mathcal{D}_0^{r+1}.$$

The first isomorphism in (10.2) is Poincaré's theorem for the open case. Its proof occupies the major part of this section.

Next we indicate the structure of $\mathcal{P}_m^r(g)$.

Theorem 10.2 *Under multiplication* $\mathcal{P}_m^r(g)$ *is a Hilbert-Lie group,* $r \leq m$.

Sketch of proof. It follows immediately from the definition, the product and quotient rule and the module structure theorem that $\mathcal{P}_m^r(g)$ is a group. $\mathfrak{B} = \{U_\delta\}_\delta > 0$,

$$U_\delta = \{\varphi \in \mathcal{P}_m^r(g) |\, |\varphi - 1|_{g,r} < \delta\},$$

is a filter basis centred at $1 \in \mathcal{P}_m^r(g)$ that satisfies all axioms for the neighbourhood filter of 1 of a topological group. Hence $\mathcal{P}_m^r(g)$ is a topological group (cf. [31]). Finally, U_δ is homeomorphic to an open ball in $\Omega^{2,r}(M)$ for $\delta > 0$ sufficiently small and has the structure of a local real Lie group. Hence $\mathcal{P}_m^r(g)$ is a Hilbert-Lie group. □

Assume as always $k \geq r > \frac{n}{2} + 1$, $g \in \mathcal{M}(I, B_k)$ and consider $\mathrm{comp}_{k+2}^r(1) \subset \mathcal{P}_{k+2}^r(g), \mathrm{comp}(g) \subset \mathcal{M}^r(I, B_k)$.

Proposition 10.3 **a)** *There is a well defined action*

$$\begin{aligned}&\mathrm{comp}_{k+2}^r(1) \times \mathrm{comp}(g) \to \mathrm{comp}(g)\\ &(\varphi', g') \to \varphi' \cdot g'.\end{aligned}$$

b) *The action is smooth, free and proper.* □

Corollary 10.4 **a)** *The orbits* $\mathrm{comp}_{k+2}^r(1) \cdot g' \subset \mathrm{comp}(g)$ *are smooth submanifolds of* $\mathrm{comp}(g)$.

b) *The quotient space* $\mathrm{comp}(g)/\mathrm{comp}_{k+2}^r(1)$ *is a smooth manifold.*

c) *The projection* $\pi : \mathrm{comp}(g) \to \mathrm{comp}(g)/\mathrm{comp}_{k+2}^r(1)$ *is a smooth submersion and has the structure of a principal fibre bundle.* □

$\mathrm{comp}(g)$ has as tangent space at $g' \in \mathrm{comp}(g)$, $T_{g'}\mathrm{comp}(g) = \Omega^r(S^2T^*, g') \cong \Omega^r(S^2T^*, g)$, where S^2T^* are the symmetric 2-fold covariant tensors. There is an L_2-orthogonal splitting

$$T_{g'}\mathrm{comp}(g) = \Omega^{r,c}(S^2T^*, g') \oplus \Omega^{r,T}(S^2T^*, g'), \tag{10.3}$$

where

$$\Omega^{r,c}(S^2T^*, g') = \{h \in \Omega^r(S^2T^*, g') | h(x) = p(x) \cdot g'(x), p \in \Omega^r(M, g'\}$$

and

$$\Omega^{r,T}(S^2T^*, g') = \{h \in \Omega^r(S^2T^*, g') \mid \operatorname{tr}_{g'} h = 0\}.$$

The decomposition (10.3) is given by (see [118] for further details)

$$h = \frac{1}{n}(\operatorname{tr}_{g'} h) \cdot g' + \left(h - \frac{1}{n}(\operatorname{tr}_{g'} h)g'\right).$$

Hence we obtain for $[g'] = \operatorname{comp}^r_{k+2}(1) \cdot g'$

$$T_{g''}(\operatorname{comp}^r_{k+2}(1) \cdot g') = \Omega^{r,c}(S^2T^*, g'')$$

and

$$T_{[g']}\operatorname{comp}(g)/\operatorname{comp}^r_{k+2}(1) = \Omega^{r,T}(S^2T^*, g').$$

Now we study the space of hyperbolic metrics for $n = 2$. We will show that for certain classes of open surfaces, a suitable metric g_0 and the space $\operatorname{comp}(g_0)_{-1} \subset \operatorname{comp}(g_0)$ of constant scalar curvature -1 holds

$$\operatorname{comp}(g_0)_{-1} \cong \operatorname{comp}(g_0)/\operatorname{comp}(1)$$

where these spaces are manifolds and $\mathcal{D}^r_0(g_0)$-equivariantly diffeomorphic to a certain component in the space of almost complex structures. The quotient space $\operatorname{comp}_{-1}(g_0)/\mathcal{D}^r_0(g_0)$ will be one of our models for the Teichmüller space.

We consider open surfaces M^2. Each such surface has ends. We admit punctures as ends. If each end is isolated, then M^2 has a finite number of ends, each of them is given by an infinite half ladder $\cong \mathop{\#}\limits_{n=1}^{\infty} T^2$, where T^2 is the 2-Torus or it is given by a puncture. If M^2 has an infinite number of ends then there exists at least one non-isolated end, i.e. an end that has no neighbourhood which is not a neighbourhood of another end. This occurs e.g. if we have repeated branchings of half ladders. In any case, such a surface can be built up by Y-pieces or so-called trumpets which we explain now. We follow the presentation given in [18].

Lemma 10.5 *Let a, b, c be arbitrary positive real numbers. There exists a right angled geodesic hexagon in the hyperbolic plane with pairwise non-adjacent sides of length a, b, c.*

□

Next we paste two copies of such a hexagon together along the remaining three sides to obtain a hyperbolic surface Y with three closed boundary geodesics of length $2a, 2b, 2c$. They determine Y up to isometry (Theorem 3.17 of [18]).

Two different Y-pieces can be glued along their boundary geodesics if they have the same length. The same holds for two "legs" of same boundary length of one Y-piece. It is a deep result of hyperbolic geometry that one obtains in this way smooth hyperbolic surfaces. Moreover, we can perform gluing with an additional twisting (cf. [18]). But here we consider gluings without twisting, at least for our starting metric g_0. As a well known matter of fact, any topologically given open surface of the above kind can be built up by Y-pieces and trumpets and we obtain in this way a hyperbolically metrized surface (M^2, g_0). The lengths of all closed boundary geodesics can be chosen in such a manner (and $\geq a > 0$) that $r_{inj}(M^2, g_0) > 0$, (cf. [38]) i.e. $g_0 \in \mathcal{M}(I, B_\infty)$.

Given an open surface M^2 of the above type, i.e. M^2 is the connected sum of a closed surface with an infinite number of half ladders with possibly infinitely many punctures, fix in this case a hyperbolic metric $g_0 \in \mathcal{M}(I, B_\infty)$. Later we must impose that these lengths must grow suitably. Consider $\mathcal{P}_\infty(g_0) = \bigcap_m \mathcal{P}_m(g_0)$, $\mathcal{P}^r_\infty(g_0)$ defined by the induced uniform structure. It is a very simple fact that $\operatorname{comp}^r_k(1, g_0) \subset \mathcal{P}^r_k(g_0)$ and $\operatorname{comp}^r_\infty(1, g_0) \subset \mathcal{P}^r_\infty(g_0)$ coincide, $k \geq 1$. We fix $r > 3$ and write $\operatorname{comp}(1) = \operatorname{comp}^r(1, g_0)$. Consider $\operatorname{comp}(g_0) \subset \mathcal{M}^r(I, B_\infty)$. As we already know, $\operatorname{comp}(1)$ acts on $\operatorname{comp}(g_0)$ and $\operatorname{comp}(g_0)/\operatorname{comp}(1)$ is a Hilbert manifold. Let $\operatorname{comp}(g_0)_{-1} \subset \operatorname{comp}(g_0)$ be the subspace of all metrics $g \in \operatorname{comp}(g_0)$ such that the scalar curvature $K(g)$ equals -1. Since we assume $r > 3 = \frac{2}{2} + 2$, g is at least of class C^2 and $K(g)$ is well defined. Usually, $K(g)$ denotes the sectional curvature. We use it for scalar curvature (which is twice the sectional curvature) because of notational convenience.

We wish to show that $\operatorname{comp}(g_0)_{-1} \subset \operatorname{comp}(g_0) \subset \mathcal{M}^r(I, B_\infty)$ is a smooth submanifold of $\operatorname{comp}(g_0)$ which is diffeomorphic to $\operatorname{comp}(g_0)/\operatorname{comp}(1)$. This is a rather deep fact which requires a series of preliminaries and is valid only under an additional spectral assumption. Let $g \in \operatorname{comp}(g_0)$. Then Δ_g maps $\Omega^r = \Omega^{2,r}(M, \nabla^{g_0}, g_0)$ into $\Omega^{r-2} \subset L_2(M, g_0)$.

Lemma 10.6 $\Delta_g + 1$ *is surjective.*

Proof Consider $\Delta_g + 1$ with domain $\Omega^r \subset \Omega^{r-2}$. Then the closure of $(\Omega^r, |\ \ |_{r-2})$ with respect to $|\cdot|_{r-2} + |(\Delta_g + 1)\cdot|_{r-2}$ is just Ω^r, i.e. $\Delta_g + 1$ is a closed operator in the Hilbert space Ω^{r-2}. Moreover, $|(\Delta_g + 1)\varphi|_{r-2} \geq c \cdot |\varphi|_{r-2}, c = 1, \varphi \in \Omega^r$. Hence $(\Delta_g + 1)\varphi_i \to \psi$ gives φ_i Cauchy and $\varphi_i \to \varphi$ in Ω^{r-2}. $\Delta_g + 1$ is closed, hence $(\Delta_g + 1)\varphi = \psi, im(\Delta_g + 1)$ closed. Finally, the orthogonal complement of $im(\Delta_g + 1)$ in Ω^{r-2} is $\{0\}$ since the adjoint (in Ω^{r-2}) operator to $\Delta_g + 1$ has no kernel. □

Let $h \in T_g \operatorname{comp}(g_0) = \Omega^r(S^2T^*, g)$. For h the divergence $\delta_g h$ is defined by $(\delta_g h)_j = \nabla^k h_{jk} = g^{ik}\nabla^g_i h_{jk}$. For $\omega = \omega_i dx^i$ a 1-form and $X_\omega = \omega^i \frac{\partial}{\partial x^i}$ the corresponding vector field, the divergence δ_w is defined by $\delta_g\omega := \delta_g X_\omega = \frac{1}{\sqrt{g}}\frac{\partial}{\partial x^i}(\omega^i\sqrt{g})$. Hence for $h \in \Omega^r(S^2T^*, g)$ the expression $\delta_g\delta_g h$ is well defined. As we already mentioned, for $r > 3 = \frac{2}{2} + 2, g \in \operatorname{comp}(g_0)$ is at least of class C^2 and the scalar curvature $K(g)$ is well defined.

Lemma 10.7 $K(g) - (-1) = K(g) - K(g_0) \in \Omega^{r-2}$.

This follows immediately from the topology in $\operatorname{comp}(g_0)$ and the definition of $K(g)$. □

Consider the C^∞-map

$$\psi : \operatorname{comp}(g_0) \to \Omega^{r-2}(M, g_0), \quad g \to K(g) - (-1).$$

Then $\operatorname{comp}(g_0)_{-1} = \psi^{-1}(0)$.

Theorem 10.8 $\operatorname{comp}(g_0)_{-1} \subset \operatorname{comp}(g_0)$ *is a smooth submanifold.*

Proof It suffices to show, that 0 is a regular value for ψ, i.e. if $K(g) = -1$ for some

g then $D\psi|_g : T_g \,\mathrm{comp}(g_0) \to \Omega^{r-2}(M, g_0)$ is surjective. Hence we have to calculate $D\psi|_g(h), h \in T_g \,\mathrm{comp}(g_0) = \Omega^r(S^2T^*, g)$. This has been done in [118],

$$D\psi|_g(h) = \Delta_g(tr_g h) + \delta_g\delta_g h + \frac{1}{2} tr_g h. \tag{10.4}$$

$D\psi|_g$ is already surjective if the restriction of $D\psi$ to h of the kind $h = \lambda \cdot g, \lambda \in \Omega^r(M)$, is surjective. Then (10.4) becomes

$$D\psi|_g(\lambda \cdot g) = \Delta_g\lambda + \lambda = (\Delta_g + 1)\lambda,$$

but $\Delta_g + 1$ is surjective according to 10.6. □

Next we prepare the proof of Poincaré's theorem which, roughly spoken, asserts $\mathrm{comp}(g_0)_{-1} \cong \mathrm{comp}(g_0)/\mathrm{comp}(1)$. Denote by $\sigma_e(\Delta)$ the essential spectrum of Δ. Here we omit the bar in the unique self-adjoint extension $\bar{\Delta}$ which equals the closure. By section 4, we have

Proposition 10.9 *$\sigma_e(\Delta_{g_0})$ is an invariant of* $\mathrm{comp}(g_0)$*, i.e. for $g \in \mathrm{comp}(g_0)$,*

$$\sigma_e(\Delta_g) = \sigma_e(\Delta_{g_0}).$$

□

Lemma 10.10 *Assume* $\inf \sigma_e(\Delta_{g_0}) > 0$*. Then* $\inf \sigma(\Delta_g) > 0$ *for all $g \in \mathrm{comp}(g_0)$, where σ denotes the spectrum.*

Proof According to 10.9, $\inf \sigma_e(\Delta_{g_0}) = \inf \sigma_e(\Delta_g)$. From $g \in \mathcal{M}(I, B_\infty), g \in \mathrm{comp}(g_0) \subset \mathcal{M}^r(I, B_\infty)$, $r > 3$ follows that g satisfies (I) and (B_0) which implies $\mathrm{vol}\,(M^2, g) = \infty$. Hence $\lambda = 0$ cannot be an eigenvalue. All other spectral values between 0 and $\inf \sigma_e(\Delta_g)$ belong to the purely discrete point spectrum $\sigma_{pd}(\Delta_g)$, i.e. $\inf \sigma(\Delta_g) > 0$. □

Now we state our main theorem, which is our version of Poincaré's lemma.

Theorem 10.11 *Assume (M^2, g_0) with g_0 smooth, $K(g_0) \equiv -1$, $r_{inj}(M^2, g_0) > 0$, $inf\sigma_e(\Delta_{g_0}) > 0$. Let $g \in \mathrm{comp}(g_0) \subset \mathcal{M}^r(I, B_\infty), r > 3$. Then there exists a unique $\varrho \in \mathrm{comp}(1) \subset \mathcal{P}^r_\infty(g_0)$ such that $K(\varrho \cdot g) \equiv -1$.*

Proof Let $\varrho = e^u$. For the existence we have to solve the PDE

$$\Delta_g u + K(g) + e^u = 0. \tag{10.5}$$

We seek for a solution $u \in \Omega^r(M, g_0)$. $u \in \Omega^r(M, g_0), r > 3$ imply $e^u - 1 \in \Omega^r$ as we will see below. (10.5) has a solution according to the general uniformization theorem. But this theorem does not provide $u \in \Omega^r$. Therefore we have to sharpen our considerations. The existence will be established by the implicit function theorem and a version of the continuity method. Consider $g_t = (1-t)g_0 + tg = g_0 + t(g - g_0) = g_0 + th \in \mathrm{comp}(g_0)$ and the map

$$F : [0,1] \times \Omega^r \to \Omega^{r-2},$$
$$(t, u) \to F(t, u) = \Delta_{g_t} u + K(g_t) + e^u = \Delta_{g_t} u + (K(g_t) - (-1)) + e^u - 1.$$

We want to show that there exists a unique $u_1 \in \Omega^r(M, g_0)$ such that $F(1, u_1) = 0$. For this we consider the set

$$\mathcal{S} = \{t \in [0,1] | \quad \text{There exists} \quad u_t \in \Omega^r \quad \text{such that} \quad F(t, u_t) = 0\}$$

and we want to show $\mathcal{S} = [0, 1]$. We start with $\mathcal{S} \neq \emptyset$. For $t = 0, g_t = g_0, K(g_0) = -1$ and $u_0 \equiv 0$ satisfies (10.5). Moreover,

$$F_u(0,0) = D_2F|_{(0,0)} = \Delta_{g_0} + 1$$

is bijective between Ω^r and Ω^{r-2}, as we have already seen. Hence there exist $\delta > 0, \varepsilon > 0$ such that for $t \in]0, \delta[$ there exists a unique $u_t \in U_\varepsilon(0) \subset \Omega^r$ with

$$F(t, u_t) = 0.$$

By the same consideration we can show that $\mathcal{S}$ is open in $[0, 1]$. To conclude $\mathcal{S} = [0, 1]$ we need to show that $\mathcal{S}$ is closed. The canonical procedure to prove this would be to prove

$$(u_{t_\nu})_\nu \text{ is a Cauchy sequence in } \Omega^r, u_{t_\nu} \to u_{t_0}, \tag{10.6}$$

$$\Delta_{g_{t_0}} u_{t_0} + K(g_{t_0}) \in e^{u_{t_0}} = 0 \tag{10.7}$$

for any sequence $t_1 < t_2 \cdots \in S, t_\nu \longrightarrow t_0$. □

We prefer a slightly other version of this argument.

Proposition 10.12 *There exists a constant* $\delta > 0$ *independent of* t_0 *such that* $t_0 \in \mathcal{S}$ *implies* $]t - \delta_0, t_0 + \delta[\cap [0, 1] \subset \mathcal{S}$.

We will see later that the proof of 10.12 is equivalent to that of (10.6) and (10.7). The proof of 10.12 is based on careful estimates in the implicit function theorem to which we now turn our attention. Roughly speaking, the proof goes as follows.

Let $t_0 \in \mathcal{S}, u_{t_0} \in \Omega^r$,

$$F(t_0, u_{t_0}) = \Delta_{g_{t_0}} u_{t_0} + K(_{g_{t_0}}) + e^{u_{t_0}} = 0.$$

Set $g(t, u) := F_u(t_0, u_{t_0})u - F(t, u)$. Then $F(t, u) = 0$ is equivalent to

$$u = F_u(t_0, u_{t_0})^{-1} g(t, u). \tag{10.8}$$

If we define $T_t u := F_u(t_0, u_{t_0})^{-1} g(t, u)$, then we are done if we can find for any $t_0 \in \mathcal{S}$ a complete metric subspace $M_{t_0,\delta_1} \subset \Omega^r(M, g_0)$ such that

$$T_t : M_{t_0,\delta_1} \to M_{t_0,\delta_1} \tag{10.9}$$

and

$$T_t \quad \text{is contracting} \tag{10.10}$$

for all $t \in]t_0 - \delta, t_0 + \delta[\cap [0, 1], \delta$ independent of t_0. Indeed, in this case T_t would have a unique fixed point u_t solving $F(t, u_t) = 0$.

We now prepare the construction of M_{t_0,δ_1} and the proof of (10.9), (10.10) by a series of estimates. First we apply the mean value theorem. From $g_u(t, v) = F_u(t_0, u_{t_0}) - F_u(t, v)$ follows

$$|g(t, u) - g(t, v)|_{r-2} \leq \sup_{0<\vartheta<1} |g_u(t, v + \vartheta(u - v))|_{r-2} \cdot |u - v|_r,$$

$$|T_t u - T_t v|_r \leq |(\Delta_{g_{t_0}} + e^{u_{t_0}})^{-1}|_{r-2,r} \cdot$$

$$\cdot \sup_{0<\vartheta<1} |(\Delta_{g_{t_0}} - \Delta_{g_t}) + ((e^{u_{t_0}} - e^{v+\vartheta(u-v)})\cdot)|_{r,r-2} \cdot |u-v|_r,$$

where $| \quad |_{i,j}$ denotes the operator norm $\Omega^i(M, g_0) \to \Omega^j(M, g_0)$. We estimate

$$|(\Delta_{g_{t_0}} + (e^{u_{t_0}}\cdot))^{-1}|_{r-2,r} \cdot |\Delta_{g_{t_0}} - \Delta_{g_t}|_{r,r-2} \tag{10.11}$$

and

$$|(\Delta_{g_{t_0}} + (e^{u_{t_0}}\cdot))^{-1}(e^{u_{t_0}}\cdot)|_{r-2,r} \cdot |(1 - e^{v-u_{t_0}+\vartheta(u-u_{t_0}-(v-u_{t_0}))}) \cdot |_{r,r-2} \tag{10.12}$$

and start with (10.11). In the sequel, the same letters for constants in different inequalities can denote different constants. The key property in all following considerations is the Lipschitz continuity of $|\Delta_{g_t}|_{i,j}$.

Lemma 10.13 *Assume g_0, g, t, t_0, r as above. Then there exists a constant $C = C(g_0, r, |g-g_0|_{g_0,r}) > 0$ such that*

$$|\Delta_{g_{t_0}} - \Delta_{g_t}|_{r,r-2} \leq C \cdot |t_0 - t|. \tag{10.13}$$

The proof is really formidable and we refer to [48] for it.

Now we continue to estimate (10.11) and have to estimate

$$|(\Delta_{g_{t_0}} + (e^{u_{t_0}}\cdot))^{-1}|_{r-2,r}$$

First we recall that Δ_{g_t} is self-adjoint on $\Omega^2(M, \Delta_{g_t}, g_t) = \Omega^2(M, \Delta_{g_0}, g_0) \subset L_2(M) = \Omega^0(M)$. For $u \in \Omega^r, r > 3$, the operator $v \to e^u \cdot v$ is symmetric and bounded on L_2. Hence $\Delta_{g_t} + e^u$ is self-adjoint.

Lemma 10.14 *There exists a constant $c > 0$ such that* $\inf \sigma(\Delta_{g_t}) \geq c, 0 \leq t \leq 1$.

Proof Assume the converse. Then there exists a convergent sequence $t_i \to t^*$ in $[0,1]$ such that $\lambda_{min}(\Delta_{g_{t_i}}) \to 0$. Here $\lambda_{min}(\Delta_{g_{t_i}})$ is the minimal spectral value of $\Delta_{g_{t_i}}$. It is > 0 and either equal to $\inf \sigma_e(\Delta_{g_t})$ or an isolated eigenvalue of finite multiplicity. According to 10.13, $\Delta_{g_{t_i}} \to \Delta_{g_{t^*}}$ in the generalized sense of [81], IV, § 2.6. Then, according to [81], V, § 4, remark 4.9, $\lambda_{min}(\Delta_{g_{t_i}}) \to \lambda_{min}(\Delta_{g_{t^*}})$, i.e. necessary $\lambda_{min}(\Delta_{g_{t^*}}) = 0$, a contradiction. □

Corollary 10.15 *For arbitrary $t \in [0,1], u \in \Omega^r$*

$$\inf \sigma(\Delta_{g_t} + e^u) \geq c,$$

$$\Delta_{g_t} + e^u = \int_c^\infty \lambda dE_\lambda(t,u),$$

$$(\Delta_{g_t} + e^u)^{-1} = \int_c^\infty \lambda^{-1} dE_\lambda(t,u),$$

$(\Delta_{g_t} + e^u)^{-1}$ is a bounded operator on L_2 and, according to [81], p. 357, (5.17), the operator norm of $(\Delta_{g_t} + e^u)^{-1}$ is $\leq \frac{1}{c}$. □

We want to prove more and to estimate

$$((\Delta_{g_t} + e^u)^{-1}|_{r-2,r}. \tag{10.14}$$

First we have to assure that (10.14) makes sense.

Lemma 10.16 *For $u \in \Omega^r, r > 3$, the map $v \to e^u \cdot v$ is a bounded map $\Omega^i \to \Omega^i, i \leq r$, with*

$$|e^u|_{i,i} \leq C(u,i) \leq C(i) \cdot sup\, e^u \cdot |u|_r. \tag{10.15}$$

□

Corollary 10.17 *The Sobolev spaces based on the operators Δ_{g_t} and $\Delta_{g_t} + e^u$ are equivalent for $i \leq r$,*

$$\Omega^i(M^2), \Delta_{g_s}, g_s) \cong \Omega^i(M^2, \Delta_{g,t} + e^u), i \leq r. \tag{10.16}$$

□

Remark 10.18 The heart of the estimate for (10.14) consists in proving that the constants arising in (10.15), (10.16) can be chosen independently of t and u if u solves

$$F(t,u) \equiv \Delta_{g_t} u + K(g_t) + e^u = 0.$$

□

Consider $\Omega^{2,r} = \Omega^r \subset \Omega^{2,2} = \Omega^2 \subset \Omega^{2,0} = \Omega^0 = L_2, \Omega^{r-2} \subset L_2$ and assume r even.

Lemma 10.19 $\Delta_{g_t} + e^u : \Omega^2 \to \Omega^0 = L_2$ *induces a bijective morphism between $\Omega^r \subset \Omega^2$ and $\Omega^{r-2} \subset \Omega^0$.*

Proof Surely, $\Delta_{g_t} + e^u$ maps $\Omega^r \subset \Omega^2$ into $\Omega^{r-2} \subset \Omega^0 = L_2$. This map is injective according to 10.14. It is surjective: Let $v \in \Omega^{r-2} \subset \Omega^0$. Then $(\Delta_{g_t} + e^u)^{-1} v \in \Omega^2$, $(\Delta_{g_t} + e^u)^i((\Delta_{g_t} + e^u)^{-1} v) = (\Delta + e^u)^{i-1} v$ is square integrable $i \leq \frac{r}{2}$. The assertion now follows from 10.17. □

Now we state our main

Proposition 10.20 *Assume $r > 3$ even. Then there exists a constant $C = C(g_0, g) > 0$, independent of t, such that*

$$|(\Delta_{g_t} + e^{u_t})^{-1}|_{r-2,r} \leq C \tag{10.17}$$

for any solution $u_t \in \Omega^r = \Omega^r(M, g_0)$ of $\Delta_{g_t} u_t + K(g_t) + e^{u_t} = 0$.

We omit the rather long and complicated proof which also uses Yau's general Schwarz lemma and refer to [43], [48] for details.

Corollary 10.21 *There exists a constant $C = C(g, g_0)$ such that*

$$|(\Delta_{g_{t_0}} + (e^{u_{t_0}} \cdot))^{-1}|_{r-2,r} \cdot |\Delta_{g_{t_0}} - \Delta_{g_t}|_{r,r-2} \leq C \cdot |t - t_0|. \tag{10.18}$$

□

The estimate of the first factor of (10.12) is already done,

$$|(\Delta_{g_{t_0}} + (e^{u_{t_0}} \cdot))^{-1} (e^{u_{t_0}} \cdot)|_{r-2,r} \leq |(\Delta_{g_{t_0}} + (e^{u_{t_0}} \cdot))^{-1}|_{r-2,r} \cdot |(e^{u_{t_0}} \cdot)|_{r-2,r-2}.$$

According to (10.17),

$$|(\Delta_{g_{t_0}} + (e^{u_{t_0}} \cdot))^{-1}|_{r-2,2} \leq C_1. \tag{10.19}$$

The proof of 10.20 contains the estimates $|\Delta^j u|_0 \leq D_j, 0 \leq j \leq \frac{r}{2}$,

$$|(e^{u_{t_0}} \cdot)|_{r-2,r-2} \leq C_2, \tag{10.20}$$

i.e.

$$|(\Delta_{g_{t_0}} + (e^{u_{t_0}} \cdot))^{-1}(e^{u_{t_0}} \cdot)|_{r-2,r} \leq C_3, \tag{10.21}$$

$C_3 = C_3(g, g_0)$ independent of t. The final estimate concerns

$$|(1 - e^{v-u_{t_0}+\vartheta(u-u_{t_0}-(v-u_{t_0}))}) \cdot |_{r,r-2}, \tag{10.22}$$

where as usual the point indicates that the corresponding expression acts by multiplication. We write

$$1 - e^{v-u_{t_0}+\vartheta(u-u_{t_0}-(v-u_{t_0}))} = -\sum_{i=1}^{\infty}[v - u_{t_0} + \vartheta(u - u_{t_0} - (v - u_{t_0})]^i/i!$$

As above, this series converges in Ω^r and for $|v - u_{t_0} + \vartheta(u - u_{t_0} - (v - u_{t_0}))|_r$ sufficiently small $|\sum_{i=1}^{\infty}[v - u_{t_0} + \vartheta(u - u_{t_0} - (v - u_{t_0}))]^i/i!|_r$ becomes arbitrarily small.

For any $f \in \Omega^r$, the operator norm of $(f\cdot) : \Omega^r \to \Omega^{r-2}, (f\cdot)w = f \cdot w$, can be estimated by $C(r) \cdot |f|_r$. This yields

Lemma 10.22 *For any* $\varepsilon_1 > 0$ *there exists* $\delta_1 > 0$ *such that*

$$|(1 - e^{v-u_{t_0}+\vartheta(u-u_{t_0}-(v-u_{t_0}))}) \cdot |_{r,r-2} \leq \varepsilon_1$$

for all u, v *with* $|u - u_{t_0}|_r, |v - u_{t_0}|_r \leq \delta_1$.

Proof Given $\varepsilon_1 > 0$, there exists δ_1' such that for $|v - u_{t_0} + \vartheta(u - u_{t_0} - (v - u_{t_0}))|_r < \delta_1'$

$$C(r) \cdot |\sum_{i=1}^{\infty}[v - u_{t_0} + \vartheta(u - u_{t_0} - (v - u_{t_0}))]^i/i!|_r \leq \varepsilon_1.$$

Set $\delta_1 = \delta_1'/4$. Then

$$\begin{aligned} &|v - u_{t_0} + \vartheta(u - u_{t_0} - (v - u_{t_0}))|_r < |v - u_{t_0}|_r + |u - u_{t_0}|_r + |v - u_{t_0}|_r = \\ &= |u - u_{t_0}|r + 2|v - u_{t_0}|_r < 2(|u - u_{t_0}|_r + |v - u_{t_0}|_r) \leq 4\delta_1 = \delta_1'. \end{aligned}$$

□

Corollary 10.23 *There exists* $\delta_1 > 0$ *such that* $|u - u_{t_0}|_r \leq \delta_1, |v - u_{t_0}|_r \leq \delta_1$ *implies*

$$|(\Delta + (e^{u_{t_0}} \cdot))^{-1}(e^{u_{t_0}} \cdot |_{r-2,r} \cdot |(1 - e^{v-u_{t_0}+\vartheta(u-u_{t_0}-(v-u_{t_0}))} \cdot |_{r,r-2} \leq \frac{1}{4}. \tag{10.23}$$

Proof Set in 10.22 $\varepsilon_1 = \frac{1}{4} \cdot \frac{1}{C_3}$, C_3 from (10.21). □

Corollary 10.24 *There exists* $\delta_1 > 0$ *such that for* $|u - u_{t_0}|_r \leq \delta_1, |v - u_{t_0}|_r \leq \delta_1$

$$|T_t u - T_t v|_r \leq (C \cdot |t - t_0| + \frac{1}{4})|u - v|_r, \tag{10.24}$$

where C *comes from lemma 10.13.*

Proof This follows immediately from (10.11), (10.12), (10.13), (10.23). □

If we choose $|t_0 - t|$ sufficiently small, then the map T_t would be contractive. But this does not make sense since until now we did not define a complete metric space on which T_t acts. This will be the next and last step in our approach. But we will use the inequality (10.24) in this step.

Proposition 10.25 *Suppose* $u_{t_0} \in \Omega^r, r > 3, \Delta_{g_{t_0}} u_{t_0} + K(g_{t_0}) + e^{u_{t_0}} = 0$. *There exist* $\delta, \delta_1 > 0$ *independent of* t_0 *such that* T_t *maps* $M_{t_0,\delta_1} = \{u \in \Omega^r | \, |u - u_{t_0}|_r \leq \delta_1\}$ *into itself for* $|t - t_0| \leq \delta$. *Moreover* T_t *is contracting.*

Proof We start estimating $T_t u - u_{t_0}$:

$$|T_t u - u_{t_0}|_r = |T_t u - T_{t_0} u_{t_0}|_r \leq |T_t u - T_t u_{t_0}|_r + |T_t u_{t_0} - T_{t_0} u_{t_0}|_r. \tag{10.25}$$

For $|u - u_{t_0}|_r \leq \delta_1, \delta_1$,

$$|T_t u - T_t u_{t_0}|_r \leq (C \cdot |t - t_0| + \frac{1}{4})|u - u_{t_0}|_r.$$

Hence for $|t - t_0| \leq \delta', |u - u_{t_0}|_r \leq \delta_1$

$$(C \cdot |t - t_0| + \frac{1}{4}) \leq \frac{1}{2}$$

and

$$|T_t u - T_t u_{t_0}|_r \leq \frac{1}{2}|u - u_{t_0}|_r \leq \frac{1}{2}\delta_1. \tag{10.26}$$

It remains to estimate $|T_t u_{t_0} - T_{t_0} u_{t_0}|_r$. But by an easy calculation

$$T_t u_{t_0} - T_{t_0} u_{t_0} = -(\Delta_{g_{t_0}} + (e^{u_{t_0}} \cdot))^{-1}((\Delta_{g_t} - \Delta_{g t_0})u_{t_0} + K(g_t) - K(g_{t_0})).$$

We are done if for $|t - t_0| \leq \delta''$

$$|(\Delta_{g_{t_0}} + (e^{u_{t_0}} \cdot))^{-1}(\Delta_{g_{t_0}} - \Delta_{g_t})u_{t_0}|_r < \delta_1/4. \tag{10.27}$$
$$|(\Delta_{g_{t_0}} + (e^{u_{t_0}} \cdot))^{-1}(K(g_{t_0}) - K(g_t))|_r < \delta_1/4. \tag{10.28}$$

The existence of such a δ'' follows immediately from 10.13, 10.20, (10.27) and from 10.20 for (10.28). Let now $\delta = \min\{\delta', \delta''\}$. Then we infer from (10.25)-(10.28)

$$|T_t u - u_{t_0}|_r \leq \delta_1,$$

i.e. $T_t : M_{t_0,\delta_1} \to M_{t_0,\delta_1}$. T_t is contractive according to (10.24) since for $|t - t_0| \leq \delta$

$$(C \cdot |t - t_0| + \frac{1}{4}) \leq \frac{1}{2}.$$

This finishes the existence proof of theorem 10.11 and yields uniqueness in a moving ball $M_{t,\delta_1}, 0 \leq t \leq 1$. We prove now the uniqueness in all of $\Omega^r = \Omega^{0,2,r}$.

Fix $x_0 \in M^2$ and denote by $d(x) = d(x, x_0)$ the Riemannian distance. Let $u, v \in \Omega^r, r > 3$, be solutions of

$$\Delta_g u + K(g) + e^u = 0.$$

We obtain $u, v, u-v$ bounded, C^2 and

$$\Delta_g(u-v) = -(e^u - e^v).$$

There are two cases.

1) $u-v$ achieves its supremum in $U_1(x_0) = \{x | d(x) \leq 1\}$. e.g. in x_1. Then $\Delta(u-v)(x_1) \geq 0, -(e^{u(x_1)} - e^{v(x_1)}) \geq 0,\ e^{u(x_1)} \leq e^{v(x_1)}, (u-v)(x_1) \leq 0$ at the supreme point x_1, hence $(u-v)(x) \leq 0$ everywhere, $u(x) \leq v(x)$.

2) Or we apply Yau's generalized maximum principle: $f \in C^2$,

$$\limsup_{d(x)\to\infty} \frac{f(x) - f(x_0)}{d(x)} \leq 0$$

and

$$\lim_{\substack{d(x)\to\infty \\ f(x)\geq f(x_0)}} \frac{K(x)(f(x) - f(x_0))}{d(x)} = 0.$$

Then there are points $(x_k)_k \subset M$ such that $\lim_{k\to\infty} f(x_k) = \sup f$, $\lim_{k\to\infty} \nabla f(x_k) = 0$ and $\limsup_{k\to\infty} \Delta f(x_k) \geq 0$.

In our case $f = u - v$. Then we have $(x_k)_k$ such that $\lim_{k\to\infty} (u-v)(x_k) = \sup(u-v)$, $\lim_{k\to\infty} \nabla(u-v)(x_k) = 0$, $\limsup \Delta(u-v)(x_k) \geq 0$, hence $\limsup(e^v - e^u)(x_k) \geq 0$, $\limsup(v-u)(x_k) \geq 0$, $\limsup(u-v)(x_k) \leq 0$, $\sup(u-v) \leq 0, u \leq v$ everywhere.

Quite similar $v \leq u$, i.e. $u = v$. This finishes uniqueness and the proof of theorem 10.11. □

We see, the proof of 10.11 exactly follows the scheme presented in the preceding section.

Remark 10.26 A seemingly more direct approach proving $\mathcal{S} = [0,1]$ would amount to prove the following assertion. Assume $t_1 < t_2 < \ldots < t_0, t_\nu \to t_0$, $\Delta_{g_{t_0}} u_{t_v} + e^{u_{t_\nu}} = 0$. Then

a) $(u_{t_\nu})_\nu$ is a Cauchy sequence with respect to $|\quad|_r$.

b) $u_{t_\nu} \longrightarrow u_{t_0}$

c) $\Delta_{g_{t_0}} u_{t_0} + K(g_{t_0}) + e^{u_{t_0}} = 0$.

But writing down a straightforward approach proving a), c) leads immediately to the key estimates performed by us.

11 Harmonic maps

Concerning harmonic maps, we have in view the treatment of the corresponding PDE, not the very interesting class of geometric theorems and examples. Many of them are contained in the contribution of J. Wood. Eells and Sampson presented in [42] a particular interesting and beautiful method to solve non-linear PDE, namely to connect this PDE with a heat flow, a non-linear evolution equation. This method, now is often used, e.g. in gauge theory and as Ricci flow which we will discuss in section 15.

Let (M^n, g), $(N^{n'}, h)$ be Riemannian manifolds, M^n closed. The case M^n open will be completely discussed in the forthcoming paper [50]. If $f \in C^\infty(M, N)$ then

$df \in \Gamma(T^*M \otimes f^*TN) \equiv C^\infty(T^*M \otimes f^*TN)$. ∇^g and $f^*\nabla^h$ induce metric connections ∇ in all tensor bundles $T^q_s(M) \otimes f^*T^u_v(N)$. Therefore $\nabla^m df$ is well defined. Let us introduce the energy density of f,

$$e(f) := \frac{1}{2}|df|_{T^*M\otimes f^*TN},$$

and the energy $E(f)$,

$$E(f) := \int\limits_M e(f)\,\mathrm{dvol}_x(g).$$

As well known, the Euler-Lagrange equations for $E(\cdot)$ are

$$\tau(f) := \mathrm{tr}_g \nabla df = 0. \tag{11.1}$$

f is called harmonic if $\tau(f) = 0$. Examples are harmonic functions, geodesics and minimal submanifolds.

Now, the most interesting question is the question for the existence. There are many answers. We present one of them. The criterion for our choice was the proof of the corresponding existence theorem which relies on the heat flow and so exhibits another specific method to solve non-linear PDEs on manifolds. This proof goes back to Eells/Sampson in [42].

Theorem 11.1 *Suppose (M^n, g), $(N^{n'}, h)$ closed, $(N^{n'}, h)$ non-positively curved. Then any continous $f_0 : M \longrightarrow N$ is homotopic to a harmonic map.*

The proof will be performed by means of the heat flow and the idea of this goes back to Eells/Sampson. Consider the initial value problem

$$\begin{aligned} \frac{\partial}{\partial t} f(x,t) &= \tau(f(x,)), \\ f(x,0) &= f_0. \end{aligned} \tag{11.2}$$

The proof of theorem 11.1 consists of 5 steps,

1) the existence of $f(x,t)$ for small t,
2) the existence of $f(x,t)$ for all t,
3) the existence of $\lim\limits_{t\to\infty} f(x,t) = f(x,\infty) = f(x)$
4) $f(x)$ is harmonic,
5) $f(x)$ is homotopic to f_0.

In the sequel, we always suppose that (N, h) has non-negative curvature. We start with the first step. Let $x \in M$, $f(x) = y \in N^{n'}$, $x^1, \ldots, x^n$ coordinates about x, $y^1, \ldots, y^{n'}$ coordinates about y. First we assume $f_0 \in C^{2+\alpha}(M,N) \equiv {}^{b,2+\alpha}\Omega(M,N)$, where the latter can be defined quite analogous to $\Omega^{p,r}(M,N)$ in section 3.

In local coordinates (11.2) looks

$$\begin{aligned} &\left(\frac{\partial}{\partial t} f(x,t) - \tau f(x,t)\right)^i \\ &\quad = \frac{\partial}{\partial t} f^i - \left(g^{\alpha\beta}\frac{\partial^2 f^i}{\partial x^\alpha \partial x^\beta} - g^{\alpha\beta}\Gamma^\gamma_{\alpha\beta}(g)\frac{\partial f^i}{\partial x^\gamma} + g^{\alpha\beta}\Gamma^i_{jk}(h)\frac{\partial f^j}{\partial x^\alpha}\frac{\partial f^k}{\partial x^\beta}\right) \\ &\quad = 0. \end{aligned} \tag{11.3}$$

Taking the linearization of (11.3), we see immediately that (11.2) is a quasi-linear parabolic system. The linearization has a solution by standard theorems, the implicit function theorem in the Banach category then yields a solution $f(x,t)$ of (11.2) with $f(x,0) = f_0$ for $0 \le t \le \varepsilon$. The same argument yields that

$$\mathcal{S} = \{T \in]0,\infty[\mid \text{(11.2) has a solution for } 0 \le t \le T\}$$

is open.

To accomplish step two, we would be done if the set $\mathcal{S}$ would be closed. This requires a series of lemmas.

Lemma 11.2 *Suppose* $u(x,t) \in C^2$, $u \ge 0$

$$-\Delta u - \frac{\partial}{\partial t} u \ge -cu \quad \text{on } [0,T]$$

and

$$\frac{d}{dt}\int_M u(x)\,\mathrm{dvol}_x(g) \le 0,$$

where $-\Delta = -\nabla^*\nabla \le 0$. *Let* $0 < R < \min\{r_{\mathrm{inj}}(M), \frac{\pi}{2\Lambda}\}$, $\Lambda = (\max|\text{sectional curvature}|)^{\frac{1}{2}}$. *Then*

$$u(x,t) \le c(TR^{-n-2} + T^{-\frac{n}{2}})\int_M u(y,0)\,\mathrm{dvol}_x(g).$$

Furthermore, for any $t_0 < T$*, in particular* $t_0 = 0$*,*

$$u(x,t) \le cR^{-2}\sup_{y\in M} u(y,t_0).$$

We refer to [80], p. 82–86, for the proof. □

Lemma 11.3 *If* $f(x,t)$ *is a solution of (11.2), then*

$$-\Delta e(f) - \frac{\partial}{\partial t}e(f) = |\nabla df|^2 + \frac{1}{2}\langle df\,\mathrm{Ric}^{\,M}(e_\alpha), df e_\alpha\rangle - \langle R^N_m(df e_\alpha, df e_\beta) df e_\beta, df e - \alpha\rangle,$$

where $e_1,\dots,e_n \in T_xM$ *is an orthonormal basis.*

This is a simple calculation. □

Corollary 11.4 $-\Delta e(f) - \frac{\partial}{\partial t}e(f) \ge -ce(f)$. □

Lemma 11.5 *If* $f(x,t)$ *is a solution of (11.2), then* $E(f(\cdot,t))$ *is a decreasing function of* t*.*

Proof

$$\frac{d}{dt}E(f(\cdot,t)) = \frac{d}{dt}\frac{1}{2}\int (df,df)\,\mathrm{dvol}_x(g) = \int \left(\nabla_{\frac{\partial}{\partial t}} df, df\right)\,\mathrm{dvol}_x(g) =$$

$$= \int \left(d\frac{\partial}{\partial t}f, df\right)\,\mathrm{dvol}_x(g) = -\int \left(\frac{\partial}{\partial t}f, \tau(f)\right)\,\mathrm{dvol}_x(g) = -\int \left|\frac{\partial}{\partial t}f\right|^2\,\mathrm{dvol}_x(g)$$

□

Lemma 11.6 *Let f be a solution of (11.2) on $[0,T]$ and $0 < R < \min\left\{r_{\text{inj}}(M), \frac{\pi}{1\Lambda}\right\}$. Then, for all $x \in M$*

$$e(f) \leq c(TR^{-n-2} + T^{-\frac{n}{2}}) \int\limits_M e(f)(y,0)\,\mathrm{dvol}_y(g)$$

and for every $t_0 < T$, in particular $t_0 = 0$,

$$e(f)(x,T) \leq cR^{-2} \sup_{y \in M} e(f)(y,t_0).$$

Proof Set $u(x,t) = e(f)(x,t)$. Then 11.4, 11.5 assure the assumptions of lemma 11.2, apply 11.2. □

We cite a lemma of Hartman (cf. [75]).

Lemma 11.7 *Let $f(x,t,s)$ be a smooth family of solutions of (11.2) depending on a parameter s and having initial values $f(x,0,s) = g(x,s)$, $0 \leq s \leq s_0$. Then for every $s \in [0,s_0]$ there holds*

$$\sup_{x \in M} g_{ij} f(x,t,s) \left(\frac{\partial f^i}{\partial s} \cdot \frac{\partial f^j}{\partial s} \right)$$

is non-increasing in t. Hence also

$$\sup_{x \in M, s \in [0,s_0]} \left(g_{ij} \frac{\partial f^i}{\partial s} \cdot \frac{\partial f^j}{\partial s} \right)$$

is non-increasing in t.

We refer to [80] for the proof. □

Corollary 11.8 *Suppose the hypotheses of 11.6. Then*

$$\sup_{x \in M} \mathrm{dist}(f(x,t,0), f(x,t,1))$$

is non-increasing in t for $t \in [0,T]$. □

Lemma 11.9 *Suppose the hypotheses of 11.6. Then for all $t \in [0,T[$ and $x \in M$*

$$\left| \frac{\partial f(x,t)}{\partial t} \right| \leq \sup_{x \in M} \left| \frac{\partial}{\partial t} f(x,0) \right|.$$

Proof Set $f(x,t,s) = f(x,t+s)$ and apply 11.7 at $s = 0$. □

Lemma 11.10 *Suppose the hypothesis of 11.6. Then for every $\alpha \in]0,1[$*

$$|f(\cdot,t)|_{C^{2+2}(M,N)} + \left| \frac{\partial f}{\partial t}(\cdot,t) \right|_{C^\alpha(M,N)} \leq c, \tag{11.4}$$

where c depends on α, the initial value $g(\cdot) = f(\cdot,0)$ and the geometry of M and N.

Proof (11.2) means in local coordinates

$$g^{\alpha\beta}\frac{\partial^2 f^i}{\partial x^\alpha \partial x^\beta} - g^{\alpha\beta}\Gamma(g)^\gamma_{\alpha\beta}\frac{\partial f^i}{\partial x^\gamma} = g^{\alpha\beta}\Gamma(h)^i_{jk}\frac{\partial f^j}{\partial x^\alpha}\frac{\partial f^k}{\partial x^\beta} + \frac{\partial f^i}{\partial t}. \tag{11.5}$$

According to 11.6 – 11.9, for a fixed neighbourhood $B_\varrho(x) \times [t_0, t_1]$, $f(x,t)$ will stay in the given chart and, moreover, the right hand side of (11.5) is bounded. Hence $|f(\cdot,t)|_{C^{1+\alpha}(M,N)}$ is bounded by elliptic regularity. This implies that the right hand side of

$$\frac{\partial f^i}{\partial t} - g^{\alpha\beta}\frac{\partial^2 f^i}{\partial x^\alpha \partial x^\beta} + g^{\alpha\beta}\Gamma(g)^\gamma_{\alpha\beta}\frac{\partial f^i}{\partial x^\gamma} = g^{\alpha\beta}\frac{\partial f^j}{\partial x^\alpha}\frac{\partial f^k}{\partial x^\beta}$$

is bounded in $C^\alpha(M,N)$ which finally yields (11.4) (cf. theorem 2.2.1 in [80]). □

Proposition 11.11 *The equation (11.2) has a solution for all* $t \in [0,\infty[$.

Proof We have seen already that the set $\mathcal{S}$ above is non-empty and open. Lemma 11.10 implies the closedness. □

This finishes step two.

Lemma 11.12 *There exists a sequence* $t_n \longrightarrow \infty$ *such that* $\lim\limits_{n\to\infty} \frac{\partial f}{\partial t}(x,t_n) = 0$ *uniformly in* $x \in M$.

Proof $E(f(\cdot,t))$ is always non-negative, $\left|\frac{\partial f}{\partial t}\right|$ has a time independent C^α-bound and according to the proof of 11.5,

$$\frac{\partial}{\partial t}E(f(\cdot,t)) = -\int_M \left|\frac{\partial f(x,t)}{\partial t}\right|^2 \mathrm{dvol}_x(g).$$

□

The $C^{2+\alpha}$-bounds for $f(\cdot,t)$ in lemma 11.10 imply that (possibly passing to a subsequence) $f(\cdot,t_n)$ converges for $t_n \to \infty$ uniformly to a harmonic map. Put in corollary 11.8

$$\begin{aligned} f_0(x,0) &= f(x,0,0) = f(x,t_n), \\ f_0(x,s_0) &= f(x,0,\xi_0) = f(x). \end{aligned}$$

We infer from the uniform convergence that some $f(\cdot,t_n)$ are homotopic to f. The same holds for all t since $f(x,t)$ is continuous in t. $f(x)$ is harmonic and a time independent solution of (11.2), hence $f(x,t,s_0) = f(x)$ for all t. We apply corollary 11.8 and obtain

$$\mathrm{dist}(f(x,t_n+t), f(x)) \le d(f(x,t_n), f(x)) \quad \text{for all } t \ge 0.$$

Hence choice of a subsequence is not necessary and $\lim\limits_{t\to\infty} f(x,t) = f(x)$ uniformly. We have accomplished the steps 3 – 5 in the case $f_0 \in C^{2+\alpha}$. If f_0 is only continuous, that it is a very simple standard approximation argument that f_0 is homotopic to a $C^{2+\alpha}$-map. This finishes the proof of theorem 11.1. □

Our goal in this section was to present a striking example for the heat flow method. By the same method one can attack to Yang-Mills equation.

12 Non-linear field theories

In this section, we present quite another approach to certain non-linear field equation. We consider those field theories whose field equations are essentially expressed in geometric notions. Then one searches for those geometries which satisfy the field equations and singles out the others. In the case of a system, one does this step by step, equation by equation. Even the underlying manifold (class of manifolds) is not fixed.

This approach has been essentially and very successfully elaborated by Friedrich, Agricola, Ivanov, Kim and others. We refer to [1], [2], [60], [61]. Historically, this kind of approach is not new. It was a striking break-through as the gauge potential of gauge theory (of physicists) has been recognized has a connection in a fibre bundle. Then mathematicians were able to apply the powerful tools of differential geometry. Another similar example was 1994 Seiberg-Witten theory which we cannot discuss here for reasons of space. We refer to [51], [85], [88]. In the sequel, we will briefly discuss the Einstein-Dirac equation on Riemannian spin manifolds and the type II B string theory.

Let (M^n, g) be a Riemannian spin manifold, $R(g)$ its scalar curvature, $D = D_g$ the Dirac operator acting on spinar fields ψ, and let $\varepsilon = \pm 1$ and $\lambda \in \mathbb{R}$ be two real parameters. Consider for an arbitrary open, bounded set U the Lagrange functional

$$W(g,\psi) := \int_U (R(g) + \varepsilon\{\lambda(\psi,\psi) - (D_g\psi,\psi)\}) \,\mathrm{dvol}_x(g).$$

The Euler-Lagrange equations are the Dirac equation

$$D_g\psi = \lambda\psi \tag{12.1}$$

and the Einstein equation

$$\mathrm{Ric}\,(g) - \frac{1}{2}R(g) = \frac{\varepsilon}{4}T_{(g,\psi)}, \tag{12.2}$$

where the energy-momentum tensor $T_{(g,\psi)}$ is given by

$$T_{(g,\psi)}(X,Y) := (X \cdot \nabla^g_Y \psi + Y \cdot \nabla^g_X \psi, \psi).$$

$R(g)$ and λ are related by

$$R = \mp \frac{\lambda}{n-2}|\psi|^2.$$

ψ is called an Einstein spinor for the eigenvalue λ. By rescaling the spinor field we can assume that the parameter ε equals ± 1, i.e. we have finally to consider the nonlinear system

$$D\psi = \lambda\psi, \quad \mathrm{Ric}\,(g) - \frac{1}{2}R(g) = \pm\frac{1}{4}T_\psi.$$

Example 12.1 Suppose (M^n, g) carries a Killing spinor φ of positive (resp. negative) Killing number $b \in \mathbb{R}$, i.e. $\nabla_X\psi = bX \cdot \psi$ for all $X \in TM$. Then $\psi := \sqrt{4(n-1)(n-2)|b|}\varphi/|\varphi|$ is a positive (resp. negative) Einstein spinor for the eigenvalue $\lambda = -nb$. In this case (M^n, g) is an Einstein manifold with $\mathrm{Ric} = 4(n-1)b^2 g$. □

Remark 12.2 For any Riemannian surface (M^2, g) we have $\mathrm{Ric} - \frac{1}{2}Sg = 0$. Consequently, we always assume that the dimension of the manifold is at least 3. □

In the sequel, we follow [61].

Definition. Let (M^n, g) bea Riemannian spin manifold whose scalar curvature S does not vanish at any point. A non-trivial spinor field ψ will be called a *weak Killing spinor (WK-spinor) with WK-number* $\lambda \in \mathbb{R}$ if ψ is a solution of the first order differential equation

$$\nabla_X\psi = \frac{n}{2(n-1)R} dR(X)\cdot\psi + \frac{2\lambda}{(n-2)R}\mathrm{Ric}\,(X)\cdot\psi - \frac{\lambda}{n-2}X\cdot\psi + \frac{1}{2(n-1)R}X\cdot dS\cdot\psi.$$

Remark 12.3 The notion of a WK-spinor is meaningful even in case that the WK-number λ is a complex number. In this section we consider only the case that $\lambda \neq 0$ is real. However, the examples of Riemannian spaces M^n with imaginary Killing spinors (see [61]) show that Riemannian manifolds admitting WK-spinors with imaginary Killing numbers exist. □

In case (M^n, g) is Einstein, the above equation reduces to $\nabla_X\psi = -\frac{\lambda}{n}X\cdot\psi$ and coincides with the Killing equation. Together with the following theorem, this justifies the name; however, notice that the vector field $V_\psi(X) = \sqrt{-1}\langle X\cdot\psi, \psi\rangle$ associated to a WK-spinor is in general not a Killing vector field. Using the formula $R\psi = -\sum_{u=1}^{n} E_u\cdot\mathrm{Ric}\,(E_u)\cdot\psi$, one checks easily that every WK-spinor of WK-number λ is an eigenspinor of the Dirac operator with eigenvalue λ. WK-spinors occur in the limiting case of an eigenvalue estimate for the Dirac operator and they are closely related to the Einstein spinors, as will be explained in the next theorem.

Theorem 12.4 *Let ψ be a WK-spinor on (M^n, g) of WK-number λ with $\lambda R < 0$ (resp. $\lambda R > 0$). Then $\frac{|\psi|^2}{R}$ is constant on M^n and $\varphi = \sqrt{\frac{(n-2)|R|}{|\lambda||\psi|^2}}\psi$ is a positive (resp. negative) Einstein spinor to the eigenvalue λ, i.e. (12.1) and (12.2) are solved.* □

For this reason, we ask for the existence of WK-spinors. In the case $n = \dim M = 3$, the existence of an WK-spinor and an Einstein spinor are equivalent in the following sense.

Theorem 12.5 *Suppose that the scalar curvature R of (M^3, g) does not vanish at any point. Then (M^3, g) admits a WK-spinor of WK-number λ with $\lambda R < 0$ (resp. $\lambda R > 0$) if and only if (M^3, g) admits a positive (resp. negative) Einstein spinor with the same eigenvalue λ.*

Now the procedure of Friedrich/Kim is to single out those Riemannian manifolds which admit WK-spinors or which don't admit WK-spinors, respectively. We present some of their theorems.

Theorem 12.6 *Let (M^n, g) be compact with positive scalar curvature R. If $|\mathrm{Ric}\,|^2 \geq \frac{1}{4}(n^2 - 5n + 8)R^2$ at all points, then (M^n, g) does not admit WK-spinors.* □

Theorem 12.7 *Let (M^n, g) be a conformally flat or Ricci-parallel Riemannian spin manifold with constant scalar curvature $R \neq 0$ and suppose that it admits a WK-spinor. Then the following two equations hold at any point of M^n:*

a) $4R\mathrm{Ric}^2 - \{n(n-3)R^2 - 4|\mathrm{Ric}\,|^2\}\mathrm{Ric} - (n-3)R^3\,\mathrm{id} = 0$

b) $4|\mathrm{Ric}\,|^4 - 4R\{\mathrm{tr}(\mathrm{Ric}^3)\} - n(n-3)R^2|\mathrm{Ric}\,|^2 + (n-3)R^4 = 0.$

In particular, the Ricci tensor is non-degenerate at any point for $n \geq 4$. □

As an immediate consequence of the preceding theorem, we shall list some sufficient conditions for a product manifold not to admit WK-spinors.

Corollary 12.8 *Let (M^p, g_M) and (N^q, g_N) be the Riemannian spin manifolds. The product manifold $(M^p \times N^q, g_M \times g_N)$ does not admit WK-spinors in any of the following cases:*

a) *(M^p, g_M) and (N^q, g_N) are both Einstein and the scalar curvatures $R_M \equiv R(g_M)$, $R_N \equiv R(g_N)$ are positive $(p, q \geq 3)$.*

b) *(M^p, g_M) is Einstein with $R_M > 0$ and (N^2, g_N) is the 2-dimensional sphere of constant curvature $(p \geq 3)$.*

c) *(M^2, g_M) and (N^2, g_N) are spheres of constant curvature.*

d) *(M^p, g_M) is Einstein and (N^q, g_N) is a q-dimensional flat torus $(q \geq 1, p \geq 3)$.* □

Next we present three theorems which contain under certain additional assumption necessary conditions for the existence of a WK-spinor.

Theorem 12.9 *Let (M^n, g) be conformally flat, Ricci parallel and with non-zero scalar curvature $(n \geq 4)$. If M^n admits a WK-spinor , then*

a) *(M^n, g) is Einstein, if $R > 0$,*

b) *the equation $|\mathrm{Ric}\,|^2 = \frac{n^3-4n^2+3n+4}{4(n-1)} R^2$ holds if $R < 0$.* □

Theorem 12.10 *Suppose that (M^p, g_M) as well as (N^q, g_N) are Einstein and that $R_M \neq 0$, $R_N \neq 0$, $R = R_M + R_N \neq 0$ $(p, q \geq 3)$. If the product manifold $M^p \times N^q$ admits WK-spinors, then either $(p-2)R_M + pR_N = 0$ or $qR_M + (q-2)R_N = 0$ holds.* □

Theorem 12.11 *Let (M^p, g_M) be an Einstein space with scalar curvature $R_M \neq 0$ and (N^q, g_N) be non-Einstein with constant scalar curvature $R_N \neq 0$ $(p, q \geq 3)$. Suppose that $R_M + R_N \neq 0$ and $M^p \times N^q$ admits a WK-spinor. Then we have $(p - 29R_M + pR_N = 0$.* □

The following three theorems single out classes of manifolds which do not admit WK-spinors.

Theorem 12.12 *A manifold (M^n, g) of constant curvature $R \neq 0$ and with a parallel 1-form does not admit WK-spinors $(n \geq 3)$.* □

Theorem 12.13 *Suppose that the scalar curvature R_M of (M^p, g_M) as well as the scalar curvature R_N of (N^q, g_N) are constant and non-zero $(p, q \geq 3)$. Furthermore, suppose the scalar curvature $R = R_M + R_N$ of the product $(M^p \times N^q, g_M \times g_N)$ is not zero. If neither (M^p, g_M) or (N^q, g_N) is Einstein, then the product manifold $(M^p \times N^q, g_M \times g_N)$ does not admit WK-spinors.*

□

Theorem 12.14 *Suppose the scalar curvature R_M of (M^p, g_M), $(P \geq 3)$ is constant and non-zero. If the scalar curvature R_N of (N^q, g_N) $(q \geq 1)$ equals identically zero, then the product manifold $(M^p \times N^q, g_M \times g_N$ does not admit WK-spinors.* □

But Friedrich/Kim show in [61] that special types of product manifolds admit Einstein spinors which are not WK-spinors.

Summarizing, we see that (12.1), (12.2) have not been solved by purely analytical methods, e.g. by deformation of the equations, the continuity method, the heat flow or something like that, but (12.1), (12.2) have been solved by finding out appropriate geometries which offer natural solutions. For this approach, the authors had to reformulate (12.1), (12.2) and potential geometries such that they can be connected. We completely suppressed the corresponding (sometimes very hard) calculations and refer to [61].

The second example is the type II B string theory which consists of a Riemannian manifold, a metric connection ∇ with totally skew-symmetric torsion T and a non-trivial spinor field ψ, where these objects are related by the equations

$$\mathrm{Ric}^{\nabla} = 0, \quad \nabla\Psi = 0, \quad , \delta(T) = 0, \quad T \cdot \Psi = \mu\Psi. \tag{12.3}$$

Here $\delta(T)$ denotes the divergence. There holds $\delta^{\nabla}(T) = \delta^{g}(T)$ since ∇ is a metric connection with totally skew-symmetric torsion. μ can be an arbitrary function, but we will restrict it soon. Again the whole procedure consists in a step-by-step finding geometries which satisfy one, two, ... equations of (12.3).

Theorem 12.15 *Let* $(M^n, g, \nabla, T, \Psi, \mu$ *be a solution of*

$$\nabla\Psi = 0, \quad \delta(T) = 0, \quad T \cdot \Psi = \mu\Psi$$

and assume that the spinor field Ψ *is non-trivial. Then the function* μ *is constant.* □

It is possible to impose for the Ricci tensor the weaker condition

$$\mathrm{div}(R^{\nabla}) = 0.$$

Then the last 3 equations of (12.3) imply conditions for $\mathrm{div}(\mathrm{Ric}^{\nabla})$.

Theorem 12.16 *Let* $(M^n, g, \nabla, T, \Psi, \mu$ *be a solution of*

$$\nabla\Psi = 0, \quad \delta(T) = 0, \quad T \cdot \Psi = \mu\Psi$$

and assume that the spinor field Ψ *is non-trivial. Then the Riemannian and the* ∇*-divergence of the Ricci tensor* Ric^{∇} *coincide,* $\mathrm{div}^{g}(\mathrm{Ric}^{\nabla}) = \mathrm{div}^{\nabla}(\mathrm{Ric}^{\nabla})$. *Moreover,* $\mathrm{div}(\mathrm{Ric}^{\nabla})$ *vanishes if and only if* $\delta^{\nabla}(dT) \cdot \Psi = 0$ *holds.* □

In the case (M^n, g) compact and $\mu = 0$, we have an overview of the geometries in question, if $n \leq 8$.

Theorem 12.17 *Let* (M^n, g) *be closed and suppose for a non-trivial spinor field* Ψ

$$\mathrm{Ric}^{\nabla} = 0, \quad \nabla\Psi = 0, \quad T \cdot \Psi = 0.$$

Then $T = 0$ *and* ∇ *is the Levi-Civita connection.* □

A systematic approach to find out the appropriate geometries is contained in [1], [60]. The starting point is the holonomy group of a G-structure. First the authors characterize all G-structures admitting a G-connection with a totally skew-symmetric torsion tensor T. Then the authors apply this general method to the subgroup $G_2 \subset SO(7)$. The next is to ask for those connections which additionally admit a parallel spinor. One singles out the classes of geometries without such a spinor and the remaining classes remain under investigation. The last step is to investigate the validity of $\mathrm{Ric}^{\nabla} = 0$.

Start with the group $G_2 \subset SO(2)$. The group G_2 is the isotropy group of the 3-form in seven variables

$$\omega^3 := e_1\wedge e_2\wedge e_7 + e_1\wedge e_3\wedge e_5 - e_1\wedge e_4\wedge e_6 - e_2\wedge e_3\wedge e_6 - e_2\wedge e_4\wedge e_5 + e_3\wedge e_4\wedge e_7 + e_5\wedge e_6\wedge e_7.$$

The 3-form ω^3 corresponds to a real spinor $\Psi_0 \in \Delta_7$ and, therefore, G_2 can be defined as the isotropy group of a non-trivial real spinor.

We consider 7-manifolds.

Theorem 12.18 *Let (M^7, g, ω^3) be a 7-dimensional Riemannian manifold with a nearly parallel G_2-structure $(\Gamma = \lambda \cdot \mathrm{id})$. Then there exists a unique affine connection ∇ such that*

$$\nabla\omega^3 = 0 \quad \textit{and} \quad T \textit{ is a 3-form.}$$

The torsion tensor is given by the formula $6 \cdot T = (d\omega^3, *\omega^3) \cdot \omega^3$. T is ∇-parallel and coclosed, $\nabla T = \delta T = 0$. □

Corollary 12.19 *Let (M^7, g, ω^3) be a 7-dimensional nearly parallel G_2-manifold. Then the triple $(M^7, g, T^* := 3 \cdot T)$ is a solution of the string equations with constant dilation:*

$$\mathrm{Ric}^g_{ij} - \frac{1}{4}T^*_{imn}T^*_{jmn} = 0, \quad \delta^g(T^*) = 0.$$

□

A *cocalibrated* G_2-structure is defined by the condition that ω^3 is coclosed, $\delta^g(\omega^3) = 0$.

Theorem 12.20 *Let $(M^7, g, \omega^3, \nabla)$ be a 7-dimensional compact nearly parallel G_2-manifold and ∇ be the unique G_2-connection with totally skew-symmetric torsion. Then every ∇-harmonic spinor Ψ is ∇-parallel. Moreover, the space of ∇-parallel spinors is one-dimensional.*

□

We conclude this section with some 5-dimensional examples.

Theorem 12.21 *Let $(M^5, g, \xi, \eta, \varphi)$ be a simply connected 5-dimensional Sasakian spin manifold and consider the unique linear connection ∇ with totally skew-symmetric torsion preserving the Sasakian structure. There exists a ∇-parallel spinor in the subbundle defined by the algebraic equation $\xi \cdot \Psi = i \cdot \Psi$ if and only if the Riemannian Ricci tensor of M^5 has the eigenvalues (6, 6, 6, 6, 4). A ∇-parallel spinor of this algebraic type is an eigenspinor of the Riemannian Dirac operator, $D^g\Psi = \pm 3 \cdot \Psi$. In case M^5 is compact, any ∇-harmonic spinor Ψ is ∇-parallel.*

□

Certain S^1-bundles over 4-dimensional Kähler-Einstein manifolds supply examples.

Theorem 12.22 *Let $(M^5, g, \xi, \eta, \varphi$ be a 5-dimensional Sasakian spin manifold and consider the unique linear connection ∇ with totally skew-symmetric torsion preserving the Sasakian structure. If there exists a ∇-parallel spinor in the subbundle defined by the algebraic equation $d\eta \cdot \Psi = 0$, then the Riemannian Ricci tensor of M^5 has the eigenvalues (−2, −2, −2, −2, 4. Any ∇-parallel spinor in this 2-dimensional subbundle satisfies the equations*

$$\nabla^g_\xi = 0, \quad \nabla^g_X\Psi = \frac{1}{2}\varphi(X) \cdot \xi \cdot \Psi = -\frac{i}{2}\varphi(X) \cdot \Psi, \quad d\eta \cdot \Psi = 0.$$

In particular, it is harmonic with respect to the Riemannian connection. Any ∇-harmonic spinor Ψ on a compact manifold M^5 satisfying the algebraic condition $d\eta \cdot \Psi = 0$ is ∇-parallel.

□

We refer to [1], [2], [60], [61] for the proofs of this section. These papers contain quite a lot of further interesting results.

13 Gauge theory

Gauge theory brought a striking break-through: the moduli space of solutions of an equation of mathematical physics on a closed 4-manifold M^4 reflects deep topological features of this M^4. An essential step in this approach was the proof that the moduli space was not empty. This has been established by Taubes in [116], [117]. With other words, he solved the self-duality equation. His approach consists in two main steps:

1) grafting of the t' Hooft solution to the M^4,

2) taking this grafted solution as the first approximation of a global solution on M^4.

We extended in [49] this procedure to open manifolds and present here an outline of our approach.

First we recall for clarity and completeness very briefly the basic notions of gauge theory.

Let (M^n, g) be a Riemannian manifold, G a compact Lie group, $\mathfrak{g}$ its Lie algebra, and $\varrho : G \longrightarrow O(E^N)$ a faithful orthogonal representation. Consider the map $Ad : G \longrightarrow \mathrm{Aut}(G)$, $Ad(a)(g) = a^{-1}ga$, its derivative ad and $\varrho' : G \longrightarrow \mathfrak{o}(E) = \mathrm{Lie}[O(E)]$, $\varrho'(a)e = \frac{d}{dt}\varrho[\exp(ta)]e|_{t=0}$, $e \in E^N$. The bundles $G_P = P \times_{Ad} G$, $\mathfrak{g}_P = P \times_{ad} \mathfrak{g}$, $E = P \times_{\varrho} E^N$, $O_E = P \times_{\varrho} O(E^N)$ and $\mathfrak{o} = P \times_{\varrho'} \mathfrak{o}(E)$ are associated with P. There are embeddings $i_1 : G_P \longrightarrow O_E$ and $i_2 : \mathfrak{g}_P \longrightarrow \mathfrak{o}_E$. Set $G_E = \mathrm{Im}\, i_1$, $\mathfrak{g}_E = \mathrm{Im}\, i_2$. Then $G_P \cong G_E$, $\mathfrak{g}_P \cong \mathfrak{g}_E$. Here $G_{E,x}$ denotes the group of all orthogonal transformations of E_x, and $\mathfrak{g}_{E,x}$ is the algebra of all skew symmetric endomorphisms of E_x. A bundle automorphism $f : P \longrightarrow P$ over id_M with $f(u \cdot a) = f(u) \cdot a$, $u \in P$, $a \in G$, is called a gauge transformation. The set of all gauge transformations forms a group, the gauge group $\mathcal{G}_P$. Consider further $\tilde{\mathcal{G}}_P = C^\infty(G_P)$ and $\hat{\mathcal{G}}_P = \{f : P \longrightarrow G | f(u \cdot a) = a^{-1}f(u)a\}$. Then $\mathcal{G} \cong \hat{\mathcal{G}} \cong \tilde{\mathcal{G}}$ as groups. The isomorphisms are given by $f \in \mathcal{G} \longrightarrow \tilde{f} = [(u, \hat{f}(u))] \in \mathcal{G}_P$, where $\hat{f}(u)$ is defined by $f(u) = u \cdot \hat{f}(u)$, and $f \in \mathcal{G}_P \longrightarrow \hat{f} \in \hat{\mathcal{G}}_P$. A connection for P or E, respectively, is given by:

1) a smooth field of horizontal subspaces $H_u \subset T_uP$, $H_{u,a} = (Ra)_* H_u$; or

2) a connection form $\omega : TP \longrightarrow \mathfrak{g}$, $R_a^*\omega = ad(a^{-1})\omega$, $\omega(A^*) = A$, where A^* is a fundamental vector field, $A_u^*\varphi = \frac{d}{dt}\varphi[u \cdot \exp(tA)]|_{t=0}$; or

3) a field of horizontal subspaces in TE, compatible with the representation; or

4) a covariant derivative $\nabla^\omega : \Omega^0(E) \longrightarrow \Omega^1(E)$, $\nabla^\omega(f \cdot e) = df \otimes e + f \cdot \nabla^\omega e$, where $f \in C^\infty(M)$, $e \in \Omega^0(E)$.

∇^ω is a metric connection since E comes from an orthogonal (or unitary) representation compatible with the connection.

The equivalence 1) $\longleftrightarrow$ 3) is given by the mapping of horizontal curves into horizontal spaces, 1) $\longleftrightarrow$ 2) by $H_u = \ker\omega$ and 3) $\longleftrightarrow$ 4) by $\nabla_x e = \pi_v e_*(X)$, where π_v denotes the projection onto the vertical subspaces. Locally the description is given as follows. Let $\{(U_\alpha, \varphi_\alpha)\}_\alpha$ be an atlas of bundle charts for $P(M,G)$, $\varphi_\alpha : \pi^{-1}(U_\alpha) \xrightarrow{\cong} U_\alpha \times G$, $\varphi_\alpha(u) = (\pi(u), h_\alpha(u))$ and $\sigma_\alpha : U_\alpha \longrightarrow \pi^{-1}(U_\alpha)$, $\sigma_\alpha(x) = u \cdot h_\alpha^{-1}(u)$ local sections and $\psi_{\alpha\beta} : U_\alpha \cap U_\beta \longrightarrow G$, $\psi_{\alpha\beta}(x) = h_\alpha(u)h_\beta(u)^{-1}$, $u \in \pi^{-1}(x)$, transition functions. If $\theta : TG \longrightarrow \mathcal{G}$ is the canonical form, then with $\omega_\alpha = \sigma_\alpha^*\omega$ and $\theta_{\alpha\beta} = \psi_{\alpha\beta}^*\theta$, we obtain by differentiation of $d_\beta(x) = \sigma_\alpha(x) \cdot \psi_{\alpha\beta}(x)$

$$\begin{aligned} \omega_\beta &= \sigma_\beta^* \omega = \mathrm{ad}[\psi_{\alpha\beta}(x)^{-1}]\omega_\alpha + \theta_{\alpha\beta}, \\ \omega &= \mathrm{ad}(h_\alpha^{-1})\pi^*\omega_\alpha + h_\alpha^*\theta. \end{aligned}$$

Denote by $\mathcal{C}_P$ the set of all connections ("gauge potentials") on P. $\mathcal{G}$ acts on $\mathcal{C}_P$ from the right by $\omega \cdot f := f^*\omega$. In a local chart, f defines local maps $f_\alpha : U_\alpha \longrightarrow G$ by $f[\sigma_\alpha(x)] = \sigma_\alpha(x) \cdot f_\alpha(x)$, and $\sigma_\alpha^*(\omega \cdot f) = \sigma_\alpha^*(f^*\omega) = (f\sigma_\alpha)^*\omega = \mathrm{ad}(f_\alpha^{-1})\omega_\alpha + f_\alpha^*\theta$, i.e.

$$(\omega \cdot f)_\alpha = \mathrm{ad}(f_\alpha^{-1})\omega_\alpha + f_\alpha^*\theta.$$

Denote by $\mathcal{C}_E$ the set of all metric connections ∇^ω on E. Clearly, $\mathcal{C}_P \cong \mathcal{C}_E$, since ϱ was a faithful representation. $\mathcal{G}_E = C^\infty(G_E)$ acts on $\mathcal{C}_E$ by $\nabla^{(\omega\cdot f)} = f^{-1} \circ \nabla^\omega \circ f$, i.e. on sections σ,

$$\nabla^{(\omega \circ f)}\sigma = f^{-1}\nabla^\omega(f\sigma).$$

Let $\omega, \omega' \in \mathcal{C}_P$. Then $\omega - \omega' \in \Omega^1(\mathfrak{g}_P)$, $\nabla^\omega - \nabla^{\omega'} \in \Omega^1(\mathfrak{g}_E)$ and $\mathcal{C}_P \cong \mathcal{C}_E$ is an affine space with vector space $\Omega^1(\mathfrak{g}_P) \cong \Omega^1(\mathfrak{g}_E)$. Denote by $D^\omega = \pi_h \circ d$ the covariant differentiation, then $D^\omega\omega = d\omega + \frac{1}{2}[\omega, \omega] = R^\omega \in \Omega^2(\mathfrak{g}_P)$ is the curvature of ω, in physical terminology the field strength of the gauge potential ω. The functional

$$\omega \in \mathcal{C}_P \cong \mathcal{C}_E \longrightarrow YM(\omega) = \frac{1}{2}\int |R^\omega|_x^2 \,\mathrm{dvol}_x(g)$$

is called the Yang-Mills functional. It is clear, that $YM(\omega)$ will most often be ∞ on open manifolds. Therefore, on open manifolds we must restrict an attention to connections with finite Yang-Mills action. But that we get for bounded geometry still the reasonable space

$$\mathcal{C}_E^{2,r}(B_k, f, 2) = \sum_{i \in I} \mathrm{comp}^{2,r}(\nabla_i)$$

is just the content of proposition 3.33. We assume in the sequel (I), (B_k) and restrict to a component in $\mathcal{C}_E^{2,r}(B_k, f, 2)$.

The stationary connections of $YM(\omega)$ are called Yang-Mills connections or Yang-Mills potentials. Define for $\omega \in \mathcal{C}_P$ the Laplacian as in I 1 by $\Delta^\omega = d^\omega\delta^\omega + \delta^\omega d^\omega : \Omega^q(\mathfrak{g}_E) \longrightarrow \Omega^q(\mathfrak{g}_E)$. The important case here is $q = 2$. The following conditions are equivalent:

1) ω is a Yang-Mills potential,

2) $\delta^\omega R^\omega = 0$,

3) $\Delta^\omega R^\omega = 0$. (13.1)

This follows immediately from the variation of $YM(\omega)$; $\delta^\omega R^\omega = 0$ and the Euler equations, always $d^\omega R^\omega = 0$, and therefore $\delta^\omega R^\omega = 0$ and $\Delta^\omega R^\omega = 0$ are equivalent. $\delta^\omega R^\omega = 0$ are called the sourceless Yang-Mills equations. Assume now that $n = 4$. Then the Hodge $*$-operator $* : \Lambda^2 T*M \longrightarrow \Lambda^2 T*M$, $*(e_i \wedge e_j) = \operatorname{sign}\binom{1234}{ijkm} e_k \wedge e_m$, satisfies $*^2 = id$ and induces therefore an orthogonal splitting,

$$\Lambda^2 = \Lambda^2_+ \oplus \Lambda^2_-, \quad \Omega^2(\mathfrak{g}_E) = \Omega^2_+(\mathfrak{g}_E) \oplus \Omega^2(\mathfrak{g}_E),$$

where $\Lambda^2_\pm$ corresponds to the eigenvalue ± 1 of $*$. This implies a splitting $R^\omega = R^\omega_+ + R^\omega_-$, $*R^\omega_\pm = \pm R^\omega_\pm$, and $|R^\omega|^2_x = |R^\omega_+|^2 + |R^\omega_-|^2$. The connection ∇^ω is called self-dual or anti-self-dual if $R^\omega_- = 0$ or $R^\omega_+ = 0$, respectively. As is well known, the first Pontrjagin number of E

$$p_1(E) = \frac{1}{8\pi^2} \int\limits_{M^4} (|R^\omega_+|^2 - |R^\omega_-|^2) \,\mathrm{dvol}$$

is an integer on compact manifolds and independent of ω. As a result,

$$4\pi^2 |p_1(E)| \leq YM(\omega)$$

with equality if and only if $R^\omega_+ = 0$ or $R^\omega_- = 0$, i.e. $YM(\omega)$ has an absolute minimum for self-dual or anti-self-dual ω, and ω is therefore a Yang-Mills connection. Self-dual or anti-self-dual solutions to (13.1) are called instantons. Until now the main interest in solving (13.1) has been devoted to instantons.

Remark 13.1 For an oriented four-manifold, (M^4, g), the $*$-operator is conformally invariant; in particular, ω is self-dual with respect to g if and only if it is self-dual with respect to $e^\varphi \cdot g$.

Examples 13.2 1) Let $G = U(1)$. Then $P \longrightarrow M$ is an electromagnetic bundle. If (M^4, g) is a Lorentz manifold, then

$$d^\omega R^\omega = 0, \quad \delta^\omega R^\omega = 0$$

are Maxwell's equations.

2) Let $\varrho_n : U(1) \longrightarrow U(1) \subset Gl(1, \mathbb{C})$ be given by $\varrho_n(z) = z^n$, $z \in U(1)$. The sections of $C^\infty(P \times_{\varrho_n} \mathbb{C})$ may be considered as wave functions of scalar particles with charge = $n\cdot$ elementary charge.

3) Probably, the most important case is $G = SU(N)$, $\varrho = \mathrm{ad} : SU(N) \longrightarrow Aut[\mathfrak{s}u(N)]$. Every section of $C^\infty[\mathfrak{s}u(N)_P] \equiv \Omega^0[\mathfrak{s}u(N)_P]$ is a standard Higgs field. The Euler equations of the functional

$$\mathcal{A}(\omega, \varphi) = \frac{1}{2} \int (|R^\omega|^2 + |\nabla^\omega \varphi|^2 + \frac{\lambda}{4}(|\varphi| - m^2)) \,\mathrm{dvol}_x(g),$$

$$(d^\omega) * R^\omega = [\nabla^\omega \varphi, \varphi], \quad (d^\omega) * d^\omega \varphi = \frac{\lambda}{2}(|\varphi|^2 - m^2)\varphi$$

are called Yang-Mills-Higgs equations.

4) Consider $M^4 = S^4$, $G = SU(2) \cong Spin(3) \cong Sp(1) \cong S^3$, the quaternions $\mathbb{H}$, the Hopf bundle $P[S^4, Sp(1)] = S^7 \xrightarrow{Sp(1)} P^1(\mathbb{H}) = S^4$, $\varrho : Sp(1) \xrightarrow{id} Sp(1) \subset Gl(\mathbb{H})$ and the associated quaternion line bundle E. Then $-1 = \chi(E) = c_2(E) = -\frac{1}{2}p_1(E)$. Let $p = (0, \ldots, 0, 1) \in S^4$, $x : S^4 \setminus \{p\} \longrightarrow \mathbb{R}^4 \cong \mathbb{H}$ the stereographic projection, and $ds^2 = 4|dx|^2/(1 + |x|^2)^2$ the induced metric on $S^4 \setminus \{p\}$. According to the remark above, a connection is self-dual with respect to ds^2 if and only if it is self-dual with respect to the Euclidean metric $|dx|^2$. Restricted to $S^4 \setminus \{p\} \cong \mathbb{R}^4 \cong \mathbb{H}$, the bundle E is trivial, $E \cong \mathbb{R}^4 \times \mathbb{H} \longrightarrow \mathbb{R}^4 \cong \mathbb{H}$. Choose as reference connection the flat connection d and consider for $\lambda > 0$ the connection $\nabla^\lambda = d + A^\lambda$, where

$$A^\lambda = \Im \left(\frac{x \cdot d\overline{x}}{\lambda^2 + |x|^2} \right).$$

Here we write $x \in \mathbb{R}^4$ as $x = x_0 + x_1 \cdot i + x_2 \cdot j + x_3 \cdot k \in \mathbb{H}$. A^λ is an $\Im(\mathbb{H})$-valued $= \mathfrak{sp}(1)$-valued one-form. An easy calculation yields

$$R^\lambda \equiv R^{\nabla^\lambda} = dA^\lambda + [A^\lambda, A^\lambda] = \frac{\lambda^2 dx \wedge d\overline{x}}{\lambda^2 + |x|^2},$$

and R^λ is self-dual. With respect to ds^2 above,

$$|R^\lambda|^2_x = \frac{3}{2} \cdot \lambda^4 \left(\frac{1 + |x|^2}{\lambda^2 + |x|^2} \right)^4.$$

Moreover, $A^\lambda = (r^\lambda) * A^1$, $R^\lambda = (r^\lambda) * R^1$, where $r^\lambda : R^4 \longrightarrow R^4$ is defined by

$$r^\lambda(x) = \frac{1}{\lambda} \cdot x.$$

A^λ can be extended to the whole of S^4. The connection ∇^λ is called the 't Hooft solution. □

Now we sketch how to attack the main problem-finding solutions of (13.1) on open manifolds, or a little less, to find instantons and to describe the moduli space.

Clearly the problem is gauge-invariant, since $|R^{(\omega \cdot f)}|_x = |f^{-1} R^\omega f|_x = |R^\omega|_x$, i.e. the Yang-Mills functional is a functional on the orbit space $\mathcal{C}_P/\mathcal{G}_P$. Therefore, for a clear analytical approach, we must endow the configuration space = orbit space with an appropriate structure. This is a longer history. We refer to [49], [53] for details and list up here only the main steps.

1) Suppose that (M^n, g) and the bundle $(P(M,G), \omega)$ have bounded geometry. Then using formulas of [78], the total space $(P, g_\omega = g_{p,\omega} = \pi^* g_M + g_F)$ has bounded geometry, where $g_F(X,Y) = g_{\mathfrak{g}}(\omega(X), \omega(Y))$. g_ω is the Kaluza-Klein metric. The proof of this fact is presented in [49], p. 3955 – 3958.

2) If $\omega, \omega' \in \mathcal{C}_P(B_k)$, $k \geq r > \frac{n}{p} + 1$, belong to the same component in $\mathcal{C}_P^{p,r}(B_k)$ then $g_\omega, g_{\omega'}$ belong to the same component in $\mathcal{M}(P)^{p,r}(I, B_k)$.

3) According to step 1 and 2 and section 3, $\mathcal{D}^{p,r}(\mathrm{comp}(\omega)) := \mathcal{D}^{p,r}(P, \mathrm{comp}(g_\omega)) := \mathcal{D}^{p,mr}(P, g_\omega)$ is well defined.

4) Now one can show that

$$\mathcal{G}_P^{2,r}(\omega) = \{f \in \mathcal{D}^{2,r}(P, g_\omega) | f \text{ covers } \mathrm{id}_M \text{ and } f(ua) = f(u)a \text{ for all } u \in P \text{ and } a \in G\}$$

is a C^{1+k-r} submanifold of $\mathcal{D}^{2,r}(P, g_\omega)$.

5) One defines for $k \geq r > \frac{n}{2} + 1$

$$\tilde{\mathcal{G}}_P^{2,r}(\omega) = \Omega^{2,r}(G_P, \omega) = \{f \in \Omega^{2,r}(M, G_P, \omega) | f \text{ is a section }\}$$

and shows that $\tilde{\mathcal{G}}_P^{2,r}(\omega)$ is a C^{1+k-r} Hilbert Lie group.

6) Define for $k \geq r > \frac{n}{2} + 1$

$$\hat{\mathcal{G}}_P^{2,r}(\omega) = \{f \in \Omega^{2,r}(M, G_P, \omega) | f(u \cdot a) = a^{-1} f(u) a\}.$$

The map $\mathcal{G}_P^{2,r}(\omega) \ni f \longrightarrow f \in \hat{\mathcal{G}}_P^{2,r}(\omega)$ is an algebraic isomorphism and one can prove that it is an isomorphism of C^{1+k-r} Hilbert Lie groups.

7) One can identify $\mathcal{G}_P^{2,r}(\omega)$, $\tilde{\mathcal{G}}_P^{2,r}(\omega)$ and $\hat{\mathcal{G}}_P^{2,r}(\omega)$ as C^{1+k-r} Hilbert Lie groups. In [53] p. 273–279 there is another sequence of arguments for this, proving even that $\mathcal{G}_P^{2,r}(\omega)$ is a smooth Hilbert Lie group (see step 10).

8) Concerning the action of $\mathcal{G}_P^{2,r+1}(\omega)$, it is clear that we have an action

$$\mathrm{comp}(f) \times \mathrm{comp}(\omega) \longrightarrow \mathrm{comp}(f^* \omega).$$

Set

$$\mathcal{G}_{p,\mathrm{comp}(\omega)}^{2,r+1}(\omega) = \{f \in \mathcal{G}_P^{2,r+1}(\omega) | f \text{ leaves } \mathrm{comp}(\omega) \text{ fixed}\}.$$

$\mathcal{G}_{p,\mathrm{comp}(\omega)}^{2,r+1}(\omega)$ is a union of components of $\mathcal{G}_P^{2,r+1}(\omega)$ and therefore a closed submanifold, even a closed subgroup, i.e. we have an action

$$\mathcal{G}_{p,\mathrm{comp}(\omega)}^{2,r+1}(\omega) \times \mathrm{comp}^{2,r}(\omega) \longrightarrow \mathrm{comp}^{2,r}(\omega), \tag{13.2}$$

in particular this holds for $\omega \in \mathcal{C}_P^{2,r}(B_k, 2, f)$.

9) The action (13.2) is of class C^{k+r-1} and closed.

10) If we follow the approach of [53] then the corresponding action is even smooth. For this we describe the action as left action in the following form

$$f\omega := (f^{-1})_* \omega = \omega \circ d(f^{-1}).$$

In terms of $\hat{f} \in \hat{\mathcal{G}}$ this means

$$f\omega = Ad(\hat{f})\omega + (\hat{f}^{-1})^* \theta,$$

where θ denotes the Maurer-Cartan form of G. Define

$$\nabla^\omega \hat{f} := (\hat{f}^{-1})^* \theta \circ \mathrm{proj}_h^\omega = -dR_{\hat{f}}^{-1} \circ \hat{f}.$$

Then

$$f\omega = \omega + \nabla^\omega \hat{f}.$$

We define a new $\hat{\mathcal{G}}_{p,\mathrm{comp}(\omega)}^{2,r+1}(\omega)$ by

$$\hat{\mathcal{G}}_{p,\mathrm{comp}(\omega)}^{2,r+1}(\omega) = \{\hat{f} \in \hat{\mathcal{G}}_P^{2,r+1}(\omega) | \nabla^\omega \hat{f} \in \Omega^{1,2,r}(\mathfrak{g}_P, \nabla^\omega)\}$$

with a topology, roughly spoken, coming from Campell-Hausdorff series. We refer to [53], p. 273 – 279 for details.

11) This $\hat{\mathcal{G}}_{p,\mathrm{comp}(\omega)}^{2,r+1}(\omega)$ acts smoothly and closed on $\mathrm{comp}^{2,r}(\omega)$.

12) To establish good properties of the configuration space, one needs the second countability of $\hat{\mathcal{G}}^{2,r+1}_{p,\text{comp}(\omega)}(\omega)$ which is in general not satisfied since $\hat{\mathcal{G}}^{2,r+1}_{p,\text{comp}(\omega)}(\omega)$ can have uncountably many components. It is easy to see that each element of the centre of G generates one component. Therefore we restrict in the sequel ourselves to the subgroup of $\hat{\mathcal{G}}^{2,r+1}_{p,\text{comp}(\omega)}(\omega)$ which consists of the components of $\hat{\mathcal{G}}^{2,r+1}_{p,\text{comp}(\omega)}(\omega)$ generated by the centre of G if the centre is countable or which equals to the component of the identity if the centre is uncountable. In the case $G = SU(2)$ the centre consists of two elements, i.e. we get two components, each of them satisfies second countability. We denote the new subgroup of $\hat{\mathcal{G}}^{2,r+1}_{p,\text{comp}(\omega)}(\omega)$ again by the same symbol.

13) The final step towards the general structure of the configuration space consists in

Theorem 13.3 *Assume* (M^n, g) *open with* (I) *and* (B_k), $k-1 \geq r > \frac{n}{2}+2$, $\omega \in \mathcal{C}^{2,r}_P(B_k, 2, f)$ *and* $\inf \sigma_e(\Delta_0^\omega|_{(\ker \Delta_0^\omega)^\perp} > 0)$. *Then the configuration space*

$$\text{comp}^{2,r}(\omega)/\hat{\mathcal{G}}^{2,r+1}_{p,\text{comp}(\omega)}(\omega)$$

has the structure of a stratified space. The strata are labelled by the conjugacy classes of symmetry groups of connections.

We refer to [53], p. 283 – 285 for the proof. □

Now we want to show that $\mathcal{M}^{2,r}_{\text{comp}(\omega)}$ is nonempty. We restrict our attention to the case $n = 4$, $G = SU(2)$. Then the bundle $P(M^4, SU(2))$ is trivial (cf. [112]). Nevertheless the real L_2-Pontrjagin numbers

$$p_1(P, \text{comp}(\omega)) = \frac{1}{8\pi^2}\int(|R^\omega_+|^2 - |R^\omega_-|^2)\,\text{dvol}$$

can be nonzero.

Up to a large degree we imitate Taubes' existence proof of transplanting the 't Hooft solution from S^4 to M^4 and finding an instanton in the neighborhood of the transplanted solution. Nice versions of this proof are contained in Taubes' original paper ([116]), and in the books by Freed-Uhlenbeck and Lawson ([58] and [83]). The main assumption for a compact M^4 besides simply-connectedness was the positive-definiteness of the intersection form. This was used in the proof by inferring that there do not exist anti-self-dual harmonic two-forms. In the noncompact case one had to translate this to the positive-definiteness of the L_2-intersection form

$$(\varphi, \varphi') \longrightarrow \int_M \varphi \wedge \varphi',$$

φ, φ' being L_2-harmonic two-forms on (M^4, g). But this is not enough. We also need the condition that the bottom of the essential spectrum outside 0 of the Laplace operator $\Delta_2 = \Delta_2(M^4, g)$ on two-forms is greater than zero, $\inf \sigma_e(\Delta_2|_{(\ker \Delta_2)^\perp}) > 0$. More precisely, it would be sufficient to claim this for the Laplace operator acting on anti-self-dual two-forms, i.e. we assume that

$$\inf \sigma_e(\Delta_2|_{(\ker \Delta_2)^\perp}) > 0, \quad \bullet_{L_2} > 0, \tag{13.3}$$

or somewhat less

$$\inf \sigma_e(\Delta_{2,-}|_{(\ker \Delta_{2,-})^\perp}) > 0, \quad \bullet_{L_2} > 0.$$

An example for the condition $\inf \sigma_e(\Delta_2|_{(\ker \Delta_2)^\perp}) > 0$ is given by hyperbolic four-space $(M^4, g) = H^4_{-1}$. In this case the spectrum satisfies $\sigma(\Delta_2) = \{0\} \cup [2, \infty[$, where 0 belongs to the point spectrum and has infinite multiplicity. (13.3) has also a combinatorial formulation concerning as well the spectral gap as the positive definiteness of the intersection form. These properties are invariants of uniform triangulations with sufficiently small mesh. We refer to [19].

Fix now $p \in M^4$, $2\varepsilon < r_{\text{inj}}(M^4, g)$, $(x^1, \ldots, x^4)$ geodesic normal coordinates on $B_{2\varepsilon}(p)$ centered at p. Let $y : S^4 \setminus \{q_0\} \longrightarrow \mathbb{R}^4$ be the stereographic projection and $\varphi \in C^\infty([0, \infty[)$, $\varphi = 1$ on $[0, \varepsilon]$, $\varphi > 0$ on $[0, 2\varepsilon[$, $\varphi \equiv 0$ on $[2\varepsilon, \infty[$, $|d\varphi| = \frac{1}{\varepsilon} + 1$. φ is a smooth approximation of

$$\tilde{\varphi}(x) = \begin{cases} 1, & 0 \le x \le \varepsilon \\ \frac{2\varepsilon - x}{\varepsilon}, & \varepsilon \le x \le 2\varepsilon, \\ 0, & x > 2\varepsilon \end{cases} .$$

In what follows we use only first derivatives of φ and work with $\varphi = \tilde{\varphi}$. Define $F : M^4 \longrightarrow S^4$, $F \equiv \{q_0\}$ in $M^4 \setminus B_{2\varepsilon}(p)$, $F = y = \frac{x}{\varphi(x)}$ in $B_{2\varepsilon}(p)$. Consider the Hopf bundles $P_0 = S^7 \overset{SU(2)}{\longrightarrow} S^4$ of instanton number 1, the associated quaternionic line bundle E_0, the 't Hooft connection $\nabla_0^\lambda = d + A^\lambda$ with curvature

$$R_0^\lambda = \frac{\lambda^2 dy \wedge d\overline{y}}{(\lambda^2 + |y|^2)^2}$$

and the pull-back $P = F^* P_0$, $E = F^* E_0$, $\nabla^\lambda = F^* \nabla_0^\lambda \cdot \nabla^\lambda$ is flat in $M \setminus B_{2\varepsilon}(p)$ and in $B_{2\varepsilon}(p)$:

$$R^\lambda = \frac{\lambda^2}{\lambda^2 \varphi^2 + |x|^2} (\varphi^2 dx \wedge d\overline{x} - 2\varphi d\varphi \wedge \Im x d\overline{x}).$$

Since $d\varphi = 0$ in $B_\varepsilon(p)$ and $|x d\varphi| \approx 1$ in $B_{2\varepsilon}(p) \setminus B_\varepsilon(p)$, we obtain

$$R^\lambda = \frac{\lambda^2 dx \wedge d\overline{x}}{(\lambda^2 + |x|^2)} \quad \text{for } |x| \le \varepsilon$$

and $|R^\lambda| \le C \left(\frac{\lambda}{|x|^2} \right)$ for $\varepsilon \le |x| \le 2\varepsilon$, $c = c(\lambda)$.

The instanton number of the induced bundle P and $\text{comp}(\nabla^\lambda) \subset \mathcal{C}(B_k, 2, f)$ is

$$p_1(P, \text{comp}(\nabla^\lambda)) = \frac{1}{8\pi^2} \int (|R_+^{\nabla^\lambda}|^2 - |R_-^{\nabla^\lambda}|^2) \, \text{dvol}_x(g).$$

According to the properties of normal coordinates, $g_{ij}(0) = \delta_{ij}$, $ds^2 = |dx|^2 + O(|x|^2)$, for ε very small $|g_{ij}(x) - \delta_{ij}| \le \alpha |x|^2$, $|x| \le 2\varepsilon$, $\alpha \ll 1$.

The metric dx^2 corresponds to the Euclidean star operator $*_e$, and ds^2 to the Riemannian star operator $*$. Then $| *_e - * | = \alpha |x|^2$ for $|x| \le 2\varepsilon$. R is $*_e$-invariant over $B(p)$, which implies that

$$|R^\lambda - * R^\lambda| \le c' \frac{\lambda^2 |x|^2}{(\lambda^2 + |x|^2)^2} \le c', \quad |x| \le 2\varepsilon.$$

In the sequel we choose $\varepsilon = \sqrt{\lambda} \ll 1$. Assume now that $1 < p < \infty$. Then there exist constants c_1 and c_2 such that for all $\lambda \ll 1$

$$|R^\lambda|_p \le c_1 \cdot \lambda^{\frac{4}{p-2}}, \quad |R_-^\lambda|_p \le c_2 \cdot \lambda^{\frac{2}{p}}.$$

Consider now the operator

$$P^\nabla : \Omega_-^{2,2,4}(\mathfrak{g}_E, \nabla) \longrightarrow \Omega_-^{2,2,2}(\mathfrak{g}_E, \nabla), \quad P^\nabla := d_-^\nabla \delta^\nabla, \quad d_-^\nabla := \pi_- d^\nabla,$$

where we identify as always $\omega \longleftrightarrow \nabla(\omega) = \nabla$, $d^\omega = d^\nabla$, $d_-^\omega = d_-^\nabla$ etc. This is a linear, second order elliptic nonnegative operator. For any pair $u_1, u_2 \in \Omega_-^{2,2,4}(\mathfrak{g}_E)$, $\langle P^\nabla u_1, u_2\rangle = \langle \delta^\nabla u_1, \delta^\nabla u_2\rangle = \langle d^\nabla u_1, d^\nabla u_2\rangle$ since $*u_i = -u_i$. Define

$$\mu(\nabla) := \mu_1(\nabla) := \inf_{u \in \Omega_-^{2,2,4}(\mathfrak{g}_E, \nabla)} \frac{|\delta^\nabla u|_2^2}{|u|_2^2},$$

where here and in the sequel $|\,|_p \equiv |\,|_{L_p}$. Replacing $\Omega_-^{2,2,4}(\mathfrak{g}_E, \nabla)$ by $\Omega_-^{2,2,4}(M \times \mathbb{R})$ and P^∇ by $\Delta_{2,-}$ we infer from assumption (13.3) that there are no L_2-harmonic anti-self-dual two-forms. This implies

$$\mu := \inf_{u \in \Omega_-^{2,2,4}(M \times \mathbb{R})} \frac{|\delta u|_2^2}{|u|_2^2} > 0. \tag{13.4}$$

If $\mu = 0$ then we would obtain a Weyl sequence $(u_\nu)_\nu$ for $\lambda = 0$, $(\Delta_{2,-} - 0)u_\nu \longrightarrow 0$, $0 \in \sigma_e(\Delta_{2,-})$. According to the spectral gap, $0 \in \sigma_p(\Delta_{2,-})$ which contradicts $\ker(\Delta_{2,-}) = 0$.

Our first task is to prove a similar inequality for $P^\lambda = P^{\nabla^\lambda}$ and $\mu(\lambda) \equiv \mu(\nabla^\lambda)$.

Lemma 13.4 *More precisely, there exists a constant $\mu_0 > 0$ such that for all λ sufficiently small*

$$\mu(\lambda) \ge \mu_0. \tag{13.5}$$

and c_1, c_2 such that for $\lambda \ll 1$

$$|R^\lambda|_p \le c_1 \lambda^{\frac{4}{p-2}}, \quad |R_-^\lambda|_p \le c_2 \lambda^{\frac{p}{2}}.$$

We refer to [49], [83] for the proof.

Now we come to the heart of the existence proof. Fixing a smooth connection $\nabla \in \mathcal{C}_P(B_4, 2, f)$, we want to find a connection $\nabla' = \nabla + A$ so that $R_-^{\nabla'} = 0$, i.e. we want to find $A \in \Omega^{1,2,4}(\mathfrak{g}_E, \nabla)$ satisfying

$$R_-^\nabla + d_-^\nabla A + [A, A]_- = 0 \tag{13.6}$$

(we put the factor $\frac{1}{2}$ into the definition of $[\,,\,]$). We make the ansatz $A = \delta^\nabla u$, $u \in \Omega^{2,2,5}(\mathfrak{g}_E, \nabla)$, which implies that

$$d_-^\nabla \delta^\nabla u = -R_-^\nabla - [\delta^\nabla u, \delta^\nabla u]_-. \tag{13.7}$$

Equation (13.7) will be solved by an iterative scheme,

$$d_-^\nabla \delta^\nabla U_k = b_k, \quad k = 1, 2, 3, \ldots, \tag{13.8}$$

where $b_1 = -R^\nabla$,

$$b_k = -2\left[\sum_{j=1}^{k-2} \delta^\nabla U_j, \delta^\nabla U_{k-1}\right]_- - [\delta^\nabla U_{k-1}, \delta^\nabla U_{k-1}]_- \quad \text{for } k > 1. \tag{13.9}$$

Considering

$$u_m = \sum_{k=1}^{m} U_k, \tag{13.10}$$

we conclude that

$$d_-^\nabla \delta^\nabla u_m = -R_-^\nabla - [\delta^\nabla u_{m-1}, \delta^\nabla u_{m-1}]_-. \tag{13.11}$$

Formally we obtain for $m \longrightarrow \infty$ a solution of (13.7). But this argument has to be made more precise, i.e. we have to ensure a reasonable convergence for the series.

Remark 13.5 We note that $\nabla \in \mathcal{C}_P(B_4, f)$ implies that $R^\nabla \in \Omega^{2,p,0}(\mathfrak{g}_E, \nabla)$ for all $p \geq 2$.

A key role is played by the following theorem, which ensures the solvability of the starting equation.

Theorem 13.6 *Assume that $\mu(\nabla) > 0$. Then for any smooth $b \in \Omega^{2,2,2}(\mathfrak{g}_E, \nabla)$, there is a unique smooth solution to the equation*

$$d_-^\nabla \delta^\nabla U = b. \tag{13.12}$$

This solution satisfies, with

$$\beta^2 = \frac{1}{\mu(\nabla)}[1 + \mu(\nabla) + |R_-^\nabla|_3^3], \tag{13.13}$$

the estimates

$$\begin{aligned} |\delta^\nabla U|_2 &\leq c \cdot \beta \cdot |b|_{4/3}, && (13.14)\\ |\delta^\nabla U|_{2,1} &\leq c[|b|_2 + \beta|b|_{4/3}(1 + |R^\nabla|_4)], && (13.15)\\ |\delta^\nabla U|_4 &\leq c[|b|_2 + \beta|b|_{4/3}(1 + |R^\nabla|_4)], && (13.16)\end{aligned}$$

where c is a constant depending only on (M^4, g). $|b|_{4/3}$ is well defined, since by assumption $b \in \Omega_-^{2,2,2}(\mathfrak{g}_E, \nabla)$ and according to the Sobolev embedding theorem for manifolds of bounded geometry. □

Set

$$\varepsilon(\nabla) = |R_-^\nabla|_2 + \beta \cdot |R_-^\nabla|_{4/3}(1 + |R^\nabla|_4). \tag{13.17}$$

For ε sufficiently small the iterative scheme will work, as the following theorem shows. Let c be the above constant, $\varepsilon \leq \min\{1/2c\beta, 1/2c, 1/2\}$ and $\varepsilon(\nabla) \leq \varepsilon^5$. Then there exist smooth solutions U_k, $k = 1$, to the rekursive sequence of equations (13.8) and (13.9), which satisfy

$$|\delta^\nabla U_k|_2 \leq \varepsilon^{k+4}(1 + |R^\nabla|_4)^{-1}, \tag{13.18}$$

$$\begin{aligned} |\delta^\nabla U_k|_{2,1} &\le \varepsilon^{k+4}, & (13.19)\\ |\delta^\nabla U_k|_4 &\le \varepsilon^{k+4}. & (13.20)\end{aligned}$$

We start the proof with $k = 1$. For $k = 1$, (13.18) – (13.20) are an immediate consequence of (13.14) – (13.16). Assume the assertion for $k - 1$. Then

$$\begin{aligned} |b_k|_1 &\le 4\sum_{j=1}^{k-2} |\delta^\nabla U_j|_2 \cdot |\delta^\nabla U_{k-1}|_2, \\ |b_k|_2 &\le 4\sum_{j=1}^{k-2} |\delta^\nabla U_j|_4 \cdot |\delta^\nabla U_{k-1}|_4. \end{aligned} \tag{13.21}$$

By the induction hypothesis

$$\begin{aligned} |b_k|_1 &\le \frac{4}{(1+|R^\nabla|_4)^2} \cdot \varepsilon^{k+3} \sum_{j=1}^{k-2} \varepsilon^{j+4} \le \frac{\varepsilon^{k+5}}{(1+|R^\nabla|_4)^2}, \\ |b_k|_2 &\le \varepsilon^{k+5}. \end{aligned} \tag{13.22}$$

We infer from Hölder's inequality that

$$|b_k|_{4/3} \le |b_k|_1^{1/2} \cdot |b_k|_2^{1/2} \le \frac{\varepsilon^{5+k}}{1+|R^\nabla|_4}, \tag{13.23}$$

plug that into (13.14) – (13.16) and obtain the desired result.

Now we show that $u = \sum_k U_k$ sums up to a solution of the self-duality equations. Assume that $\varepsilon(\nabla) \le \varepsilon^5$ (ε as above). Then the sequence $(u_m)_m = (\sum_{k=1}^m U_k)_m$ converges to a smooth solution of (13.7), i.e. $\tilde\nabla = \nabla + \delta^\nabla u$ is self-dual. Moreover,

$$|\tilde\nabla - \nabla|_{2,1} + |\tilde\nabla - \nabla|_4 \le 2\varepsilon(\nabla). \tag{13.24}$$

To prove this, we need the following lemma.

Lemma 13.7 *There is a constant $c > 0$, independent of ∇, such that if $\mu(\nabla) > 0$ then*

$$\frac{1}{c\beta(\nabla)} |\nu|_{2,1} \le |\delta^\nabla \nu|_2 \le c|\nu|_{2,1} \tag{13.25}$$

for all $\nu \in \Omega_-^{2,2,1}(\mathfrak{g}_E, \nabla)$.

With (13.25) we can replace the good estimate (13.18) for $|\delta U_k|_2$ by an estimate for $|U_k|_2$, and it is now evident that the sequence $(u_m)_m = (\sum_{k=1}^m U_k)_m$ is Cauchy with respect to $|\,|_{2,1}$. Hence it converges to $u = \lim_m u_m$ in $|\,|_{2,1}$. We still have to show that u satisfies the required equation and that $u \in C^\infty$. For this consider the sequence

$$\sigma_m := d_-^\nabla \delta^\nabla u_m + [\delta^\nabla u_m, \delta^\nabla u_m]_- + R_-^\nabla.$$

We know that $d_-^\nabla \delta^\nabla u_m \longrightarrow d_-^\nabla \delta^\nabla u$ in $|\,|_2$. Moreover, according to (13.20), δu_m is Cauchy in the $|\,|_4$ norm. By Hölder's inequality, $[\delta^\nabla u_m, \delta^\nabla u_m]_- \longrightarrow [\delta^\nabla u, \delta^\nabla u]_-$ with respect to $|\,|_2$. By construction of the U_k, $\tilde{\sigma}_m = d_-^\nabla \delta^\nabla u_m + [\delta^\nabla u_{m-1}, \delta^\nabla u_{m-1}]_- + R_-^\nabla = 0$ for each m. According to $|\tilde{\sigma}_m - \sigma_m|_2 = O(\varepsilon^m)$, we have

$$d_-^\nabla \delta^\nabla u + [\delta^\nabla u, \delta^\nabla u_-] + R_-^\nabla = \lim_m \sigma_m = \lim \tilde{\sigma}_m = 0,$$

i.e. u is a weak solution to (13.7). Standard elliptic regularity theory now shows that $u \in C^\infty$.

Two points are still open, namely (13.25) and the existence of a starting solution U satisfying (13.14) – (13.16). For the proof of 13.7, see [83] (pp. 82, 83). There is nothing to add and nothing to change. Consider the equation $P^\nabla U = b$, $P^\nabla = d_-^\nabla \delta^\nabla = (\delta^\nabla)^* \delta^\nabla$, a linear self-adjoint elliptic operator which is strictly positive since $\mu(\nabla) > 0$. Since (M^4, g) and (P, ∇) have bounded geometry, there exists a good heat kernel of P^∇ and moreover a good Green's kernel, because $\mu(\nabla) > 0$, by theorem 6.55. This implies the existence of U. Standard elliptic regularity implies that $U \in C^\infty$. For the estimates (13.14) – (13.16) see [83] (pp. 83, 84).

All in all, we sketched the proof of the following main theorem.

Theorem 13.8 *Let (M^4, g) be open, oriented, complete with $(B_k(M))$, (I), $k > 3$, $\inf \sigma_e(\Delta_{2,-}|_{(\ker \Delta_{2,-})^\perp}) > 0$, $\bullet_{L_2} > 0$, $G = SU(2)$, $P = F^* P_0$. Then $\mathrm{comp}(\nabla^\lambda) \subset \mathcal{C}_P(B_k, f)$ contains a self-dual connection.*

Proof According to (13.5) there is a constant $\beta_0 > 0$ such that $1/\beta_0 \le 1/\beta(\nabla^\lambda)$ for all $\lambda \ll 1$, and $\varepsilon(\nabla^\lambda) \le c_0 \cdot \lambda^{1/2}$ for some constant c_0. Consequently, the hypotheses for the validity of (13.18) – (13.20) are satisfied for all $\lambda \ll 1$. □

Now we want to support theorem 13.8 by

Example 13.9 A first class of examples comes immediately from the compact case. Let (M_0^4, g_0) be a compact oriented Riemannian manifold with boundary ∂M_0^4, $H^1(\partial M_0^4) = H^2(\partial M_0^4) = 0$ and positive definite intersection form. Set $(M^4, g) = (M_0^4 \cup \partial M_0^4 \times [0, \infty[, g)$, where $g|_{M_0^4} = g_0$ and $g|_{\partial M_0^4 \times [1,\infty[} = dr^2 + g_0|_{M_0^4}$, i.e. (M^4, g) is a manifold with cylindrical ends. Then $\inf \sigma_e(\Delta_2(M^4, g)) = \inf \sigma(\Delta_2(\partial M_0^4)) > 0$, $\mathcal{H}^{2,2,\{d\}}(M^4) = H^{2,2,\{d\}}(M^4) \cong \mathrm{Im}\,(H^2(M_0^4, \partial M_0^4) \longrightarrow H^2(M_0^4)) \cong H^2(M_0^4)$ and the L_2-intersection form of M^4 is equivalent to that of M_0^4 (cf. [6]). Take e.g. a closed manifold $(\overline{M}^4, \overline{g})$ with positive intersection form, delete from it small disjoint open discs $B_1, \ldots, B_s$ and set $M_0^4 = \overline{M}^4 \setminus \bigcup_{\sigma=1}^{s} B_\sigma$. □

14 Fluid dynamics

As we indicated in the introduction and partially presented in section 7, the transform of an PDE into an ODE on an infinite-dimensional manifold is one of the standard techniques in solving PDE's. A striking example for this approach are several important equations of fluid dynamics. According to our knowledge, V. J. Arnold was the first who established such an approach. This has been essentially elaborated by Ebin/Marsden in the seminal paper [40]. But they restricted themselves to the compact case. Now we are able to extend their approach to the open case. This is possible since now we have a solid theory of diffeomorphism groups and their geometry for open manifolds. Such a theory was for

Ebin/Marsden 1970 not available. To establish such a theory took us a couple of years and was possible since we developed the theory of completed diffeomorphism groups also for underlying open manifolds. The key is the identity component of the completed group $\mathcal{D}^{\infty,r}_{\mu,0}$.

From now on we assume $p = 2$, (M^n, g) with (I) and (B_k), $k \geq m \geq r > \frac{n}{2} + 1$ and write $\mathcal{D}^r \equiv \mathcal{D}^{2,r}$ and $\mathcal{D}^r_0$ shall denote the component of the identity, $\mathcal{D}^r_0 = \mathrm{comp}(\mathrm{id}) \subset \mathcal{D}^r$. Let $\omega \in {}^{b,m}\Omega^q(M)$ be a C^m-bounded q-form on M and denote $D^r_\omega = \{f \in D^r | f^*\omega = \omega\}$. Here we assume $r \leq m \leq k$. Then $\mathcal{D}^r_{\omega,0} = \mathcal{D}^r_\omega \cap \mathcal{D}^r_0$ is a group. Assume additionaly that ω is closed. We want to show that $\mathcal{D}^r_{\omega,0}$ is a good submanifold of $\mathcal{D}^r_0$. The most important examples are $\omega =$ volume form μ or $\omega =$ symplectic form on a symplectic manifold. It is quite natural and helpful, in particular for integration theory of Hamiltonian systems, to assume on open symplectic manifolds that ω is adapted to a metric of bounded geometry by requiring $\omega \in {}^{b,m}\Omega^2(M)$, $2 \leq m \leq k$. This ensures the completeness of Hamiltonian vector fields, the transitivity of the flow and hence the existence of Liouville tori in the case of $\frac{1}{2} dim M$ independent integrals in involution.

We say $\omega \in {}^{b,0}\Omega^q$ is *nondegenerate* if for every $\varphi \in {}^{b,0}\Omega^{q-1}$ there exists a unique C^0-vector field X such that

$$i_X\omega \equiv \omega(X, .) = \varphi(.). \tag{14.1}$$

ω is said to be *strongly nondegenerate* if in addition

$$\inf_{x \in M} |\omega|^2_x > 0. \tag{14.2}$$

There is only a small choice. (14.1) means that X establishes for any $x \in M$ an isomorphism between T_xM and $\Lambda^{q-1}T^*_xM$. Therefore φ must be an $(n-1)$ or 1-form, $q = n$ or $q = 2$. Hence, ω must be a volume form μ or symplectic form ω satisfying (14.2) if we additionally claim the closedness. The generalization in comparison with Ebin-Marsden, [40] consists in allowing arbitrary volume forms, not only such defined by a Riemannian metric. The restriction in comparison with Ebin-Marsden [40] consists in condition (14.2).

We indicate the main steps of our approach and refer to [54] for the proofs.

Theorem 14.1 *Assume (M^n, g) with (I) and (B_k), $k \geq m \geq r > \frac{n}{2} + 1$. Let $\omega \in {}^{b,m}\Omega^q$ be closed and $f \in D^r_0(M^n, g)$. Then*

$$f^*\omega - \omega \in d\Omega^{q-1,r-1}.$$

□

Corollary 14.2 *Let $f \in \mathcal{D}^r \cap C^{\infty,r}(M, M)$ and $f' \in \mathrm{comp}(f)$, $\omega \in {}^{b,m}\Omega^q$ closed, $k \geq m \geq r > \frac{n}{2} + 1$. Then*

$$(f')^*\omega - f^*\omega \in d\Omega^{q-1,r-1}(f^*\nabla).$$

□

Lemma 14.3 *Assume $k \geq m \geq r > \frac{n}{2} + 1$ and $\mu \in {}^{b,m}\Omega^n$ a volume form with $\inf_{x\in M} |\mu|^2_x > 0$. Then $\Phi : X \longrightarrow i_X\mu$ defines an isomorphism between $\Omega^{r-1}(TM)$ and $\Omega^{q-1,r-1}$.* □

Completely parallel to lemma 14.3, we state

Lemma 14.4 *Assume $k \geq m \geq r > \frac{n}{2} + 1$ and $\omega \in {}^{b,m}\Omega^2$ a symplectic form with $\inf_{x \in M} |\omega|_x^2 > 0$. Then $X \mapsto i_X\omega$ establishes an isomorphism between $\Omega^{r-1}(TM)$ and $\Omega^{1,r-1}$.* □

Let Δ_q be the Laplace operator acting on q-forms, $\sigma_e(\Delta_q)$ its essential spectrum and $\sigma_e(\Delta_q|_{(\ker \Delta_q)^\perp})$ the essential spectrum of Δ_q restricted to the orthogonal complement of its kernel, and $\inf \sigma_e(\Delta_q|_{(\ker \Delta_q)^\perp})$ its $g.l.b$. Now we can state our first main theorem.

Theorem 14.5 *Assume (M^n, g) with (I) and (B_k), $k \geq m \geq r > \frac{n}{2} + 1$ and $\omega \in {}^{b,m}\Omega^q$, $q = n$ or $q = 2$, a closed strongly nondegenerate form with $\inf \sigma_e(\Delta_1|_{(\ker \Delta_1)^\perp}) > 0$. Then the group $\mathcal{D}^r_{\omega,0} = \mathcal{D}^r_0 \cap \mathcal{D}^r_\omega$ is a C^{k-r+1} submanifold of $\mathcal{D}^r_0$.*

Proof Consider the map $\psi : \mathcal{D}^r_0 \longrightarrow [\omega]_{r-2} := [\omega + \overline{d\Omega^{q-1,r-1}}]$, $\psi(f) := f^*\omega$. According to our spectral assumption we have $d\Omega^{q-1,r-1} = \overline{d\Omega^{q-1,r-1}}$, where we have taken the closure in $\Omega^{q,r-2}, q = 2$ or n. Hence, the affine space $[\omega]_{r-2}$ is a smooth Hilbert-manifold. We conclude from theorem 14.1 that ψ is well defined. $\mathcal{D}^r_0$ has differentiability class C^{k-r+1}. A straightforward calculation shows that ψ has differentiability class C^{k-r+1}. The map $\psi_{*,id}$ can easily be calculated

$$\psi_{*,\mathrm{id}}(X) = \lim_{t \longrightarrow 0} \frac{(\exp tX)^*\omega - \omega}{t} = L_X\omega = di_X\omega + i_X d\omega = di_X\omega,$$

$X \in T_{id}\mathcal{D}^r_0 = \Omega^r(TM)$. We conclude from lemma 14.3, 14.4 that $X \mapsto i_X\omega$ maps into $\Omega^{q-1,r}$ and conclude once again from our spectral assumption the closedness of $d\Omega^{q-1,r}$ in $\Omega^{q,r-2}$, i.e. the map $\psi_{*,id} : T_{\mathrm{id}}\mathcal{D}^r_0 \longrightarrow \overline{d\Omega^{q-1,r-1}} = T_\omega[\omega]_{r-2}$ is surjective. A simple shifting argument using $T_f\mathcal{D}^r_0 = \Omega^r(f^*TM, f^*\nabla)$ and the fact that $\Omega^r(TM)$ is mapped homeomorphically to $\Omega^r(f^*TM, f^*\nabla)$ by means of f and $\psi_{*,f}(X) = f^*(L_{X\circ f^{-1}}\omega)$ shows that ψ_* is surjective everywhere, i. e. ψ is a submersion. Hence $\mathcal{D}^r_{\omega,0} = \psi^{-1}(0)$ is a C^{k-r+1}-submanifold of $\mathcal{D}^r_0$. □

We defined for $C^{\infty,m}(M,N)$ a uniform structure $\mathcal{U}^{p,r}$. Consider now $C^{\infty,\infty}(M,N) = \bigcap_m C^{\infty,m}(M,N)$. Then we have an inclusion $i : C^{\infty,\infty}(M,N) \to C^{\infty,m}(M,N)$ and $i \times i : C^{\infty,\infty}(M,N) \times C^{\infty,\infty}(M,N) \to C^{\infty,m}(M,N) \times C^{\infty,m}(M,N)$ hence a well defined uniform structure $\mathcal{U}^{\infty,p,r} = (i \times i)^{-1}\mathcal{U}^{p,r}$. After completion we obtain once again the manifolds of mappings $\Omega^{\infty,p,r}(M,N)$, where $f \in \Omega^{\infty,p,r}(M,N)$ if and only if for every $\varepsilon > 0$ there exists an $\tilde{f} \in C^{\infty,\infty}(M,N)$ and a $Y \in \Omega^{p,r}(\tilde{f}^*T^*N)$ such that $f = \exp Y$ and $|Y|_{p,r} \leq \varepsilon$. Moreover, each connected component of $\Omega^{\infty,p,r}(M,N)$ is a Banach manifold and $T_f\Omega^{\infty,p,r}(M,N) = \Omega^{p,r}(f^*TN)$. As above we set

$$\begin{aligned}\mathcal{D}^{\infty,p,r}(M) \quad = \quad & \{f \in \Omega^{\infty,p,r}(M,M) | f \text{ is injective, subjective,} \\ & \text{preserves orientation and } |\lambda|_{min}(df) > 0\}.\end{aligned}$$

We assume $p = 2$ and write $\Omega^{\infty,r}(M,N) \equiv \Omega^{\infty,2,r}(M,N)$ and $\mathcal{D}^{\infty,r}(M) \equiv \mathcal{D}^{\infty,2,r}(M)$. The only difference between our former construction is the fact that the spaces $\Omega^{\infty,r}$ and $\mathcal{D}^{\infty,r}$ are based on maps which are bounded up to arbitrarily high order. For compact manifolds we have $C^\infty(M,N) = C^{\infty,r}(M,N) = C^{\infty,\infty}(M,N)$, $\Omega^{\infty,r}(M,N) = \Omega^r(M,N)$ and $\mathcal{D}^{\infty,r}(M) = \mathcal{D}^r(M)$ for all r. For open manifolds we have strong inclusions $C^{\infty,\infty} \subset C^{\infty,r}$ and $\mathcal{D}^{\infty,\infty} \subset \mathcal{D}^r$. It is very easy to construct a diffeomorphism $f \in C^{\infty,1}(\mathbf{R},\mathbf{R})$ such that $f \notin C^{\infty,2}(\mathbf{R},\mathbf{R})$. This supports the conjecture that the inclusion $\mathcal{D}^{r+s} \hookrightarrow \mathcal{D}^r$, $s \geq 1$ is not dense. It is very probable that the density fails as a careful analysis of our example shows. Here we restrict ourselves to $\mathcal{D}^{\infty,r}$. This space has the advantage that $\mathcal{D}^{\infty,r+s}$ is densely and continuously embedded into $\mathcal{D}^{\infty,r}$. This follows easily from the corresponding properties of Sobolev spaces. Quite analogously we define the space $\mathcal{D}^{\infty,r}_\omega$. The components of the identity have special nice properties:

Lemma 14.6 *Assume the conditions for defining $\mathcal{D}^r$ and $\mathcal{D}^r_\omega$. Then*

$$\mathcal{D}^{\infty,r}_0 = \mathcal{D}^r_0, \quad \mathcal{D}^{\infty,r}_{\omega,0} = \mathcal{D}^r_{\omega,0}.$$

Most of the interesting diffeomorphism groups are endowed with a natural Riemannian metric which is usually a weak one. For further applications, it is useful and important to know the corresponding Riemannian geometry, i.e. the curvature and the geodesics. In this section, we study the general diffeomorphism groups $\mathcal{D}^{\infty,r} \subset \mathcal{D}^r$ and the subgroups of form preserving diffeomorphisms $\mathcal{D}^{\infty,r}_\omega \subset \mathcal{D}^r_\omega$. Later on, we restrict ourselves to the component of the identity. There are several papers where this geometry already has been studied. But in all of these papers, only the case of compact manifolds M^n has been studied. Moreover, serious analytical problems arising in this investigation mostly have been suppressed. Completions have not been considered. In the noncompact case, one has to solve these difficulties. Without that, everything becomes wrong. If we start with $\mathcal{D}^r$, we would have to consider $\mathcal{D}^{r+s}$ with $k + 1 - (r + s) >$s, $s \geq 2$, $r > \frac{n}{2} + 1$ in order to apply some versions of the ω-lemma. Finally, one obtains curvature formulas for tangent vectors tangent to $\mathcal{D}^r$ at $f \in \mathcal{D}^{r+s}$. Then one would like to extend these formulas to all $f \in \mathcal{D}^r$. But this is impossible since it is not yet clear that $\mathcal{D}^{r+s} \subset \mathcal{D}^r$ is dense. Therefore, one has, at least at the final stage, to restrict everything to $\mathcal{D}^{\infty,r+s}$, $\mathcal{D}^{\infty,r}$. Then we have $\mathcal{D}^{\infty,r+s} \subset \mathcal{D}^{\infty,r}$ densely. Moreover, in this case $\{\mathcal{D}^{\infty,\infty}, \mathcal{D}^{\infty,r}, r > \frac{n}{2}+1\}$ is an ILH-group if $k = \infty$. In our considerations, we consider a finite tower of this ILH-group. As a matter of fact, the final formulas e.g. for sectional curvature coincide with those of the compact case, as it should be.

We present here a rapid, very short presentation. All details would exceed the framework of such a paper. Calculations which are parallel to the compact case are completely suppressed. Many of them are contained in [111], [112] and [113].

We start by defining the Levi-Civita connection for $\mathcal{D}^r$. Later on, we restrict everything to $\mathcal{D}^{\infty,r}$. Let M^n be a manifold. Then a connection is given by a field of horizontal subspaces of TTM or a covariant derivative ∇ or a connector map $K : TTM \longrightarrow TM$, respectively. They are related by $\nabla_X Y = K(Y_*(X))$ and K = projection onto the vertical subspaces along the horizontal ones. Locally, K can be expressed explicitly by the Christoffel symbols Γ^k_{ij}. A connector K can be independently characterized by the following properties. K is a map $T^2M \longrightarrow TM$ which satisfies the following conditions.

For all $X \in TM, K : T_X(TM) \longrightarrow T_{\pi(X)}M$ is linear.
For all $x \in M, K : T(T_xM) \longrightarrow T_xM$ is linear.
For all $X \in TM, H_X := \ker K|_{T_X TM}$ is horizontal.

Assume now (M^n, g) is open with (I) and (B_k), $k > \frac{n}{2} + 2$. Let ∇ denote the Levi-Civita connection with associated connector map K. The tangent manifold TM can be endowed with a canonical metric, the so called Sasaki metric given by:

$$g_{TM}(X, Y) = g_M(\pi_* X, \pi_* Y) + g_M(KX, KY), \quad X, Y \in T_Z TM.$$

Lemma 14.7 *The Sasaki metric g_{TM} satisfies (I) and (B_{k-1}).*

Proof (B_{k-1}) follows immediately from [14], p.130, equation (2). We proved in [49] (I) for principal fiber bundles $P(M,G)$ with respect to the Kaluza-Klein metric

$$g_\omega(X,Y) = g_M(\pi_* X, \pi_* Y) + \langle \omega(X), \omega(Y)\rangle_{\mathbf{g}},$$

g_M with (I) and $(B_k), \omega$ with (B_k). A similar proof can also be performed in our case here. □

Corollary 14.8 *For $k-2 \geq r > \frac{n}{2}+1$*

$$\Omega^r(M,TM) \text{ and } \Omega^r(TM,T^2M),$$

$$\Omega^{\infty,r}(M,TM) \text{ and } \Omega^{\infty,r}(M,T^2M)$$

are well defined. □

Proposition 14.9 *Assume $k-1 \geq r > \frac{n}{2}+1$. Then there exists a $C^{k-1-r+1} = C^{k-r}$ embedding*

$$\varphi : T\Omega^r(M,M) \longrightarrow \Omega^r(M,TM).$$

Corollary 14.10 *Assume $k-2 \geq r > \frac{n}{2}+1$. Then there exists a canonical C^{k-1-r}-embedding*

$$\psi : T^2\Omega^r(M,M) \longrightarrow \Omega^r(M,T^2M).$$

□

Remark 14.11 φ is not surjective. Let $(M^n,g) = (\mathbf{R}^n, g_{\mathbf{R}^n})$, $T\mathbf{R}^n = \mathbf{R}^n \times \mathbf{R}^n$ and e be the section $x \in \mathbf{R}^n \mapsto (x,(1,0,\ldots,0)) \in T_x\mathbf{R}^n$. Then $e \in C^{\infty,r}(\mathbf{R}^n, T\mathbf{R}^n)$, for arbitrary r, in particular $e \in \Omega^r(\mathbf{R}^n, T\mathbf{R}^n)$. If $e \in \operatorname{Im}\varphi$ then e should be a Sobolev vector field covering $id_{\mathbf{R}^n}$. This is impossible since e is not even square integrable. □

Proposition 14.12 *$T\mathcal{D}^r$ is C^{k-r}-diffeomorphic to*

$$\{X_f \in \operatorname{Im}\varphi \subset \Omega^r(M,TM) | X_f \text{ covers } f \in \mathcal{D}^r(M)\}.$$

$T^2\mathcal{D}^r$ is C^{k-r-1}-diffeomorphic to

$$\{V \in \operatorname{Im}\psi \subset \Omega^r(M,T^2M) | V \text{ covers } X_f\}.$$

Proof The assertion follows immediately from the openness of $\mathcal{D}^r \subset \Omega^r(M,M)$, $T\mathcal{D}^r \subset T\Omega^r(M,M)$, $T^2\mathcal{D}^r \subset T^2\Omega^r(M,M)$. □

Remark 14.13 All assertions above remain true if we replace $\Omega^r(M,N)$ by $\Omega^{\infty,r}(M,N)$, and $\mathcal{D}^r(M)$ by $\mathcal{D}^{\infty,r}(M)$. □

In the sequel, we identify $T\mathcal{D}^r$ or $T^2\mathcal{D}^r$ with their corresponding images.

Now we would like to define for $V \in T^2\mathcal{D}^r \subset \Omega^r(M,T^2M)$

$$\bar{K}(V) := K \circ V$$

and a covariant derivative $\bar{\nabla}$ by

$$\bar{\nabla}_{\bar{X}}\bar{Y} := \bar{K}(\bar{Y}_x(\bar{X})).$$

But, there arise several serious difficulties. $\bar{K}$ is defined by left multiplication by K. To obtain a certain differentiability class of $\bar{K}$, we should have an ω-lemma. In the compact case, all considered manifolds are Hilbert manifolds of class C^∞, an α - and ω - lemma are very easily established. $\Omega^r(M,M)$ is of class C^{k-r+1} , $\Omega^r(M,TM)$ of class $C^{k-1-r+1} = C^{k-r}$, and $\Omega^r(M,T^2M)$ of class C^{k-r-1}. The same holds for $\Omega^{\infty,r}(M,M)$, $\Omega^{\infty,r}(M,TM)$, and $\Omega^{\infty,r}(M,T^2M)$, assuming $k-2 \geq r > \frac{n}{2}+1$. Therefore, considering the C^s-property of $\omega_K, \omega_K(V) = K \circ V$, this does not make sense if $k-r-1 < s$. It does make sense for $k-1-(r+s) > s \geq 1$.

Proposition 14.14 *Assume $k-1-(r+s) > s \geq 1$, $r > \frac{n}{2}+1$. Then $\bar{K}(V) := K \circ V = \omega_K(V)$ is a C^s-map $\omega_K : T^2\mathcal{D}^{r+s} \longrightarrow T\mathcal{D}^r$ or $\omega_K : T^2\mathcal{D}^{\infty,r+s} \longrightarrow T\mathcal{D}^{\infty,r}$, respectively.*

Corollary 14.15 *If $k=\infty$ then $\mathcal{D}^r$, $T\mathcal{D}^r$, and $T^2\mathcal{D}^r$ are smooth manifolds and $\bar{K}$ is an everywhere defined smooth connection map.*

This follows from proposition 14.14 and the fact that the Christoffel symbols and hence K are bounded of arbitrarily high order. □

Now let $\bar{X} \in C^2(T\mathcal{D}^{r+s})$ be a C^2-vector field. We say $\bar{X}$ is right invariant if

$$\bar{X}(f) = \bar{X}(id) \circ f$$

for all $f \in \mathcal{D}^{r+s}$. Defining $X := \bar{X}(id)$, we can write $\bar{X}(f) = X \circ f$.

Lemma 14.16 *Let $X_f \in T_f\mathcal{D}^{r+s} = \Omega^{r+s}(f^*M)$. Then $X_f \circ f^{-1}$ is a vector field along id_M and $X_f \circ f^{-1} \in \Omega^{r+s}(TM)$. In particular, $X_f \circ f^{-1}$ is a C^2-vector field if $r+s > \frac{n}{2}+2$.*

We refer to [54] for the simple proof. □

Now, we define a weak Riemannian metric on $\mathcal{D}^{r+s}$, $\mathcal{D}^{\infty,r+s}$ and establish that $\bar{\nabla}$ is the corresponding Levi-Civita connection. Set for $V, W \in T_f\mathcal{D}^{r+s} = \Omega^{r+s}(f^*TM)$

$$\begin{aligned} \langle V, W \rangle &= g_0(V,W) = \int_M (V_{f(x)}, W_{f(x)})_{f(x)} \,\mathrm{dvol}_x(g) = \\ &= \int_M (V,W) \circ f(x) \,\mathrm{dvol}_x(g). \end{aligned} \tag{14.3}$$

The integral (14.3) converges according to the definition of $T_f\mathcal{D}^{r+s}$. Using charts in $\mathcal{D}^{r+s}$, it is easy to see that g_0 is a weak C^s metric. Remember $k-(r+s)-1 > s$.

Remark 14.17 g_0 only gives the L_2 - topology of each tangent space. To obtain the actual, i. e. the Sobolev topology, we should work with the corresponding strong Riemannian metric.

□

Lemma 14.18 *Let $\bar{X}, \bar{Y}, \bar{Z}$ be right invariant C^1 vector fields on D^{r+s}. Then*

$$g_0(\bar{\nabla}_{\bar{Z}}\bar{X}, \bar{Y}) + g_0(\bar{\nabla}_{\bar{Z}}\bar{Y}, \bar{X}) = \bar{Z}g_0(\bar{X}, \bar{Y}).$$

This is lemma 11.12 from [39]. □

Lemma 14.19 *Let $\bar{X}, \bar{Y}, \bar{Z} \in C^2(T\mathcal{D}^{r+s})$, $f \in D^{r+s}$, $r > \frac{n}{2}+1$, $s \geq 2$. Then*

$$g_0(\bar{\nabla}_{\bar{Z}}\bar{X}, \bar{Y})(f) + g_0(\bar{\nabla}_{\bar{Z}}\bar{Y}, \bar{X})(f) = \bar{Z}g_0(\bar{X}, \bar{Y})(f). \tag{14.4}$$

The same holds for $\mathcal{D}^{\infty,r+s}$, $f \in \mathcal{D}^{\infty,r+s}$.

Proof In [39], equation (14.4) was proved under the assumption that $\bar{X}(f) \circ f^{-1}, \bar{Y}(f) \circ f^{-1}, \bar{Z}(f) \circ f^{-1} \in C^\infty(TM)$. What was really needed there was $\bar{X}(f) \circ f^{-1}, \bar{Y}(f) \circ f^{-1}$, $\bar{Z}(f) \circ f^{-1} \in C^1(TM)$. But, we conclude from lemma 14.16 above $\bar{X}(f) \circ f^{-1}$, $\bar{Y}(f) \circ f^{-1}$, $\bar{Z}(f) \circ f^{-1} \in C^2(TM)$. Hence, the proof of [39] carries over to our case. It essentially uses lemma 14.18. □

Corollary 14.20 *$\bar{\nabla}$ is the Levi-Civita connection for C^s vector fields on $D^{\infty,r}$, $r > \frac{n}{2}+1$.*

Proof $\bar{\nabla}$ is defined for C^s vector fields on $D^{\infty,r+s}$, $r > \frac{n}{2}+1$, $s \geq 2$, $k-(r+s)-1 > s$. Equation (14.4) shows that under these conditions it is the Levi- Civita connection for such vector fields and g_0 restricted to $\mathcal{D}^{\infty,r+s}$. But $\mathcal{D}^{\infty,r+s}$ is dense in $\mathcal{D}^{\infty,r}$ and we can extend (14.4) to $\mathcal{D}^{\infty,r}$. □

Remark 14.21 This is the step where we must go from $\mathcal{D}^r$ to $\mathcal{D}^{\infty,r}$. $\mathcal{D}^{r+s}$ is probably not dense in $\mathcal{D}^r$ and we cannot conclude as in corollary (14.20). □

Proposition 14.22 *Let $\bar{V}_0 \in T\mathcal{D}^{\infty,r}$, $\bar{\gamma}_{\bar{V}_0}$ be the geodesic in $\mathcal{D}^{\infty,r}$ with initial condition $\bar{V}_0$. Then $\bar{\gamma}_{\bar{V}_0} : I \longrightarrow \mathcal{D}^{\infty,r}$ is given by the map*

$$t \mapsto (x \mapsto \gamma_{\bar{V}_0(x)}(t)).$$

where $\gamma_{\bar{V}_0(x)}$ is the geodesic in M with initial condition $\bar{V}_0(x) \in TM$.

Proof $(\bar{\nabla}_{\dot{\bar{\gamma}}_{\bar{V}_0}} \dot{\bar{\gamma}}_{\bar{V}_0})(t)(x) = \bar{K}(\ddot{\bar{\gamma}}_{\bar{V}_0}(t))(x) = K \circ \ddot{\bar{\gamma}}_{\bar{V}_0}(t)(x) = K(\ddot{\bar{\gamma}}_{\bar{V}_0(x)}(t)) = 0.$ □

Hence, a curve $\gamma(t)$ in $\mathcal{D}^{\infty,r}$ is a geodesic in $\mathcal{D}^{\infty,r}$ if and only if the associated curve $\gamma(t)(x)$ is a geodesic in M for all $x \in M$.

Now, we want to study the geometry of $\mathcal{D}^{\infty,r}_{\omega,0} \equiv \mathcal{D}^r_{\omega,0}$ and start with $\omega = \mu$ the volume form of (M^n, g). Then it follows immediately from the proof of theorem 14.5 that

$$T_f D^{\infty,r}_{\mu,0} = \{X | X \circ f^{-1} \in \Omega^r(TM), div(X \circ f^{-1}) = 0\}.$$

For a vector Y denote by $Y^\flat$ the corresponding 1-form, i.e. if $Y = \eta^i \frac{\partial}{\partial u^i}$ then $Y^\flat = \eta_i du^i$ and $(Y^\flat)^\sharp = Y$. Divergence freeness of Y is equivalent to $\delta Y^\flat = 0$. Under our assumption on σ_e, $\mathcal{D}^{\infty,r}_{\mu,0} = \mathcal{D}^r_{\mu,0}$ is a submanifold of $\mathcal{D}^{\infty,r}_0 = \mathcal{D}^r_0$. Hence, it is endowed with an induced weak Riemannian metric. We want to describe explicitly the corresponding Levi-Civita connection. Consider the Laplace operator $\Delta = \delta d : \Omega^r(M) \longrightarrow \Omega^{r-2}(M)$. The spectral assumption implies that $\operatorname{Im} \bar{d}$, $\operatorname{Im} \bar{\delta}$, and $\operatorname{Im} \bar{\Delta}$ are closed. Moreover, $\ker \Delta \cap L_2 = \{0\}$ according to $vol(M^n, g) = \infty$, and Δ has a bounded inverse Δ^{-1}.

Proposition 14.23 *For $V \in T_{\mathrm{id}}\mathcal{D}^{\infty,r}_0$ let $P_{\mathrm{id}}(V) := V - (d\Delta^{-1}\delta(V^\flat))^\sharp$. Then P_{id} is the orthogonal projection $P_{\mathrm{id}} : T_{\mathrm{id}}\mathcal{D}^{\infty,r}_0 \longrightarrow T_{\mathrm{id}}\mathcal{D}^{\infty,r}_{\mu,0}$ associated to g_0.*

Proof One has to show

$$\begin{aligned} &\operatorname{Im} P_{\mathrm{id}} \subset T_{\mathrm{id}} D^{\infty,r}, \\ &P_{\mathrm{id}}^2 = P_{\mathrm{id}}, \\ &P_{\mathrm{id}}(T_{\mathrm{id}} D^{\infty,r}) = T_{\mathrm{id}} D^{\infty,r}_\mu, \\ &W \in \ker P_{\mathrm{id}}, V \in \operatorname{Im} P_{\mathrm{id}} \text{ implies } g_0(V,W) = 0. \end{aligned}$$

These assertions are easy calculations which are performed in the proof of Proposition 11.26 of [39]. They hold under our assumptions also in the noncompact case.

□

Proposition 14.24 *Assume $k - r \geq 2$, $4r > \frac{n}{2} + 1$. Then P is a C^2 - morphism of fiber bundles.*

Proof As we have already seen, for $k - r \geq 2$, $r > \frac{n}{2} + 1$, $T\mathcal{D}_0^{\infty,r}$ and $T\mathcal{D}_{\mu,0}^{\infty,r}$ are manifolds of at least class C^2. First, it is clear that P_f is continuous with respect to the Ω^r-topology since P_{id} has this property and this property is preserved by right multiplication with f^{-1} and after that with f. We see immediately that $P_f^2 = P_f$. Moreover, P_f is the projection onto $T_f\mathcal{D}_{\mu,0}^{\infty,r}$ with respect to g_0 : Let $W \in (T_f\mathcal{D}_{\mu,0}^{\infty,r})^{\perp g_0}$ and $X \in T\mathcal{D}_0^{\infty,r}|_{T\mathcal{D}_{\mu,0}^{\infty,r}}$. Then $g_0(P_f X, W) = g_0(P_{\mathrm{id}}(X \circ f^{-1}) \circ f, W) = g_0(P_{\mathrm{id}}(X \circ f^{-1}) \circ f \circ f^{-1}, W \circ f^{-1})$ $= g_0(P_{\mathrm{id}}(X \circ f^{-1}), W \circ f^{-1}) = 0$, according to the fact that $W \circ f^{-1} \in (T_{\mathrm{id}}\mathcal{D}_{\mu,0}^{\infty,r})^{\perp g_0}$. Now, $T\mathcal{D}_0^{\infty,r}|_{\mathcal{D}_{\mu,0}^{\infty,r}} / T\mathcal{D}_{\mu,0}^{\infty,r}$ is of class $k - r$ since each element of the quotient is of class $k - r$ and $T\mathcal{D}_{\mu,0}^{\infty,r}$ is a submanifold of $T\mathcal{D}_0^{\infty,r}$. On $\mathcal{D}_0^{\infty,r}$ g_0 is a weak metric of class $k - r$. Hence, the isomorphism $(T\mathcal{D}_{\mu,0}^{\infty,r})^{\perp g_0} \longrightarrow T\mathcal{D}_0^{\infty,r}|_{\mathcal{D}_{\mu,0}^{\infty,r}} / T\mathcal{D}_{\mu,0}^{\infty,r}$ is of class $k - r$ and therefore also the projector P associated to the decomposition

$$T\mathcal{D}_0^{\infty,r}|_{\mathcal{D}_{\mu,0}^{\infty,r}} \cong T\mathcal{D}_{\mu,0}^{\infty,r} \oplus (T\mathcal{D}_{\mu,0}^{\infty,r})^{\perp g_0}$$

is of class $k - r$. □

Proposition 14.25 $\tilde{\nabla} := P \circ \bar{\nabla}$ *is the Levi - Civita connection for $g_0|_{D_{\mu,0}^{\infty,r}}$.*

Proof Let $X, Y \in C^2(T\mathcal{D}_{\mu,0}^{\infty,r+s})$. Then

$$\tilde{\nabla}_X Y - \tilde{\nabla}_Y X = P(\bar{\nabla}_X Y - \bar{\nabla}_Y X) = P([X, Y])$$

and $[X, Y] \in C^1(T\mathcal{D}_{\mu,0}^{\infty,r+s})$ since $\mathcal{D}_{\mu,0}^{\infty,r+s}$ is a submanifold of $\mathcal{D}_0^{\infty,r+s}$. Hence, $P([X, Y]) = [X, Y]$, and $\tilde{\nabla}_X Y - \tilde{\nabla}_Y X = [X, Y]$, i.e. $\tilde{\nabla}$ is torsion free. Let $X, Y, Z \in C^2(T\mathcal{D}_{\mu,0}^{\infty,r+s})$. We consider them as vector fields on $\mathcal{D}_0^{\infty,r+s}$, defined only on the submanifold $\mathcal{D}_{\mu,0}^{\infty,r+s}$. Then, since $\bar{\nabla}$ is the Levi - Civita connection of g_0,

$$Zg_0(X, Y) = g_0(\bar{\nabla}_Z X, Y) + g_0(\bar{\nabla}_Z Y, X). \tag{14.5}$$

The components in $(T\mathcal{D}_{\mu,0}^{\infty,r+s})^{\perp g_0}$ of $\bar{\nabla}_Z X$, $\bar{\nabla}_Z Y$ produce 0 in (14.5) since the second component of the scalar products on the right hand side of (14.5) belongs to $T\mathcal{D}_{\mu,0}^{\infty,r+s}$. Hence,

$$Zg_0(X, Y) = g_0(P(\bar{\nabla}_Z X), Y) + g_0(P(\bar{\nabla}_Z Y), X).$$

i. e.

$$Zg_0(X, Y) = g_0(\tilde{\nabla}_Z X, Y) + g_0(\tilde{\nabla}_Z Y, X).$$

□

Using density arguments, we extend all formulas from $\mathcal{D}_{\mu,0}^{\infty,r+s}$ to $\mathcal{D}_{\mu,0}^{\infty,r}$.

We come now to our application in hydrodynamics. We recall the Euler equations for an incompressible, homogeneous fluid without viscosity

$$\frac{\partial u}{\partial t} + \nabla_{u(t)} u(t) = \operatorname{grad} p, \tag{14.6}$$

$$\operatorname{div} u(t) \quad = \quad 0, \tag{14.7}$$

where $u = u(x,t)$ is a time dependent C^1 vector field on (M^n, g), $\nabla = \nabla^g$, $\operatorname{div} = \operatorname{div}_{\mathrm{dvol}_x(g)}$. Additionally, we assume $u(t) \in \Omega^r(TM)$ for all t, which means that the fluid moves very slowly at infinity, $r > \frac{n}{2} + 1$. $u(t)$ defines a 1-parameter family of diffeomorphisms f_t defined by

$$\frac{df_s}{ds}|_{s=t} = u(t) \circ f_t.$$

The f_t remain in the identity component of $D_\mu^{\infty,r}$, since $f_0 = \mathrm{id}$, $\operatorname{div} u = 0$, and $\mu = dvol_x(g)$.

Theorem 14.26 *Assume (M^n, g) with (I) and (B_k), $\inf \sigma_e(\Delta_1|_{(\ker \Delta_1)^\perp}) > 0$, $k - 2 \geq r > \frac{n}{2} + 1$. Then $u(t)$ satisfies the Euler equations (14.6) iff $\{f_t\}_t$ is a geodesic in $D_{\mu,0}^{\infty,r}$.*

Proof Under the above assumptions, the proof is the same as in the compact case. Assume (14.6) and apply P_{id} to it. This yields

$$P_{\mathrm{id}}\left(\frac{\partial u}{\partial t}\right) + P_{id}(\nabla_u u) = 0$$

and, since $\operatorname{div} u = 0$, $P_{\mathrm{id}}\left(\frac{\partial u}{\partial t}\right) = \frac{\partial u}{\partial t}$,

$$\frac{\partial u}{\partial t} = -P_{\mathrm{id}}(\nabla_u u).$$

We differentiate (14.6) and the equation $f_t \circ f_t^{-1} = \mathrm{id}$, and obtain

$$\begin{aligned} \frac{d^2 f}{ds^2}|_{s=t} &= u_*(t) \circ u(t) \circ f(t) + \frac{\partial u(t)}{\partial t} \circ f(t) = \\ &= u_*(t) \circ u(t) \circ f_t - P_{\mathrm{id}}(\nabla_{u(t)} u(t)) \circ f_t, \end{aligned}$$

where we identify $P_{\mathrm{id}}(\nabla_{u(t)} u(t)) \circ f_t$ with its horizontal lift in $T^2 D_{\mu,0}^{\infty,r-1}$. Using

$$P_{\mathrm{id}}(\nabla_{u(t)} u(t)) \circ f_t = P(\nabla_{u(t)} u(t) \circ f_t)$$

and applying $\tilde{K} = P \circ \bar{K}$, we obtain

$$\begin{aligned} \tilde{K}\left(\frac{d^2 f}{ds^2}|_{s=t}\right) &= P\bar{K}[u_*(t) \circ u(t) \circ f_t - P_{id}(\nabla_{u(t)} u(t) \circ f_t)] \\ &= P(\nabla_{u(t)} u(t) \circ f_t) - P(\nabla_{u(t)} u(t) \circ f_t) = 0, \end{aligned}$$

i.e. $\{f_t\}$ is a geodesic in $D_{\mu,0}^{\infty,r}$. We omit the converse direction and refer to [41], 187-188. □

Corollary 14.27 *Assume (M^n, g) with (I), (B_k), $\inf \sigma_e(\Delta_1|_{(\ker \Delta_1)^\perp}) > 0$. (These conditions are automatically satisfied if M^n is compact.) Then for small t the Euler equations (14.6) have a unique solution $u \in \Omega^r(TM)$ if $k \geq r > \frac{n}{2} + 1$.*

Proof This follows from the local existence of geodesics in Hilbert manifolds. □

15 The Ricci flow

After Perelman's spectacular papers on the solution of Thurston's geometrization conjecture, the Ricci flow is again in the center of the common interest. In the meantime, there appeared many papers concerning the geometric procedures of surgery and geometric implications. We will not contribute to this. The aim of this short contribution is to discuss the fundamental analytical questions connected with the Ricci flow. As we already discussed in section 11, one fundamental approach solving non-linear PDEs is to establish an appropriate non-linear evolution equation and to perform and to study $\lim\limits_{t\to\infty}$. For harmonic maps, this has been done first by Eells/Sampson in [42]. Similar approaches have been established for gauge theory.

The question for a "natural evolution" of the initial metric g_0 on a manifold M^n led R. Hamilton 1982 to the equation

$$\frac{\partial g_{ij}}{\partial t} = -2R_{ij}, \tag{15.1}$$

where $R_{ij} = R_{ij}(g) = \mathrm{Ric}\,(g)_{ij}$ is the Ricci tensor.

Hamiltion has chosen the right hand side for much the same reason that Einstein introduced the Ricci tensor into his theory of gravitation. He needed a symmetric 2-index tensor which arises naturally from the metric tensor and its first and second derivatives. The Ricci tensor is essentially the only possibility. Moreover, one has to choose the negative sign on the right hand side to obtain a well behaved forward in time equation.

Example 15.1 Let $(M^n, g) = (S^r_r, g_{standard}) = (S^n_r, r^2 \cdot g_{st}(S^n_1))$, $g_{ij} = r^2 \cdot g_{ij}(S^n_1)$, $R_{ij} = (n-1)g_{ij}(S^n_1)$. If we make the ansatz that the time dependence is contained only in $r = r(t)$, then (15.1) becomes

$$\frac{d(r^2)}{dt} = -2(n-1),$$

hence $r^2(t) = r^2(0) - 2(n-1)t$, i.e. the sphere S^n_r collapses to a point in finite time. □

This kind of collapsing can be prevented if one restricts to evolution with constant volume.

Denote by $R = R(g)$ the scalar curvature and by $r = r(g) = \frac{1}{\mathrm{vol}\,(M^n,g)} \int\limits_M R\,\mathrm{dvol}_x(g)$ its mean value. Consider the equation

$$\frac{\partial g_{ij}}{\partial t} = \frac{2}{n} r g_{ij} - 2R_{ij}. \tag{15.2}$$

Proposition 15.2 *If $g_{ij}(t)$ solves (15.2) such that* $\mathrm{vol}\,(M^n, g(t)) < \infty$ *and* $R(g(t)) \in L_1(M^n, g(t))$ *then* $\mathrm{vol}\,(M^n, g(t))$ *is constant.*

Proof $\mathrm{dvol}_x(g(t)) = \sqrt{\det g_{ij}(t)}$, $\frac{\partial}{\partial t}\log(\sqrt{\det g_{ij}(t)}) = \frac{1}{2} g^{ij} \frac{\partial}{\partial t} g_{ij} = r - R$, according to (15.2). On the other hand $\frac{\partial}{\partial t}\log\sqrt{\det g_{ij}} = \frac{\frac{\partial}{\partial t}\sqrt{\det g_{ij}}}{\sqrt{\det g_{ij}}}$. This yields

$$0 = \int\limits_M (r - R)\,\mathrm{dvol}_x(g) = \frac{\partial}{\partial t}\int\limits_M \mathrm{dvol}_x(g) = \frac{\partial}{\partial t}\mathrm{vol}\,(M, g(t)).$$

□

Corollary 15.3 *The assertion of 15.2 holds if M^n is closed.* □

(15.2) is called the normalized evolution equation. Any solution of (15.1) implies a solution of (15.2), as expressed by

Proposition 15.4 *Let $g_{ij}(t)$ be a solution of (15.1). Set*

$$\psi(t)^{-\frac{n}{2}} := \int_M \mathrm{dvol}_x(g(t)),$$

$$\tilde{t}(t) := \int_0^t \psi(\xi) d\xi,$$

$$\tilde{g}_{ij}(\tilde{t}) := \psi(t(\tilde{t})) g_{ij}(t).$$

Then $\tilde{g}_{ij}(\tilde{t})$ satisfies the differential equation

$$\frac{\partial \tilde{g}_{ij}}{\partial \tilde{t}} = \frac{2}{n} \tilde{r}(\tilde{t}) - 2\tilde{R}_{ij}(\tilde{t}).$$

The proof is an easy calculation applying the chain rule. □

Now the essential point is an existence theorem for the initial value problem

$$\frac{g_{ij}}{\partial t} = -2R_{ij}, \quad g_{ij}|_{t_0} = \mathring{g}_{ij}. \tag{15.3}$$

First we restrict to the case M^n closed. If the equation (15.1) would be strictly parabolic (i.e. its linearization is) than we would get by standard theorems for $0 \le t \le \varepsilon$ a unique solution of 15.4. Unfortunately, (15.2) is not strictly parabolic.

Consider an evolution equation

$$\frac{\partial f}{\partial t} = E(f), \tag{15.4}$$

$E(f)$ a non-linear differential operator of degree 2 in f (15.4) has the linearization

$$\frac{\tilde{f}}{\partial t} = (DE(f))(\tilde{f}),$$

where $\tilde{f}$ is a variation of f.

We say E is (strictly) parabolic if its linearization is (strictly) parabolic around an f. This is the case if and only if all eigenvalues of $\sigma DE(f)\xi$ have (strictly) positive real parts when $\xi \neq 0$.

Lemma 15.5 *Let $E(g_{ij}) = -2R_{ij}$. Then*

$$\begin{aligned} DE(g_{ij})\tilde{g}_{jk} &= -2\tilde{R}_{jk} \\ &= g^{hi} \left\{ \frac{\partial^2 \tilde{g}_{jk}}{\partial x^h \partial x^i} - \frac{\partial^2 \tilde{g}_{hj}}{\partial x^i \partial x^k} - \frac{\partial^2 \tilde{g}_{ik}}{\partial x^h \partial x^j} + \frac{\partial^2 \tilde{g}_{hi}}{\partial x^t \partial x^k} \right\} + \textit{lower order terms}. \end{aligned} \tag{15.5}$$

Proof This is an easy calculation. □

We obtain from (15.5)

$$\sigma_L DE(g_{jk})(\xi)\tilde{g}_{jk} = g^{hi}\left\{\xi_h\xi_i\tilde{g}_{jk} - \xi_i\xi_k\tilde{g}_{hj} - \xi_h\xi_j\tilde{g}_{ik} + \xi_j\xi_k\tilde{g}_{hi}\right\}.$$

Choose a coordinate system so that $g_{jk}(\text{point}) = \delta_{jk}$, $\xi_1 = 1$, $\xi_i = 0$ for $i \neq 1$. Then with $T_{jk} = \tilde{g}_{jk}$,

$$\begin{aligned}
&[\sigma_L DE(g)(\xi)T]_{jk} = T_{jk} \text{ for } j \neq 1, k \neq 1,\\
&[\sigma_L DE(g)(\xi)T]_{1k} = 0 \text{ for } k \neq 1,\\
&[\sigma_L DE(g)(\xi)T]_{11} = T_{22} + \cdots + T_{nn}.
\end{aligned}$$

Hence $\sigma DE(g)(\xi)$ is not injective, has an eigenvalue $= 0$, (15.1) is not strictly parabolic.

There are two explanations for this fact. Consider the equation $R_{ij} = 0$. The space of solutions has either zero or infinite dimension since with a solution g also f^*g, $f \in$ Diff (M), is a solution. If (15.1) would be parabolic then $R_{ij} = 0$ would be elliptic with finite dimensional space of solutions (since M^n is closed). Another explanation comes from the second Bianchi identity.

Now one has to solve (15.3), having in mind that (15.1) is not strictly parabolic. There are two existence proofs. One classical, in a certain sense canonical proof by Hamilton as presented in [74] and the other much shorter by De Turk, first presented in [39]. We give an outline of Hamilton's proof.

Consider an evolution equation $\frac{\partial f}{\partial t} = E(f)$, $E(f)$ a non-linear differential operator of order 2 in f, $E : C^\infty(M,U) \longrightarrow C^\infty(M,F)$, where $F \longrightarrow M$ is a vector bundle, $U \subset F$ open. Let $G \longrightarrow M$ be a vector bundle, $L(g)h$ a differential operator of first order in g and h, $g \in C^\infty(M,U)$, $h \in C^\infty(M,F)$, $L(g)h \in C^\infty(M,G)$. Suppose that $Q(f) := L(f)E(f)$ has at most degree 1 in f. Then $L(f)$ is called an integrability condition for $\frac{\partial f}{\partial t} = E(f)$.

We calculate the linearizations. Linearization in E yields $L(f)DE(f)\tilde{f}$, linearization in L yields $DL(f)\{E(f),\tilde{f}\}$. We get

$$L(f)DE(f)\tilde{f} + DL(f)\{E(f),\tilde{f}\} = DQf\tilde{f}.$$

$DL(f)\{E(f),\tilde{f}\}$, $DQf\tilde{f}$ have only degree 1 in $\tilde{f}$, but $L(f)DE(f)\tilde{f}$ is of degree 3 in the derivations of $\tilde{f}$

$$\begin{aligned}
&\sigma_L L(f)(\xi) \circ \sigma_L DE(f)\xi = 0,\\
&\operatorname{Im}\sigma_L DE(f)(\xi) \subseteq \ker \sigma_L L(f)(\xi).
\end{aligned}$$

Theorem 15.6 *Let M^n be closed and*

$$\frac{\partial f}{\partial t} = E(f) \tag{15.6}$$

be an evolution equation with integrability condition $L(f)$, $\operatorname{Im}\sigma_L DE(f)(\xi) \subseteq \ker\sigma_L L(f)(\xi)$. Suppose that $L(f)E(f) = Q(f)$ has degree 1 and all the eigenvalues of the eigenspaces of $\sigma_L DE(f)(\xi)$ in $\ker\sigma L(f)(\xi)$ have strictly positive real points. Then the initial value problem $f|_{t=0} = f_0$ has a unique smooth solution for a short time $0 \le t \le \varepsilon$, where ε may depend on f_0.

Outline of the proof. First consider the equation

$$\frac{\partial \overline{f}}{\partial t} - E(\overline{f}) = \overline{h}, \quad 0 \leq t \leq 1, \quad \overline{f}|_{t=0} = \overline{f}_0.$$

Assertion. Then there exists for every f_0 sufficiently near to $\overline{f}_0$ and h sufficiently near to $\overline{h}$ a unique solution f,

$$\frac{\partial f}{\partial t} - E(f) = h, \quad 0 \leq t \leq \varepsilon, \quad f|_{t=0} = f_0.$$

The assertion implies theorem 15.6: First one solves the equation $\frac{\partial f}{\partial t} - E(f)$, $f|_{t=0} = f_0$ at $t = 0$ by a function $\overline{f}$ (determine the formal Taylor series from the PDE and the initial condition). Let $\overline{h} = \frac{\partial \overline{f}}{\partial t} - E(\overline{f})$. The formal Taylor series of $\overline{h}$ vanishes at $t = 0$. Now perturb $\overline{h}$ by h such that $h = 0$ for $0 \leq t \leq \varepsilon$, ε sufficiently small. We infer from the assertion that there exists f satisfying $\frac{\partial f}{\partial t} - E(f) = h$, $f|_{t=0} = f_0$, i.e. f solves the evolution equation for $0 \leq t \leq \varepsilon$.

One has to prove the assertion. Consider for this the operator $\mathcal{E} : \mathbb{C}^\infty(M \times [0,1], F) \longrightarrow C^\infty(M \times [0,1], F) \times C^\infty(M, F)$,

$$\mathcal{E}(f) = \left(\frac{\partial}{\partial t} - E(f), f|_{t=0} \right).$$

$\mathcal{E}$ has the linearization

$$D\mathcal{E}(f)\tilde{f} = \left(\frac{\partial \tilde{f}}{\partial t} - DE(f)\tilde{f}, \tilde{f}|_{t=0} \right).$$

Aim. To show that the linearized equation $\frac{\partial \tilde{f}}{\partial t} - DE(f)\tilde{f} = \tilde{h}$, $\tilde{f}|_{t=0} = \tilde{f}_0$ has a unique solution $\tilde{f}$, and the solution $\tilde{f}$ is a smooth tame function of $\tilde{h}$ and $\tilde{f}_0$. Then, according the Nash-Moser inverse function theorem, one gets a solution for the non-linearized equation, i.e. 15.6 is done.

Hence there remains 3 tasks,

1) solution of the linearized problem,

2) proof, that the assumptions of the Nash-Moser theorem are fulfilled,

3) appplication of the Nash-Moser theorem.

For reasons of space we cannot perform these steps which are not too complicated and we refer to [74]. □

As we mentioned above, a substantially simpler and shorter existence proof has been established by DeTurck. We give a brief outline.

Fix a background metric $\hat{g}$ on M. Then

$$(\nabla^g - \nabla^{\hat{g}})^i_{jk} = D^i_{jk} = \Gamma^i_{jk} - \hat{\Gamma}^i_{jk} = \frac{1}{2} \sum_l g^{il}(g_{jl;k} + g_{kl;j} - g_{jk;l}), \tag{15.7}$$

where $; = \nabla^{\hat{g}}$, and

$$R_{ij} - \hat{R}_{ij} = \sum_k \left\{ D^k_{ij;k} - D^k_{ki;j} + \sum_l (D^k_{kl} D^l_{ji} - D^k_{jl} D^l_{ki}) \right\}. \tag{15.8}$$

Using (15.7), one can calculate the second order terms of (15.8) as

$$-\frac{1}{2}\sum_{k,l} g^{kl} g_{ij;kl} + \frac{1}{2}\sum_{k,l} g^{kl}(g_{il;jk} + g_{jl;ik} - g_{kl;ij}).$$

Define the vector field $X = X(g)$ by

$$X^p(g) = -\sum_{i,k,l} g^{pi} g^{kl}(g_{ik;l} - \frac{1}{2} g_{kl;i}$$

and the elliptic operator $F(g)$ by

$$F_{ij}(g) = \sum_{k,l} g^{kl} g_{ij;kl} + Q_{ij}(g, \nabla^{\hat{g}} g),$$

where Q is a quadratic polynomial in the indicated variables. $\mathrm{Ric}\,(g)$ then can be written as

$$\mathrm{Ric}\,(g) = \mathrm{Ric}\,(\hat{g}) - \frac{1}{2} F(g) - \frac{1}{2} L_x g. \tag{15.9}$$

Consider now the initial value problem

$$\begin{aligned} \frac{\partial \overline{g}}{\partial t} &= F(\overline{g}) - 2\mathrm{Ric}\,(\overline{g}), \\ \overline{g}|_{t=0} &= g_0. \end{aligned} \tag{15.10}$$

(15.10) is in fact a quasilinear parabolic equation and we have by standard theorems (as cited above) a solution for small t. Let $\Phi_t : M \longrightarrow M$ be the flow of the vector field $X(t) = X(\overline{g}(t))$ and set $g(t) := \Phi_t^* \overline{g}(t)$. Then

$$\frac{\partial g}{\partial t} = \Phi_t^* \left(\frac{\partial \overline{g}}{\partial t} \right) + L_{x(t)} g(t)$$

and, according to (15.9) and (15.10),

$$\Phi_t^* \left(\frac{\partial \overline{g}}{\partial t} \right) = \Phi_t^*(-2\mathrm{Ric}\,(\overline{g}) - L_x(\overline{g})) = 2\mathrm{Ric}\,(g) - L_x(g),$$

i.e.

$$\frac{\partial g}{\partial t} = -2\mathrm{Ric}\,(g(t)).$$

The reason for the non-parabolicity of $\frac{\partial g}{\partial t} = -2\mathrm{Ric}\,(g)$ was the gauge invariance under $\mathrm{Diff}\;(M)$. Fixing the gauge $\{\Phi_t\}_t$, reduced the problem to a parabolic problem.

Until now, we restricted to the case M^n closed. In [108], [109] W.-X. Shi extended the short time existence under certain conditions to open complete manifolds.

Theorem 15.7 *Let (M^n, g_0) be open, complete, satisfying $(B_0(g_0))$, $|R_{ijkl}|^2 \le k_0$. Then there exists a constant $T(n, k_0) > 0$ such that the initial value problem*

$$\frac{\partial g_{ij}}{\partial t} = -2R_{ij},$$

$$g_{ij}|_{t=0} = g_{0,ij}$$

has for $0 \le t \le T(n, k_0)$ *a smooth solution satisfying the following estimates: For any* $m \ge 0$ *there exists* $c_m = c_m(m, k_0)$ *such that*

$$\sup_{x \in M} |\nabla^m R_{ijkl}(x,t)|^2 \le \frac{c_m}{t^m}, \quad 0 < t \le T(n, k_0).$$

□

It is well known that under orthogonal transformations the curvature tensor splits into three components,

$$R_m = \{R_{ijkl}\} = W + V + U$$

where $W = \{W_{ijkl}\}$ is the Weyl conformal curvature tensor and $V = \{V_{ijkl}\}$ and $U = \{U_{ijkl}\}$ denote the traceless Ricci part and the scalar curvature part, respectively. With $\overset{\circ}{R}_{ij} = R_{ij} - \frac{1}{n} g_{ij}$ the explicit expressions are well known.

Set $\overset{\circ}{R}_m = \{\overset{\circ}{R}_{ijkl}\} = \{R_{ijkl} - U_{ijkl}\} = V + W$.

Theorem 15.8 *Let* $(M^n, g)_0$ *be open, complete,* $n \ge 3$. *For any* $c_1, c_2 > 0$ *and* $\delta > 0$ *there exists* $\varepsilon = \varepsilon(n, c_1, c_2, \delta) > 0$ *such that the following holds: If* $\mathrm{vol}_{\,g_0}(B_\gamma(x)) \ge c_1 \cdot \gamma^n$ *for all* $x \in M$ *and* $|\overset{\circ}{R}_m(g_0)|^2 \le \varepsilon R^2$, $0 < R \le c_2/\mathrm{dist}_{(X_0,X)^{2+\delta}}$ *for all* $x \in M$ *then the initial value problem*

$$\frac{\partial g_{ij}}{\partial t} = -2R_{ij},$$
$$g_{ij}|_{t=0} = g_{0,ij}$$

has a solution for all $t \ge 0$, *the metric* $g(t)$ *converges with respect to the* C^∞*-topology to a metric* $g(\infty)$ *such that* $R_m(g(\infty)) = \{R_{ijkl}(\infty)\} \equiv 0$ *on* M. □

Concerning closed 3-manifolds, the papers of Perelmann contain many striking results, but their discussion is outside of this contribution.

Our goal was a certain survey on PDEs on manifolds, and we stop here.

Bibliography

[1] I. Agricola, T. Friedrich, P.-A. Nagy, C. Puhle: On the Ricci tensor in the common sector of type II string theory *Class. Quant. Grav.* **22** (2005) 2569–2577

[2] I. Agricola, T. Friedrich: Killing spinors in supergravity with 4–fluxes *Class. Quant. Grav.* **20** (2003) 4707–4717

[3] L. Ahlfors: Some remarks on Teichmüller's space of Riemannian surfaces *Ann. Math.* **74** (1961) 171–191

[4] N. Anghel: An abstract index theorem on non–compact Riemannian manifolds *Houston J. Math.* **19** (1993) 223–237

[5] N. Aronszajn: A unique continuation theorem for solutions of elliptic partial differential equations or inequalities of second order *J. Math. Pures Appl.* **36** (1957) 235–249

[6] M. F. Atiyah, V. Patodi, I. M. Singer: Spectral asymmetry and Riemannian geometry I *Math. Proc. Cambr. Philosoph. Soc.* **77** (1975) 43–69

[7] M. F. Atiyah: Elliptic operators, discrete groups and von Neumann algebras *Asterisque* **32/33** (1976) 43–72

[8] T. Aubin: Espaces de Sobolev sur les variétés Riemanniennes *Bull. Sci. mat.* **100** (1976) 149–173

[9] R. Azencott: Behavior of diffusion semi–groups at infinity *Bull. Soc. Math.* **102** (France), (1974) 193–240

[10] Y. Baldin, M. Noronhra: Some complete manifolds with nonnegative curvature operator *Math. Z.* **195** (1987) 383–390

[11] P. Barbasch, H. Moscovici: L_2–index and the Selberg trace formula *J. Functional Analysis* **53** (1983) 151–201

[12] M. Belger, R. Schimming, V. Wünsch: A Survey on Huygens' Principle *J. Anal. Appl.* **16** (1997) 5–32

[13] L. Bers: Quasi conformal mappings and Teichmüller's theorem *In: Analytic Functions* (Princeton, 1960) 89–120

[14] D. E. Blair: *Contact manifolds in Riemannian geometry* (Lecture Notes in Mathematics 509, Berlin, 1976)

[15] A. Borel: The L_2–cohomology of negatively curved Riemannian symmetric spaces *Annales Acad. Sci. Fenn., Ser. A* **10** (1985) 95–105

[16] J. Brüning, M. Lesch: Hilbert Complexes *J. Funct. Anal.* **108** (1992) 88–132

[17] U. Bunke: *Spektraltheorie von Diracoperatoren auf offenen Mannigfaltigkeiten* (PhD thesis, Greifswald, 1991)

[18] P. Buser: *Geometry and Spectra of Compact Riemannian Surfaces* (Progress in Mathematics vol. 106, Boston, Birkhäuser Verlag, 1992)

[19] I. Buttig: Spectral approximation by L_2 discretization of the Laplace operator on manifolds of bounded geometry *Ann. Global Anal. Geom.* **6** (1988) 55-107

[20] I. Buttig, J. Eichhorn: The heat kernel for p–forms on manifolds of bounded geometry *Acta Sci. Math.* **55** (1991) 33–51

[21] G. Carron: Index theorems for noncompact manifolds *J. Reine Angew. Math.* **541** (2001) 81–115

[22] G. Carron: L_2–cohomology of manifolds with flat ends *GAFA* **13** (2003) 366–395

[23] I. Chavel: *Eigenvalues in Riemannian geometry* (New York, Academic Press, 1984)

[24] J. Cheeger: On the Hodge theory of Riemannian pseudomanifolds *In: Proc. Symp. Pure Math.* (36, 1980) 91–146

[25] J. Cheeger, M. Gromov, M. Taylor: Finite propagation speed, kernel estimates for functions of the Laplacian, and geometry of complete Riemannian manifolds *J. Diff. Geom.* **17** (1983) 15–53

[26] J. Cheeger, S. T. Yau: A lower bound for the heat kernel *Comm. Pure Appl. Math.* **34** (1981) 465–480

[27] S. Y. Cheng, P. Li, S. T. Yau: On the upper estimate of the heat kernel of complete Riemannian manifold *Amer. J. Math.* **103** (1981) 1021–1063

[28] A. Connes, H. Moscovici: The L_2.–index theorem for homogeneous spaces of Lie groups *Ann. Math.* **115** (1982) 291–330

[29] D. De Turck: Deforming metrics in direction of their Ricci tensors *J. Diff. Geom* **18** (1983) 157–162

[30] E. B. Davies: Explicit constants for Gaussian upper bounds at heat kernels *Amer. J. Math.* **109** (1987) 319–334

[31] J. Dieudonne: *Grundzüge der modernen Analysis, vol. 2* (DVW Publ. Berlin, 1975)

[32] J. Dodziuk: L_2–harmonic forms on rotationally symmetric Riemannian manifolds *Proc. Amer. Math. Soc.* **77** (1979) 395–400

[33] J. Dodziuk: Maximum principle for parabolic inequalities and the heat flow on open manifolds *Indiana Univ. Math. J.* **32** (1983) 703–716

[34] H. Donnelly, P. Li: Lower bounds for the eigenvalues of complete Riemannian manifolds *Michigan Math. J.* **29** (1982) 149–161

[35] H. Donnelly, F. Xavier: On the differential form spectrum of negatively curved Riemannian manifolds *Am. J. Math.* **106** (1984) 169–185

[36] H. Donnelly, C. Fefferman: L_2–cohomology and index theorem for the Bergman metric *Ann. Math.* **118** (1983) 593–618

[37] C. J. Earle: Teichmüller Theory *In: Discontinuous Groups and Automorphic Functions* (Academic Press, 1977) 143–162

[38] C. J. Earle, J. Eells: A fibre bundle description of Teichmüller theory *J. Diff. Geom.* **3** (1969) 19–43

[39] D. G. Ebin: *Espace de metriques Riemanniennes et movement des fluids via les varieties d'applications* (Lecture Notes, Paris, 1972)

[40] D. G. Ebin, J. E. Marsden: Groups of diffeomorphisms and the motion of an incompressible fluid *Ann. Math.* **92** (1970) 102–163

[41] D. G. Ebin, A. E. Fischer, J. E. Marsden: Diffeomorphism groups, hydrodynamics and relativity *In: Proc. 13th Biennal Seminar of Canad. Math. Congress* (Montreal, 1972) 135–279

[42] J. Eells, J. H. Sampson: Harmonic Mappings of Riemannian Manifolds *Am. J. Math.* **86** (1964) 109–160

[43] J. Eichhorn: *Global Analysis on Open Manifolds* (to appear in Nova Science Publ., New York, 544)

[44] J. Eichhorn: The Euler characteristic and signature for open manifolds *In: Suppl. ai Rendiconti del Circolo Matematico di Palermo* (Serie II, no. 16, 1987) 29–41

[45] J. Eichhorn: Elliptic operators on noncompact manifolds *In: Teubner-Texte zur Math.* (vol. 106, 1988) 4–169

[46] J. Eichhorn: The boundedness of connection coefficients and their derivatives *Math. Nachr.* **152** (1991) 145–158

[47] J. Eichhorn: The heat semigroup acting on tensors or differential forms with values in a vector bundle *Archivum Mathematicum* **27** (1991) 15–24

[48] J. Eichhorn: Poincare's theorem and Teichmüller theory for open surfaces *Asian J. Math.* **2** (1998) 355–404

[49] J. Eichhorn: Gauge theory on open manifolds of bounded geometry *Int. Journ. Mod. Physics* **7** (1992) 3927–3977

[50] J. Eichhorn: Harmonic maps between open manifolds (in preparation)

[51] J. Eichhorn, T. Friedrich: Seiberg–Witten theory *In: Symplectic singularities and geometry of gauge fields* (Proc. of the Banach Center symposium on differential geometry and mathematical physics in Spring 1995, Warsaw, Poland, Banach Center Publ. 39, 1997) 231–267

[52] J. Eichhorn: *Relative Zetra Functions, Determinants, Torsion, Index theorems and Invariants for Open Manifolds* (to appear in World Scientific Publ., Singapore, 250 pp., 2007)

[53] J. Eichhorn, G. Heber: The configuration space of gauge theory on open manifolds of bounded geometry *In: Symplectic singularities and geometry of gauge fields* (Proc. of the Banach Center symposium on differential geometry and mathematical physics in Spring 1995, Warsaw, Poland, Banach Center Publ. 39, 1997) 269–286

[54] J. Eichhorn, R. Schmid: Form preserving diffeomorphisms on open manifolds *Ann. Global Anal. and Geom.* **14** (1996) 147–176

[55] J. F. Escobar, A. Freire, M. Min–Oo: L_2–vanishing theorems in positive curvature *Indiana Univ. Math. J.* **42** (1993) 1545–1553

[56] M. Farber, S. Weinberger: On the zero–in–the–spectrum conjecture *Ann. Math.* **154** (2001) 139–154

[57] A. E. Fischer, A. J. Tromba: On a purely Riemannian proof of the structure and dimension of the unramified moduli space of a compact Riemannian surface *Math. Ann.* **267** (1984) 311–345

[58] D. Freed, K. Uhlenbeck: *Instantons and Four–Manifolds* (New York, Springer Verlag, 1984)

[59] F. G. Friedlander: *The Wave Equation on a Curved Space–Time* (Cambridge Univ. Press, 1975)

[60] T. Friedrich, S. Ivanov: Parallel spinors and connections with skew–symmetric torsion in string theory *Asian J. Math.* **6** (2002) 303–336

[61] T. Friedrich, E. C. Kim: The Einstein–Dirac equation on Riemannian spin manifolds *J. Geom. Phys.* **33** (2000) 128–172

[62] M. Gaffney: A special Stokes' theorem for complete manifolds *Ann. Math.* **60** (1954) 40–145

[63] S. Gallot, D. Meyer: Operateurs de Courbure et Laplacian des formes differentielles d'une variete Riemannienne *J. Math. Pures Appl.* **54** (1975) 259–284

[64] P. Gilkey: Spectral geometry of real and complex manifolds *In: Proc. Symp. Pure Math.* (AMS, 27, 1975) 265–280

[65] P. Gilkey: Spectral geometry of Riemannian manifolds *J. Diff. Geom.* **10** (1975) 601–618

[66] R. Greene, H. Wu: *Function theory on manifolds which posses a pole* (Lecture Notes in Mathematics 699, Berlin, 1979)

[67] A. Grigoryan: Estimates of heat kernel on Riemannian manifolds *London Math. Soc. Lecture Notes Series* **273** (Cambridge), (1999) 140–225

[68] M. Gromov: Kähler hyperbolicity and L_2–Hodge theory *J. Diff. Geom.* **33** (1991) 263–292

[69] M. Gromov, H. B. Lawson: Positive scalar curvature and the Dirac operator on complete Riemannian manifolds *Publ. Math. IHES* **58** (1983) 295–408

[70] P. Günter: *Zur Gültigkeit des Huygensschen Prinzips bei partiellen Differentialgleichungen vom normalen hyperbolischen Typus* (Ber. Verh. Sächs. Akad. Wiss. Leipzig, Math. Nat. Klasse 100, Heft 2., 1952)

[71] P. Günter: *Huygens' Principle and Hyperbolic Equations* (Perspectives in Mathematics, San Diego, Academic Press, 1988)

[72] P. Günter: Über die Darbouxsche Differentialgleichung mit variablen Koeffizienten *Math. Nachr.* **22** (1960) 285–321
[73] J. Hardamard: *Lectures on Cauchy's Problem in Linear Partial Differential Equations* (New Haven: Yak. Univ. Press, 1923)
[74] R. Hamilton: Three–manifolds with positive Ricci curvature *J. Diff. Geom.* **17** (1982) 255–306
[75] P. Hartmann: On homotopic harmonic maps *Can. J. Math.* **19** (1967) 673–687
[76] N. Higson, J. Roe, T. Schick: Spaces with vanishing L_2–homology *Geom. Dedicata* **87** (2001) 335–343
[77] W. C. Hsiang, V. Pati: L_2–cohomology of normal algebraic surfaces I *Invent. math.* **81** (1985) 395–412
[78] G. Jensen: Einstein metrics on principal fibre bundles *J. Diff. Geom.* **8** (1973) 599–614
[79] J. Jost, Z. Kang: Vanishing theorems for L_2–cohomology on infinite coverings of compact Kähler manifolds and applications in algebraic geometry *Comm. Anal. Geom.* **8** (2000) 1–30
[80] J. Jost: *Nonlinear Methods in Riemannian and Kählerian Geometry* (Birkhäuser Publ., Basel, 1988)
[81] T. Kato: *Perturbation Theory for Linear Operators* (Grundlehren der mathematischen Wissenschaften 132, Berlin, 1972)
[82] Y. A. Kordjukov: L_p–theory of elliptic differential operators on manifolds of bounded geometry *Acta Applicandae Mathematicae* **23** (1991) 223–260
[83] H. B. Lawson: *The theory of gauge fields in four dimensions* (CBMS Reg. Conf. Ser. in Mathematics 58, Providence, 1985)
[84] N. Lohoue: Estimation des projecteurs de de Rahm Hodge de certaines varietes Riemanniennes non compactes *Math. Nach.* **279** (2005) 272–298
[85] M. Marcolli: *Seiberg–Witten gauge theory* (Texts and Readmaps in Mathematics 17, New Dehli, 1999)
[86] R. Mazzeo, R. S. Phillips: Hodge theory on hyperbolic manifolds *Duke Math. J.* **60** (1990) 509–559
[87] R. Melrose: *Geometric Scattering Theory* (Cambridge University Press, 1995)
[88] J. W. Morgan: *The Seiberg–Witten equations and applications to the topology of smooth four–manifolds* (Mathematical Notes 44, Princeton, 1996)
[89] H. Moscovici: L_2–index of elliptic operators on locally symmetric spaces of finite volume *Contemp. Math.* **10** (1982) 129–138
[90] W. Müller: L_2–index theory, eta invariants and values of L_2–functions *Contemp. Math.* **105** (1990) 145–189
[91] W. Müller: Signature defects of cusps of Hilbert modular varieties and values of L–series at $s = 1$ *J. Diff. Geom.* **20** (1984) 55–119
[92] M. Nagase: L_2–cohomology and intersection homology of stratified spaces *Duke Math. J.* **50** (1983) 329–368
[93] M. Reed, B. Simon: *Methods of Modern Mathematical Physics II* (Moscow, Publishers Mir., 1982)
[94] J. Roe: An index theorem on open manifolds I *J. Diff. Geom.* **27** (1988) 87–113
[95] J. Roe: An index theorem on open manifolds II *J. Diff. Geom.* **27** (1988) 115–136

[96] J. Roe: Operator algebras and index theory on non–compact manifolds *Contemporary Mathematics* **70** (1988) 229–240
[97] J. Roe: *Index theory, coarse geometry and topology of open manifolds* (CBMS Reg. Conf. Series in Math. vol. 90, Providence, 1996)
[98] J. Roe: *Coarse cohomology and index theory on complete Riemannian manifolds* (Memoirs of the AMS 497, 1993)
[99] J. Roe: *Elliptic operators, topology and asymptotic methods* (Pitman Research Notes in Mathematics, Series 179, New York, 1988)
[100] G. Salomonsen: Equivalence of Sobolev spaces *Results Math.* **39** (2001) 115–130
[101] L. Saper, M. Stern: L_2–cohomology of arithmetic varieties *Ann. Math.* **132** (1990) 1–69
[102] L. Saper: L_2–cohomology and intersection homology of certain algebraic varieties with isolated singularities *Invent. math.* **82** (1985) 207–255
[103] R. Schimming: Das Huygenssche Prinzip bei hyperbolischen Differentialgleichungen für allgemeine Felder *Beitr. Anal.* **11** (1978) 45–90
[104] R. Schimming: Hyperbolic Huygensian differential equations of second order for multicomponent fields *Ukrain. Math. J.* **29** (1977) 351–363
[105] H. Schubert: *Topologie* (Teubner Publ., Stuttgart, 1964)
[106] B.-W. Schulze: *Pseudo–differential operators on manifolds with singularities* (North–Holland Publ., Amsterdam, 1991)
[107] B.-W. Schulze: Operatos with symbol hierarchies and iterated asymptotics *Publications RIMS, Kyoto University* **38(4)** (2002) 735–802
[108] W.-X. Shi: Deforming the metric on complete manifolds *J. Diff. Geom.* **30** (1988) 223–301
[109] W.-X. Shi: Ricci deformation of the metric on complete noncompact Riemannian manifolds *J. Diff. Geom.* **30** (1989) 303–394
[110] M. A. Shubin: Weak Bloch property and weight estimates for elliptic operators *Ecole Polytechnique, Seminaire* (1989–1990)
[111] N. K. Smolentzew: The curvature of diffeomorphism groups and the space of volume elements *Sib. Math. J.* **33** (1992) 135–140
[112] N. K. Smolentzew: The biinvariant metric on the group of symplectic diffeomorphisms and the equation $\frac{\partial}{\partial t}\Delta F = \{\Delta F, F\}$ *Sib. Math. J.* **27** (1986) 150–156
[113] N. K. Smolentzew: The curvature of classical diffeomorphism groups *Sib. Math. J.* **35** (1994) 169–176
[114] R. S. Strichartz: Analysis of the Laplacian on the complete Riemannian manifold *J. Functional Analysis* **52** (1983) 48–79
[115] R. S. Strichartz: L_p–contractive projections and the heat semigroup for differential forms *J. Funct. Analysis* **65** (1986) 348–357
[116] H. C. Taubes: Self–dual connections on non–self–dual manifolds *J. Diff. Geom.* **17** (1982) 139–170
[117] H. C. Taubes: Self–dual connections on 4–manifolds with indefinite intersection matrix *J. Diff. Geom.* **19** (1984) 517–560
[118] A. J. Tromba: *Teichmüller Theory in Riemannian Geometry* (ETH Zürich Lecture Notes, Birkhäuser Publ., Basel, 1992)
[119] N. T. Varopoulos: Brownian motion and random walks on manifolds *Ann. Inst. Fourier* **34** (Grenoble), (1984) 243–269

[120] S. T. Yau: On the heat kernel of a complete Riemannian manifold *J. Math. Pures Appl.* **57** (1978) 191–201

[121] G. Yu: K–theoretic indices of Dirac type operators on complete manifolds and the Roe algebra *K–Theory* **11** (1997) 1–15

[122] S. Zelditch: Lectures on wave invariants *In: Spectral theory and geometry* (Proc. of the ICMS instructional conference, London Mathematical Society Leture Notes Series 273, 1999) 284–328

[123] S. Zucker: L_2–cohomology of warped products and arithmetic groups *Invent. Math.* **70** (1982) 169–218

[124] S. Zucker: L_2–cohomology of locally symmetric manifolds of finite volume *Duke Math. J.* **50** (1983) 625–647

[125] S. Zucker: L_2–cohomology and intersection homology of locally symmetric varieties *In: Proc. Symp. Pure Math.* (40, Part 2, 1983) 675–680

[126] S. Zucker: L_2–cohomology and intersection homology of locally symmetric varieties II *Composito Math.* **59** (1986) 339–398

JÜRGEN EICHHORN

Institut für Mathematik und Informatik, Universität Greifswald,
D–17487 Greifswald, Germany

E-mail: `eichhorn@uni-greifswald.de`

Handbook of Global Analysis
Demeter Krupka and David Saunders

The spectral geometry of operators of Dirac and Laplace type

P. Gilkey

Contents

1 Introduction

The field of spectral geometry is a vibrant and active one. In these brief notes, we will sketch some of the recent developments in this area. Our choice is somewhat idiosyncratic and owing to constraints of space necessarily incomplete. It is impossible to give a complete bibliography for such a survey. We refer Carslaw and Jaeger [41] for a comprehensive discussion of problems associated with heat flow, to Gilkey [54] and to Melrose [91] for a discussion of heat equation methods related to the index theorem, to Gilkey [56] and to Kirsten [84] for a calculation of various heat trace and heat content asymptotic formulas, to Gordon [66] for a survey of isospectral manifolds, to Grubb [73] for a discussion of the pseudo-differential calculus relating to boundary problems, and to Seeley [116] for an introduction to the pseudo-differential calculus. Throughout we shall work with smooth manifolds and, if present, smooth boundaries. We have also given in each section a few additional references to relevant works. The constraints of space have of necessity forced us to omit many more important references than it was possible to include and we apologize in advance for that.

We adopt the following notational conventions. Let (M,g) be a compact Riemannian manifold of dimension m with smooth boundary ∂M. Let Greek indices μ,ν range from 1 to m and index a local system of coordinates $x=(x^1,...,x^m)$ on the interior of M. Expand the metric in the form $ds^2=g_{\mu\nu}dx^\mu\circ dx^\nu$ were $g_{\mu\nu}:=g(\partial_{x_\mu},\partial_{x_\nu})$ and where we adopt the *Einstein convention* of summing over repeated indices. We let $g^{\mu\nu}$ be the inverse matrix. The Riemannian measure is given by $dx:=g\ dx^1...dx^m$ for $g:=\sqrt{\det(g_{\mu\nu})}$.

Let ∇ be the Levi-Civita connection. We expand $\nabla_{\partial_{x_\nu}}\partial_{x_\mu}=\Gamma_{\nu\mu}{}^\sigma\partial_{x_\sigma}$ where $\Gamma_{\nu\mu}{}^\sigma$ are the *Christoffel symbols*. The *curvature operator* $\mathcal{R}$ and corresponding *curvature tensor* R are may then be given by $\mathcal{R}(X,Y):=\nabla_X\nabla_Y-\nabla_Y\nabla_X-\nabla_{[X,Y]}$ and given by $R(X,Y,Z,W):=g(\mathcal{R}(X,Y)Z,W)$.

We shall let Latin indices i,j range from 1 to m and index a local orthonormal frame $\{e_1,...,e_m\}$ for the tangent bundle of M. Let R_{ijkl} be the components of the curvature tensor relative to this base; the *Ricci tensor* ρ and the *scalar curvature* τ are then given by setting $\rho_{ij}:=R_{ikkj}$ and $\tau:=\rho_{ii}=R_{ikki}$. We shall often have an auxiliary vector bundle V and an auxiliary connection given on V. We use this connection and the Levi-Civita connection to covariantly differentiate tensors of all types and we shall let ';' denote the components of multiple covariant differentiation.

Let dy be the measure of the induced metric on the boundary ∂M. We choose a local orthonormal frame near the boundary of M so that e_m is the inward unit normal. We let indices a,b range from 1 to $m-1$ and index the induced local frame $\{e_1,...,e_{m-1}\}$ for the tangent bundle of the boundary. Let $L_{ab}:=g(\nabla_{e_a}e_b,e_m)$ denote the *second fundamental form*. We sum over indices with the implicit range indicated. Thus the geodesic curvature κ_g is given by $\kappa_g:=L_{aa}$. We shall let ':' denote multiple tangential covariant differentiation with respect to the Levi-Civita connection of the boundary; the difference between ';' and ':' being, of course, measured by the second fundamental form.

2 The geometry of operators of Laplace and Dirac type

In this section, we shall establish basic definitions, discuss operators of Laplace and of Dirac type, introduce the DeRham complex, and discuss the Bochner Laplacian and the Weitzenböch formula; [55] provides a good reference for the material of this section.

Let D be a second order partial differential operator on the space of smooth sections $C^\infty(V)$ of a vector bundle V over M. Expand $D=-\{a^{\mu\nu}\partial_{x_\mu}\partial_{x_\nu}+a^\sigma\partial_{x_\sigma}+b\}$ where the coefficients $\{a^{\mu\nu},a^\mu,b\}$ are smooth endomorphisms of V; we suppress the fiber indices. We say that D is an *operator of Laplace type* if $a^{\mu\nu}=g^{\mu\nu}\,\mathrm{id}$. A first order operator A on $C^\infty(V)$ is said to be an operator of *Dirac type* if A^2 is an operator of Laplace type. If we expand $A=\gamma^\nu\partial_{x_\nu}+\gamma_0$, then A is an operator of Dirac type if and only if the endomorphisms γ^ν satisfy the *Clifford commutation relations* $\gamma^\nu\gamma^\mu+\gamma^\mu\gamma^\nu=-2g^{\mu\nu}\,\mathrm{id}$.

Let A be an operator of Dirac type and let $\xi=\xi_\nu dx^\nu$ be a smooth 1-form on M. We let $\gamma(\xi)=\xi_\nu\gamma^\nu$ define a *Clifford module structure* on V; this is independent of the particular coordinate system chosen. We can always choose a fiber metric on V so that γ is skew-adjoint. We can then construct a unitary connection ∇ on V so that $\nabla\gamma=0$. Such a connection is called *compatible*. If ∇ is compatible, we expand $A=\gamma^\nu\nabla_{\partial_{x_\nu}}+\psi_A$; the endomorphism ψ_A is tensorial and does not depend on the particular coordinate system chosen; it does, of course, depend on the particular compatible connection chosen.

2.1 The DeRham complex

The prototypical example is given by the exterior algebra. Let $C^\infty(\Lambda^p M)$ be the space of smooth p forms. Let $d : C^\infty(\Lambda^p M) \to C^\infty(\Lambda^{p+1} M)$ be *exterior differentiation* and let $\delta = d^*$ be the adjoint operator, *interior differentiation*. If ξ is a cotangent vector, let $\mathrm{ext}(\xi) : \omega \to \xi \wedge \omega$ denote exterior multiplication, and let $\mathrm{int}(\xi)$ be the dual, interior multiplication. Let $\gamma(\xi) := \mathrm{ext}(\xi) - \mathrm{int}(\xi)$ define a Clifford module on the exterior algebra $\Lambda(M)$. Since $d + \delta = \gamma(dx^\nu)\nabla_{\partial_{x_\nu}}$, $d + \delta$ is an operator of Diract type. The associated Laplacian $\Delta_M := (d+\delta)^2 = \Delta_M^0 \oplus ... \oplus \Delta_M^p \oplus ... \oplus \Delta_M^m$ decomposes as the direct sum of operators of Laplace type Δ_M^p on the space of smooth p forms $C^\infty(\Lambda^p M)$. One has $\Delta_M^0 = -g^{-1}\partial_{x_\mu} g g^{\mu\nu}\partial_{x_\nu}$.

It is possible to write the p-form valued Laplacian in an invariant form. Extend the Levi-Civita connection to act on tensors of all types. Let $\tilde\Delta_M \omega := -g^{\mu\nu}\omega_{;\mu\nu}$ define the *Bochner* or *reduced Laplacian*. Let $\mathcal{R}$ give the associated action of the curvature tensor. The *Weitzenböck* formula then permits us to express the ordinary Laplacian in terms of the Bochner Laplacian in the form $\Delta_M = \tilde\Delta_M + \frac{1}{2}\gamma(dx^\mu)\gamma(dx^\nu)\mathcal{R}_{\mu\nu}$. This formalism can be applied more generally:

Lemma 2.1 *Let D be an operator of Laplace type on a Riemannian manifold. There exists a unique connection ∇ on V and there exists a unique endomorphism E of V so that $D\phi = -\phi_{;ii} - E\phi$. If we express D locally in the form $D = -\{g^{\mu\nu}\partial_{x_\nu}\partial_{x_\mu} + a^\mu\partial_{x_\mu} + b\}$ then the connection 1-form ω of ∇ and the endomorphism E are given by*

$$\omega_\nu = \tfrac{1}{2}\left(g_{\nu\mu}a^\mu + g^{\sigma\varepsilon}\Gamma_{\sigma\varepsilon\nu}\,\mathrm{id}\right) \quad \textit{and} \quad E = b - g^{\nu\mu}\left(\partial_{x_\nu}\omega_\mu + \omega_\nu\omega_\mu - \omega_\sigma\Gamma_{\nu\mu}{}^\sigma\right) .$$

Let V be equipped with an auxiliary fiber metric. Then D is self-adjoint if and only if ∇ is unitary and E is self-adjoint. We note that if D is the *Spin Laplacian*, then ∇ is the spin connection on the spinor bundle and the Lichnerowicz formula [86] yields, with our sign convention, that $E = -\frac{1}{4}\tau\,\mathrm{id}$ where τ is the scalar curvature.

3 Heat trace asymptotics for closed manifolds

Throughout this section, we shall assume that D is an operator of Laplace type on a closed Riemannian manifold (M, g). We shall discuss the L^2 spectral resolution if D is self-adjoint, define the heat equation, introduce the heat trace and the heat trace asymptotics, present the leading terms in the heat trace asymptotics, and discuss the form valued Laplacian; [41, 54, 116] are good references for the material of this section and other references will be cited as needed.

We suppose that D is self-adjoint. There is then a *complete spectral resolution* of D on $L^2(V)$. This means that we can find a complete orthonormal basis $\{\phi_n\}$ for $L^2(V)$ where the ϕ_n are smooth sections to V which satisfy the equation $D\phi_n = \lambda_n\phi_n$. Let $||_k$ denote the C^k-norm.

Theorem 3.1 *Let $\phi \in L^2(V)$. Expand $\phi = \sum_{n=1}^\infty c_n\phi_n$ in the L^2 sense where one has $c_n := \int_M(\phi, \phi_n)$. If $\phi \in C^\infty(V)$, then this series converges in the C^k topology for any k; $\phi \in C^\infty(V)$ if and only if $\lim_{n\to\infty} n^k c_n < \infty$ for any k. The set of eigenvalues is discrete. Each eigenvalue appears with finite multiplicity and there are only a finite number of negative eigenvalues. If we enumerate the eigenvalues so that $\lambda_1 \le \lambda_2 \le ...$,* *then* $\lambda_n \sim n^{2/m}$ *as* $n \to \infty$. *There exist constants* $\nu_k > 0$ *and* $C_k > 0$ *so that one has norm estimates* $||\phi_n||_k \le C_k n^{\nu_k}$ *for all* k, n.

This yields the familiar *Weyl asymptotic formula* [127] giving the eigenvalue growth. For example, if $D = -\partial_\theta^2$ on the circle, then the eigenvalues grow quadratically since the associated spectral resolution is given by $\{n^2, \frac{1}{\sqrt{2\pi}} e^{in\theta}\}_{n\in\mathbb{Z}}$. The L^2 expansion of Theorem 3.1 in this setting then becomes the usual Fourier series expansion for ϕ and one has the familiar result that a function on the circle is smooth if and only if its Fourier coefficients are rapidly decreasing.

Let the initial temperature distribution be given by $\phi \in L^2(V)$. Impose the classical time evolution for the subsequent temperature distribution without additional heat input:

$$(\partial_t + D)u = 0 \text{ for } t > 0 \quad \text{and} \quad \lim_{t\downarrow 0} u(t,\cdot) = \phi \text{ in } L^2 .$$

Then $u(t,\cdot) = e^{-tD}\phi$ where e^{-tD} is given by the functional calculus. This operator is infinitely smoothing; we have $u(t,x) = \int_M K(t,x,\tilde{x})\phi(\tilde{x})d\tilde{x}$ for a smooth kernel function K. If D is self-adjoint, let $\{\lambda_n, \phi_n\}$ be a spectral resolution of D. Then

$$K(t,x,\tilde{x}) := \sum_n e^{-t\lambda_n}\phi_n(x) \otimes \phi_n(\tilde{x}) : V_{\tilde{x}} \to V_x .$$

Theorem 3.1 implies this series converges uniformly in the C^k topology for $t \ge \varepsilon > 0$.

Let $F \in C^\infty(\mathrm{End}(V))$ be an auxiliary endomorphism used for localizing; F is often referred to as a *smearing endomorphism*. The localized heat trace $\mathrm{Tr}_{L^2}\left\{Fe^{-tD}\right\}$ is analytic for $t > 0$. As $t \downarrow 0$, there is a complete asymptotic expansion [117]

$$\mathrm{Tr}_{L^2}\left\{Fe^{-tD}\right\} \sim \sum_{n=0}^{\infty} a_n(F,D)t^{(n-m)/2} .$$

The coefficients $a_n(F,D)$ are the *heat trace asymptotics*; $a_n(F,D) = 0$ if n is odd. In Section 5 we will consider manifolds with boundary and the corresponding invariants are non-trivial for n both even and odd. There exist locally computable endomorphisms $e_n^M(D)(x)$ of V which are defined for all $x \in M$ so that

$$a_n(F,D) = \int_M \mathrm{Tr}\{Fe_n^M(D)\}(x)dx . \tag{3.a}$$

The invariants $e_n^M(D)$ are uniquely characterized by Equation (3.a).

We use Lemma 2.1 to express $D = D(g,\nabla,E)$ where ∇ is a uniquely defined connection on V and where E is a uniquely defined auxiliary endomorphism of V. Let Ω_{ij} be the endomorphism valued components of the curvature defined by the connection ∇.

Theorem 3.2 *Let* $F \in C^\infty(\mathrm{End}(V))$ *be a smearing endomorphism.*

(1) $a_0(F,D) = (4\pi)^{-m/2}\int_M \mathrm{Tr}\{F\}dx.$

(2) $a_2(F,D) = (4\pi)^{-m/2}\frac{1}{6}\int_M \mathrm{Tr}\{F(6E + \tau\,\mathrm{id})\}dx.$

(3) $a_4(F,D) = (4\pi)^{-m/2}\frac{1}{360}\int_M \mathrm{Tr}\{F(60E_{;kk} + 60\tau E + 180E^2$
$+12\tau_{;kk}\,\mathrm{id} + 5\tau^2\,\mathrm{id} - 2|\rho|^2\,\mathrm{id} + 2|R|^2\,\mathrm{id} + 30\Omega_{ij}\Omega_{ij})\}dx.$

(4) $a_6(F,D)=\int_M \mathrm{Tr}\{F((\frac{18}{7!}\tau_{;iijj}+\frac{17}{7!}\tau_{;k}\tau_{;k}-\frac{2}{7!}\rho_{ij;k}\rho_{ij;k}-\frac{4}{7!}\rho_{jk;n}\rho_{jn;k}$
$+\frac{9}{7!}R_{ijkl;n}R_{ijkl;n}+\frac{28}{7!}\tau\tau_{;nn}-\frac{8}{7!}\rho_{jk}\rho_{jk;nn}+\frac{24}{7!}\rho_{jk}\rho_{jn;kn}+\frac{12}{7!}R_{ijkl}R_{ijkl;nn}$
$+\frac{35}{9\cdot 7!}\tau^3-\frac{14}{3\cdot 7!}\tau|\rho|^2+\frac{14}{3\cdot 7!}\tau|R|^2-\frac{208}{9\cdot 7!}\rho_{jk}\rho_{jn}\rho_{kn}-\frac{64}{3\cdot 7!}\rho_{ij}\rho_{kl}R_{ikjl}$
$-\frac{16}{3\cdot 7!}\rho_{jk}R_{jnli}R_{knli}-\frac{44}{9\cdot 7!}R_{ijkn}R_{ijlp}R_{knlp}-\frac{80}{9\cdot 7!}R_{ijkn}R_{ilkp}R_{jlnp})\,\mathrm{id}$
$+\frac{1}{45}\Omega_{ij;k}\Omega_{ij;k}+\frac{1}{180}\Omega_{ij;j}\Omega_{ik;k}+\frac{1}{60}\Omega_{ij;kk}\Omega_{ij}+\frac{1}{60}\Omega_{ij}\Omega_{ij;kk}-\frac{1}{30}\Omega_{ij}\Omega_{jk}\Omega_{ki}$
$-\frac{1}{60}R_{ijkn}\Omega_{ij}\Omega_{kn}-\frac{1}{90}\rho_{jk}\Omega_{jn}\Omega_{kn}+\frac{1}{72}\tau\Omega_{kn}\Omega_{kn}+\frac{1}{60}E_{;iijj}+\frac{1}{12}EE_{;ii}$
$+\frac{1}{12}E_{;ii}E+\frac{1}{12}E_{;i}E_{;i}+\frac{1}{6}E^3+\frac{1}{30}E\Omega_{ij}\Omega_{ij}+\frac{1}{60}\Omega_{ij}E\Omega_j+\frac{1}{30}\Omega_{ij}\Omega_{ij}E$
$+\frac{1}{36}\tau E_{;kk}+\frac{1}{90}\rho_{jk}E_{;jk}+\frac{1}{30}\tau_{;k}E_{;k}-\frac{1}{60}E_{;j}\Omega_{ij;i}+\frac{1}{60}\Omega_{ij;i}E_{;j}+\frac{1}{12}EE\tau$
$+\frac{1}{30}E\tau_{;kk}+\frac{1}{72}E\tau^2-\frac{1}{180}E|\rho|^2+\frac{1}{180}E|R|^2)\}dx.$

There are formulas available for a_8 and a_{10}; we refer to Amsterdamski, Berkin, and O'Connor[1], to Avramidi [9], and to van de Ven [124] for further details.

There is also information available about the general form of the heat trace asymptotics a_n for all values of n; we refer to Avramidi [10] and to Branson et al. [36] for further details. These formulas play an important role in the compactness results we shall discuss presently in Theorem 4.6. Let D be an operator of Laplace type on a closed Riemannian manifold M. Let $\Delta E=-E_{;kk}$. Set $\epsilon_n=(-1)^n/\{2^{n+1}\cdot 1\cdot 3\cdot\ldots\cdot(2n+1)\}$.

Theorem 3.3 *Let '+...' denote lower order terms.*

(1) *If* $n\ge 1$, *then* $a_{2n}(F,D)=\epsilon_n(4\pi)^{-m/2}\int_M \mathrm{Tr}\{F(-(8n+4)\Delta^{n-1}E$
$-2n\Delta^{n-1}\tau\,\mathrm{id}+\ldots\}dx.$

(2) *If* $n\ge 3$, *then* $a_{2n}(D)=\epsilon_n(4\pi)^{-m/2}\,\mathrm{Tr}\{(n^2-n-1)|\nabla^{n-2}\tau|^2\,\mathrm{id}$
$+2|\nabla^{n-2}\rho|^2\,\mathrm{id}+4(2n+1)(n-1)\nabla^{n-2}\tau\cdot\nabla^{n-2}E$
$+2(2n+1)\nabla^{n-2}\Omega\cdot\nabla^{n-2}\Omega+4(2n+1)(2n-1)\nabla^{n-2}E\cdot\nabla^{n-2}E+\ldots\}dx.$

We note that Polterovich [109, 110] has introduced a formalism for computing in closed form the heat trace asymptotics a_n for all n.

If one specializes these formulas for a_0, a_2, and a_4 to the case in which D is the form valued Laplacian, one has the following result of Patodi [106]. Introduce constants:

$$\begin{aligned}
&c(m,p)=\tfrac{m!}{p!(m-p)!},\\
&c_0(m,p)=c(m,p)-6c(m-2,p-1),\\
&c_1(m,p)=5c(m,p)-60c(m-2,p-1)+180c(m-4,p-2),\\
&c_2(m,p)=-2c(m,p)+180c(m-2,p-1)-720c(m-4,p-2),\\
&c_3(m,p)=2c(m,p)-30c(m-2,p-1)+180c(m-4,p-2)\,.
\end{aligned}$$

Theorem 3.4 (1) $a_0(\Delta_M^p)=(4\pi)^{-m/2}c(m,p)\,\mathrm{Vol}(M)$.

(2) $a_2(\Delta_M^p)=(4\pi)^{-m/2}\frac{1}{6}c_0(m,p)\int_M\tau dx.$

(3) $a_4(\Delta_M^p)=(4\pi)^{-m/2}\frac{1}{360}\int_M\{c_1(m,p)\tau^2+c_2(m,p)\rho^2+c_3(m,p)R^2\}dx.$

Such formulas play an important role in the study of spectral geometry. There is a long history involved in computing these invariants. Weyl [127] discovered the leading term in the asymptotic expansion, a_0. Minakshisundaram and Pleijel [93, 94] examined the asymptotic expansion for the scalar Laplacian in some detail. The a_2 and a_4 terms in the asymptotic expansion were investigated by McKean and Singer [90] in the scalar case and by Patodi [105] for the form valued Laplacian. The a_6 term for the scalar Laplacian was determined by Sakai [111] and the general expression for a_2, a_4 and a_6 for arbitrary operators of Laplace type was worked out in [53]. As noted above, there are formulas for a_8 and a_{10}. The literature is a vast one and we refer to [54, 56] more details and additional references.

We now discuss the relationship between the heat trace asyptotics and the eta and zeta functions in a quite general context. Let P be a positive, self-adjoint elliptic partial differential operator on a closed Riemannian manifold M. Then e^{-tP} is an infinitely smoothing operator which is given by a smooth kernel function. Let Q be an auxiliary partial differential operator. Then $\mathrm{Tr}_{L^2}\{Qe^{-tP}\}$ is analytic for $t > 0$ and as $t \downarrow 0$, there is a complete asymptotic expansion with locally computable coefficients:

$$\mathrm{Tr}_{L^2}\{Qe^{-tP}) \sim \sum_{n=0}^{\infty} a_n(P,Q)t^{(n-m-\mathrm{ord}(Q))/\,\mathrm{ord}(P)}$$

The generalized zeta function is given by:

$$\zeta(s,P,Q) := \mathrm{Tr}_{L^2}(QP^{-s}) \quad \text{for} \quad \Re(s) >> 0\,.$$

The Mellin transform may be used to relate the zeta function to the heat kernel. Let Γ be the classical Gamma function. We refer to Seeley [116, 117] for the proof of Assertions (1) and (2) and to [50] for the proof of Assertion (3) in the following result. Assertion (2) generalizes eigenvalue growth estimates of Weyl [127] given previously in Theorem 3.1.

Theorem 3.5 (1) *If* $\mathrm{Re}(s) >> 0$, *then* $\zeta(s,P,Q) = \Gamma(s)^{-1}\int_0^\infty t^{s-1}\,\mathrm{Tr}_{L^2}(Qe^{-tP})dt$. $\Gamma(s)\zeta(s,P,Q)$ *has a meromorphic extension to the complex plane with isolated simple poles at* $s = (m + \mathrm{ord}(Q) - n)/\,\mathrm{ord}(P)$ *for* $n = 0, 1, \ldots$ *and*

$$\mathrm{Res}_{s=(m+\mathrm{ord}(Q)-n)/\,\mathrm{ord}(P)}\,\Gamma(s)\zeta(s,P,Q) = a_n(P,Q).$$

(2) *The leading heat trace coefficient* $a_0(P)$ *is non-zero. Let* $\lambda_1 \le \ldots \le \lambda_n \le \ldots$ *be the eigenvalues of* P. *Then* $\lim_{n\to\infty} n\lambda_n^{-m/\,\mathrm{ord}(P)} = \Gamma(\frac{m}{\mathrm{ord}(P)})^{-1}a_0(P)$.

(3) *Let* $A(t)$ *and* $B(t)$ *be polynomials of degree* $a \ge 0$ *and* $b > 0$ *where* B *is monic. There are constants so* $a_n(B(P),A(P)) = \Sigma_{k\le k(n)}c(k,n,m,A,B)a_k(P)$.

4 Hearing the shape of a drum

Let $\mathrm{Spec}(D) = \{\lambda_1 \le \lambda_2 \le \ldots\}$ denote the set of eigenvalues of a self-adjoint operator of Laplace type, repeated according to multiplicity. One is interested in what geometric and topological properties of M are reflected by the spectrum. Good references for this section are [26, 54, 66]; other references will be cited as appropriate.

One says that M and $\tilde{M}$ are *isospectral* if $\text{Spec}(\Delta^0_M) = \text{Spec}(\Delta^0_{\tilde{M}})$; p-*isospectral* refers to Δ^p. M. Kac [81] in his seminal article raised the question of determining the geometry, at least in part, of the underlying manifold from the spectrum of the scalar Laplace operator Δ^0_M. It is not possible in general to completely determine the geometry:

Theorem 4.1 (1) *Milnor [92]: There exist isospectral non isometric flat tori of dimension 16.*

(2) *Vigneras [125]: There exist isospectral non-isometric hyperbolic Riemann surfaces. Furthermore, if $m \geq 3$, there exist isospectral hyperbolic manifolds with different fundamental groups.*

(3) *Ikeda [79]: There exist isospectral non-isometric spherical space forms.*

(4) *Urakawa [123]: There exist regions Ω_i in flat space for $m \geq 4$ which are isospectral for the Laplacian with both Dirichlet and Neumann boundary conditions but which are not isometric.*

These examples listed above come in finite families. We say that a family of metrics g_t on M is a non-trivial family of isospectral manifolds if (M, g_t) and (M, g_s) are isospectral for every s, t, but (M, g_t) is not isometric to (M, g_s) for $s \neq t$.

Theorem 4.2 (1) *Gordon-Wilson [67]: There exists a non-trivial family of isospectral metrics on a smooth manifold M which are not conformally equivalent.*

(2) *Brooks-Gordon [37]: There exists a non-trivial family of isospectral metrics on a smooth manifold M which are conformally equivalent.*

There is a vast literature in the subject. In particular, Sunada [121] gave a general method for attacking the problem which has been exploited by many authors.

Despite this somewhat discouraging prospect, there are a number of positive results available. For example Berger [27] and Tanno [122] showed that a sphere or projective space is characterized by its spectral geometry, at least in low dimensions:

Theorem 4.3 *Let $\mathcal{M}_i$ and $\mathcal{M}_2$ be closed Riemannian manifolds of dimension $m \leq 6$ which are isospectral. If M_1 has constant sectional curvature c, so does M_2.*

Patodi [106] showed additional geometrical properties are determined by the form valued Laplacian. The following is an easy consequence of Theorem 3.4.

Theorem 4.4 *Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be closed Riemannian manifolds which are p-isospectral for $p = 0, 1, 2$. Then:*

(1) *If M_1 has constant scalar curvature $\tau = c$, then so does M_2.*

(2) *If M_1 is Einstein, so is M_2.*

(3) *If M_1 has constant sectional curvature c, then so does M_2.*

For manifolds with boundary, suitable boundary conditions must be imposed. Formulas that will be discussed presently in Section 5 have been used by Park [104] to show:

Theorem 4.5 *Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be compact Einstein Riemannian manifolds with smooth boundaries with the same constant scalar curvatures $\tau_{\mathcal{M}_1} = \tau_{\mathcal{M}_2}$. Also assume that $\mathcal{M}_1$ and $\mathcal{M}_2$ are isospectral for both Neumann and Dirichlet boundary conditions. Then:*

(1) *If $\mathcal{M}_1$ has totally geodesic boundary, then so does $\mathcal{M}_2$.*

(2) *If $\mathcal{M}_1$ has minimal boundary, then so does $\mathcal{M}_2$.*

(3) *If $\mathcal{M}_1$ has totally umbillic boundary, then so does $\mathcal{M}_2$.*

(4) *If $\mathcal{M}_1$ has strongly totally umbillic boundary, then so does $\mathcal{M}_2$.*

There are also a number of compactness results. Theorem 3.3 plays a central role in the following results:

Theorem 4.6 (1) *Osgood, Phillips, and Sarnak [102]: Families of isospectral metrics on Riemann surfaces are compact modulo gauge equivalence.*

(2) *Brooks, Perry, and Yang [39] and Chang and Yang [42]: If $m = 3$, then families of isospectral metrics within a conformal class are compact modulo gauge equivalence.*

(3) *Brooks, Perry, and Petersen [38]: Isospectral negative curvature manifolds contain only a finite number of topological types.*

5 Heat trace asymptotics of manifolds with boundary

In previous sections, we have concentrated on closed Riemannian manifolds. Let D be an operator of Laplace type on a compact Riemannian manifold M with smooth boundary ∂M. Good basic references for the material of this section are [56, 73, 84]. Many authors have contributed to the material discussed here; we refer in particular to the work of [40, 76, 78, 82, 83, 90, 93, 94, 96, 120, 127].

We impose suitable boundary conditions $\mathcal{B}$ to have a well posed problem; $\mathcal{B}$ must satisfy a condition called the *strong Lopatenski-Shapiro condition*. We shall suppress technical details for the most part in the interests of simplicity. The boundary conditions we shall consider have physical underpinnings. Dirichlet boundary conditions correspond to immersing the boundary in ice water; Neumann boundary conditions correspond to an insulated boundary. Robin boundary conditions are a generalization of Neumann boundary conditions where the heat flow across the boundary is proportional to the temperature on the boundary. Transmission boundary conditions arise in the study of heat conduction problems between closely coupled membranes. Transfer boundary conditions arise in the study of branes. Both these conditions reflect the heat flow between two inhomogeneous mediums coupled along a common boundary or brane. Transmission boundary conditions correspond to having the two components pressed tightly together. By contrast, heat transfer boundary conditions correspond to a loose coupling between the two components. We refer to Carslaw and Jaeger [41] for further details.

Through out the remainder of this section, we let $F \in C^\infty(\text{End}(V))$ define a localizing or smearing endomorphism and let $\mathcal{B}$ denote a suitable boundary operator; in what follows, we shall give a number of examples. Let $D_\mathcal{B}$ be the realization of an operator D of Laplace type with respect $\mathcal{B}$; the domain of $D_\mathcal{B}$ is then the set of all functions ϕ in a suitable Schwarz space so that ϕ satisfies the appropriate boundary conditions, i.e. so that $\mathcal{B}\phi = 0$. Greiner [68, 69] and Seeley [118, 119] showed that there was a full asymptotic expansion as $t \downarrow 0$ of the form:

$$\mathrm{Tr}_{L^2}\{Fe^{-tD_{\mathcal{B}}}\} \sim \sum_{n=0}^{\infty} a_n(F,D,\mathcal{B})t^{(n-m)/2}\,.$$

There are locally computable endomorphisms $e_n(D)(x)$ defined on the interior and locally computable endomorphisms $e_{n,k}^{\partial M}(D,\mathcal{B})(y)$ defined on the boundary so that

$$\begin{aligned} a_n(F,D,\mathcal{B}) &= \int_M \mathrm{Tr}\{Fe_n^M(D)\}(x)dx \\ &+ \sum_{k=0}^{n-1}\int_{\partial M} \mathrm{Tr}\{(\nabla_{e_m}^k F)e_{n,k}^{\partial M}(D,\mathcal{B})\}(y)dy\,. \end{aligned}$$

The invariants $e_n^M(D)$ and $e_{n,k}^{\partial M}(D,\mathcal{B})$ are uniquely characterized by this identity; the interior invariants $e_n^M(D)$ are not sensitive to the boundary condition and agree with those considered previously in Equation (3.a). The remainder of Section 5 is devoted to giving explicit combinatorial formulas for these invariants.

A function ϕ satisfies Dirichlet boundary conditions if ϕ vanishes on ∂M. Thus the Dirichlet boundary operator is defined by:

$$\mathcal{B}\phi := \phi|_{\partial M}\,. \tag{5.a}$$

Theorem 5.1 *[Dirichlet boundary conditions] Let* $F \in C^\infty(\mathrm{End}(V))$.

(1) $a_0(F,D,\mathcal{B}) = (4\pi)^{-m/2}\int_M \mathrm{Tr}\{F\}dx$.

(2) $a_1(F,D,\mathcal{B}) = -(4\pi)^{-(m-1)/2}\frac{1}{4}\int_{\partial M}\mathrm{Tr}\{F\}dy$.

(3) $a_2(F,D,\mathcal{B}) = (4\pi)^{-m/2}\frac{1}{6}\int_M \mathrm{Tr}\{F(6E+\tau)\}dx + (4\pi)^{-m/2}\frac{1}{6}\int_{\partial M}\mathrm{Tr}\{2FL_{aa} -3F_{;m}\}dy$.

(4) $a_3(F,D,\mathcal{B}) = -\frac{1}{384}(4\pi)^{-(m-1)/2}\int_{\partial M}\mathrm{Tr}\{96FE + F(16\tau + 8R_{amam} + 7L_{aa}L_{bb} -10L_{ab}L_{ab}) - 30F_{;m}L_{aa} + 24F_{;mm}\}dy$.

(5) $a_4(F,D,\mathcal{B}) = (4\pi)^{-m/2}\frac{1}{360}\int_M \mathrm{Tr}\{F(60E_{;kk} + 60\tau E + 180E^2 + 30\Omega^2 + 12\tau_{;kk} +5\tau^2 - 2|\rho^2| + 2|R^2|)\}dx + (4\pi)^{-m/2}\frac{1}{360}\int_{\partial M}\mathrm{Tr}\{F(-120E_{;m} + 120EL_{aa} -18\tau_{;m} + 20\tau L_{aa} + 4R_{amam}L_{bb} - 12R_{ambm}L_{ab} + 4R_{abcb}L_{ac} + 24L_{aa:bb} +\frac{40}{21}L_{aa}L_{bb}L_{cc} - \frac{88}{7}L_{ab}L_{ab}L_{cc} + \frac{320}{21}L_{ab}L_{bc}L_{ac}) + F_{;m}(-180E - 30\tau -\frac{180}{7}L_{aa}L_{bb} + \frac{60}{7}L_{ab}L_{ab}) + 24F_{;mm}L_{aa} - 30F_{;iim}\}dy$.

Neumann boundary conditions are defined by the operator $\mathcal{B}_N\phi := \phi_{;m}|_{\partial M}$; the associated boundary conditions define a perfectly insulated boundary with no heat flow across the boundary. It is convenient in many applications to consider slightly more general conditions called Robin boundary conditions that permit the heat flow to be proportional to the temperature. Let S be an auxiliary endomorphism of V over ∂M. The Robin boundary operator is defined by:

$$\mathcal{B}_S\phi := (\phi_{;m} + S\phi)|_{\partial M}\,. \tag{5.b}$$

Theorem 5.2 *[Robin boundary conditions] Let* $F \in C^\infty(\text{End}(V))$.

(1) $a_0(F,D,\mathcal{B}_S) = (4\pi)^{-m/2}\int_M \text{Tr}\{F\}dx$.

(2) $a_1(F,D,\mathcal{B}_S) = (4\pi)^{(1-m)/2}\frac{1}{4}\int_{\partial M}\text{Tr}\{F\}dy$.

(3) $a_2(F,D,\mathcal{B}_S) = (4\pi)^{-m/2}\frac{1}{6}\int_M \text{Tr}\{F(6E+\tau)\}dx + (4\pi)^{-m/2}\frac{1}{6}\int_{\partial M}\text{Tr}\{F(2L_{aa} +12S) + 3F_{;m}\}dy$.

(4) $a_3(F,D,\mathcal{B}_S) = (4\pi)^{(1-m)/2}\frac{1}{384}\int_{\partial M}\text{Tr}\{F(96E + 16\tau + 8R_{amam} + 13L_{aa}L_{bb} +2L_{ab}L_{ab} + 96SL_{aa} + 192S^2 + F_{;m}(6L_{aa} + 96S) + 24F_{;mm}\}dy$.

(5) $a_4(F,D,\mathcal{B}_S) = (4\pi)^{-m/2}\frac{1}{360}\int_M \text{Tr}\{F(60E_{;kk} + 60\tau E + 180E^2 + 30\Omega^2 + 12\tau_{;kk} +5\tau^2 - 2|\rho|^2 + 2|R|^2)\}dx + (4\pi)^{-m/2}\frac{1}{360}\int_{\partial M}\text{Tr}\{F(240E_{;m} + 42\tau_{;m} +24L_{aa:bb} + 120EL_{aa} + 20\tau L_{aa} + 4R_{amam}L_{bb} - 12R_{ambm}L_{ab} + 4R_{abcb}L_{ac} +\frac{40}{3}L_{aa}L_{bb}L_{cc} + 8L_{ab}L_{ab}L_{cc} + \frac{32}{3}L_{ab}L_{bc}L_{ac} + 360(SE + ES) + 120S\tau +144SL_{aa}L_{bb} + 48SL_{ab}L_{ab} + 480S^2L_{aa} + 480S^3 + 120S_{:aa}) + F_{;m}(180E +30\tau + 12L_{aa}L_{bb} + 12L_{ab}L_{ab} + 72SL_{aa} + 240S^2) + F_{;mm}(24L_{aa} + 120S) +30F_{;iim}\}dy$.

When discussing the Euler characteristic of a manifold with boundary in Section 6 subsequently, it will useful to consider absolute and relative boundary conditions. Let r be the geodesic distance to the boundary. Near the boundary, decompose a differential form $\omega \in C^\infty(\Lambda(M))$ in the form $\omega = \omega_1 + dr \wedge \omega_2$ where the ω_i are tangential differential forms. We define the relative boundary operator $\mathcal{B}_r$ and the absolute boundary operator $\mathcal{B}_a$ for the operator $d + \delta$ by setting:

$$\mathcal{B}_r(\omega) = \omega_1|_{\partial M} \text{ and } \mathcal{B}_a(\omega) = \omega_2|_{\partial M}. \tag{5.c}$$

There are induced boundary conditions for the associated Laplacian $(d + \delta)^2$. They are defined by the operator $\bar{\mathcal{B}}_{r/a}\phi := \mathcal{B}_{r/a}\phi \oplus \mathcal{B}_{r/a}(d + \delta)\phi$.

The boundary conditions defined by the operators $\bar{\mathcal{B}}_{r/a}$ provide examples of a more general boundary condition which are called *mixed boundary conditions*. We can combine Theorems 5.1 and 5.2 into a single result by using such boundary conditions. We assume given a decomposition $V|_{\partial M} = V_+ \oplus V_-$. Extend the bundles $V_\pm$ to a collared neighborhood of ∂M by parallel translation along the inward unit geodesic rays. Set $\chi := \Pi_+ - \Pi_-$. Let S be an auxiliary endomorphism of V_+ over ∂M. The *mixed boundary operator* may then be defined by setting

$$\mathcal{B}_{\chi,S}\phi := \Pi_+(\phi_{;m} + S\phi)|_{\partial M} \oplus \Pi_-\phi|_{\partial M}\,. \tag{5.d}$$

One sets $\chi = \text{id}$, $\Pi_+ = \text{id}$, and $\Pi_- = 0$ to obtain the Robin boundary operator of Equation (5.b); one sets $\chi = -\,\text{id}$, $\Pi_+ = 0$, and $\Pi_- = \text{id}$ to obtain the Dirichlet boundary operator of Equation (5.a). The formulas of Theorem 5.1 and Theorem 5.2 then be obtained by this specialization.

Theorem 5.3 *[Mixed boundary conditions] Let* $F = f\,\text{id}$ *for* $f \in C^\infty(M)$. *Then:*

(1) $a_0(F,D,\mathcal{B}_{\chi,S})=(4\pi)^{-m/2}\int_M \mathrm{Tr}\{F\}dx.$

(2) $a_1(F,D,\mathcal{B}_{\chi,S})=(4\pi)^{-(m-1)/2}\frac{1}{4}\int_{\partial M}\mathrm{Tr}\{F\chi\}dy.$

(3) $a_2(F,D,\mathcal{B}_{\chi,S})=(4\pi)^{-m/2}\frac{1}{6}\int_M \mathrm{Tr}\{F(6E+\tau)\}dx$
$+(4\pi)^{-m/2}\frac{1}{6}\int_{\partial M}\mathrm{Tr}\{2FL_{aa}+3F_{;m}\chi+12FS\}dy.$

(4) $a_3(F,D,\mathcal{B}_{\chi,S})=(4\pi)^{-(m-1)/2}\frac{1}{384}\int_{\partial M}\mathrm{Tr}\{F(96\chi E+16\chi\tau+8\chi R_{amam}$
$+[13\Pi_+-7\Pi_-]L_{aa}L_{bb}+[2\Pi_++10\Pi_-]L_{ab}L_{ab}+96SL_{aa}+192S^2$
$-12\chi_{:a}\chi_{:a})+F_{;m}([6\Pi_++30\Pi_-]L_{aa}+96S)+24\chi F_{;mm}\}dy.$

(5) $a_4(F,D,\mathcal{B}_{\chi,S})=(4\pi)^{-m/2}\frac{1}{360}\int_M \mathrm{Tr}\{F(60E_{;kk}+60\tau E+180E^2$
$+30\Omega^2+12\tau_{;kk}+5\tau^2-2|\rho|^2+2|R|^2)\}dx+(4\pi)^{-m/2}\frac{1}{360}\int_M \mathrm{Tr}\{F([240\Pi_+$
$-120\Pi_-]E_{;m}+[42\Pi_+-18\Pi_-]\tau_{;m}+120EL_{aa}+24L_{aa:bb}+20\tau L_{aa}$
$+4R_{amam}L_{bb}-12R_{ambm}L_{ab}+4R_{abcb}L_{ac}+720ES+120S\tau+[\frac{280}{21}\Pi_+$
$+\frac{40}{21}\Pi_-]L_{aa}L_{bb}L_{cc}+[\frac{168}{21}\Pi_+-\frac{264}{21}\Pi_-]L_{ab}L_{ab}L_{cc}+[\frac{224}{21}\Pi_++\frac{320}{21}\Pi_-]$
$\times L_{ab}L_{bc}L_{ac}+144SL_{aa}L_{bb}+48SL_{ab}L_{ab}+480S^2L_{aa}+480S^3+120S_{:aa}$
$+60\chi\chi_{:a}\Omega_{am}-12\chi_{:a}\chi_{:a}L_{bb}-24\chi_{:a}\chi_{:b}L_{ab}-120\chi_{:a}\chi_{:a}S)+F_{;m}(180\chi E$
$+30\chi\tau+[\frac{84}{7}\Pi_+-\frac{180}{7}\Pi_-]L_{aa}L_{bb}+240S^2+[\frac{84}{7}\Pi_++\frac{60}{7}\Pi_-]L_{ab}L_{ab}$
$+72SL_{aa}-18\chi_{:a}\chi_{:a})+F_{;mm}(24L_{aa}+120S)+30F_{;iim}\chi\}dy.$

(6) $a_5(F,D,\mathcal{B}_{\chi,S})=(4\pi)^{-(m-1)/2}\frac{1}{5760}\int_{\partial M}\mathrm{Tr}\{F\{360\chi E_{;mm}+1440E_{;m}S$
$+720\chi E^2+240\chi E_{:aa}+240\chi\tau E+48\chi\tau_{;ii}+20\chi\tau^2-8\chi\rho_{ij}\rho_{ij}$
$+8\chi R_{ijkl}R_{ijkl}-120\chi\rho_{mm}E-20\chi\rho_{mm}\tau+480\tau S^2+12\chi\tau_{;mm}$
$+24\chi\rho_{mm:aa}+15\chi\rho_{mm;mm}+270\tau_{;m}S+120\rho_{mm}S^2+960SS_{:aa}$
$+16\chi R_{ammb}\rho_{ab}-17\chi\rho_{mm}\rho_{mm}-10\chi R_{ammb}R_{ammb}+2880ES^2$
$+1440S^4+(90\Pi_++450\Pi_-)L_{aa}E_{;m}+(\frac{111}{2}\Pi_++42\Pi_-)L_{aa}\tau_{;m}$
$+30\Pi_+L_{ab}R_{ammb;m}+240L_{aa}S_{:bb}+420L_{ab}S_{:ab}+390L_{aa:b}S_{:b}$
$+480L_{ab:a}S_{:b}+420L_{aa:bb}S+60L_{ab:ab}S+(\frac{487}{16}\Pi_++\frac{413}{16}\Pi_-)L_{aa:b}L_{cc:b}$
$+(238\Pi_+-58\Pi_-)L_{ab:a}L_{cc:b}+(\frac{49}{4}\Pi_++\frac{11}{4}\Pi_-)L_{ab:a}L_{bc:c}$
$+(\frac{535}{8}\Pi_+-\frac{355}{8}\Pi_-)L_{ab:c}L_{ab:c}+(\frac{151}{4}\Pi_++\frac{29}{4}\Pi_-)L_{ab:c}L_{ac:b}$
$+(111\Pi_+-6\Pi_-)L_{aa:bb}L_{cc}+(-15\Pi_++30\Pi_-)L_{ab:ab}L_{cc}$
$+(-\frac{15}{2}\Pi_++\frac{75}{2}\Pi_-)L_{ab:ac}L_{bc}+(\frac{945}{4}\Pi_+-\frac{285}{4}\Pi_-)L_{aa:bc}L_{bc}$
$+(114\Pi_+-54\Pi_-)L_{bc:aa}L_{bc}+1440L_{aa}SE+30L_{aa}S\rho_{mm}+240L_{aa}S\tau$
$-60L_{ab}\rho_{ab}S+180L_{ab}SR_{ammb}+(195\Pi_+-105\Pi_-)L_{aa}L_{bb}E$
$+(30\Pi_++150\Pi_-)L_{ab}L_{ab}E+(\frac{195}{6}\Pi_+-\frac{105}{6}\Pi_-)L_{aa}L_{bb}\tau$
$+(5\Pi_++25\Pi_-)L_{ab}L_{ab}\tau+(-\frac{275}{16}\Pi_++\frac{215}{16}\Pi_-)L_{aa}L_{bb}\rho_{mm}$
$+(-\frac{275}{8}\Pi_++\frac{215}{8}\Pi_-)L_{ab}L_{ab}\rho_{mm}+(-\Pi_+-14\Pi_-)L_{cc}L_{ab}\rho_{ab}$
$+(\frac{109}{4}\Pi_+-\frac{49}{4}\Pi_-)L_{cc}L_{ab}R_{ammb}+16\chi L_{ab}L_{ac}\rho_{bc}$

$$
\begin{aligned}
&+(\tfrac{133}{2}\Pi_+ + \tfrac{47}{2}\Pi_-)L_{ab}L_{ac}R_{bmmc} - 32\chi L_{ab}L_{cd}R_{acbd}\\
&+\tfrac{315}{2}L_{cc}L_{ab}L_{ab}S + (\tfrac{2041}{128}\Pi_+ + \tfrac{65}{128}\Pi_-)L_{aa}L_{bb}L_{cc}L_{dd}\\
&+150L_{ab}L_{bc}L_{ac}S + (\tfrac{417}{32}\Pi_+ + \tfrac{141}{32}\Pi_-)L_{cc}L_{dd}L_{ab}L_{ab}\\
&+1080L_{aa}L_{bb}S^2 + 360L_{ab}L_{ab}S^2 + (\tfrac{375}{32}\Pi_+ - \tfrac{777}{32}\Pi_-)L_{ab}L_{ab}L_{cd}L_{cd}\\
&+\tfrac{885}{4}L_{aa}L_{bb}L_{cc}S + (25\Pi_+ - \tfrac{17}{2}\Pi_-)L_{dd}L_{ab}L_{bc}L_{ac} + 2160L_{aa}S^3\\
&+(\tfrac{231}{8}\Pi_+ + \tfrac{327}{8}\Pi_-)L_{ab}L_{bc}L_{cd}L_{da} - 180E^2 + 180\chi E\chi E - 120S_{:a}S_{:a}\\
&+720\chi S_{:a}S_{:a} - \tfrac{105}{4}\Omega_{ab}\Omega_{ab} + 120\chi\Omega_{ab}\Omega_{ab} + \tfrac{105}{4}\chi\Omega_{ab}\chi\Omega_{ab} - 45\Omega_{am}\Omega_{am}\\
&+180\chi\Omega_{am}\Omega_{am} - 45\chi\Omega_{am}\chi\Omega_{am} + 360(\Omega_{am}\chi S_{:a} - \Omega_{am}S_{:a}\chi)\\
&+45\chi\chi_{:a}\Omega_{am}L_{cc} - 180\chi_{:a}\chi_{:b}\Omega_{ab} + 90\chi\chi_{:a}\chi_{:b}\Omega_{ab} + 90\chi\chi_{:a}\Omega_{am;m}\\
&+120\chi\chi_{:a}\Omega_{ab:b} + 180\chi\chi_{:a}\Omega_{bm}L_{ab} + 300\chi_{:a}E_{:a} - 180\chi_{:a}\chi_{:a}E - 90\chi\chi_{:a}\chi_{:a}E\\
&+240\chi_{:aa}E - 30\chi_{:a}\chi_{:a}\tau - 60\chi_{:a}\chi_{:b}\rho_{ab} + 30\chi_{:a}\chi_{:b}R_{mabm} - \tfrac{675}{32}\chi_{:a}\chi_{:a}L_{bb}L_{cc}\\
&-\tfrac{75}{4}\chi_{:a}\chi_{:b}L_{ac}L_{bc} - \tfrac{195}{16}\chi_{:a}\chi_{:a}L_{cd}L_{cd} - \tfrac{675}{8}\chi_{:a}\chi_{:b}L_{ab}L_{cc} - 330\chi_{:a}S_{:a}L_{cc}\\
&-300\chi_{:a}S_{:b}L_{ab} + \tfrac{15}{4}\chi_{:a}\chi_{:a}\chi_{:b}\chi_{:b} + \tfrac{15}{8}\chi_{:a}\chi_{:b}\chi_{:a}\chi_{:b} - \tfrac{15}{4}\chi_{:aa}\chi_{:bb} - \tfrac{105}{2}\chi_{:ab}\chi_{:ab}\\
&-15\chi_{:a}\chi_{:a}\chi_{:bb} - \tfrac{135}{2}\chi_{:b}\chi_{:aab}\} + F_{;m}\{(\tfrac{195}{2}\Pi_+ - 60\Pi_-)\tau_{;m} + 240\tau S - 90\rho_{mm}S\\
&+270S_{:aa} + (630\Pi_+ - 450\Pi_-)E_{;m} + 1440ES + 720S^3 + (90\Pi_+ + 450\Pi_-)\\
&\times L_{aa}E + (-\tfrac{165}{8}\Pi_+ - \tfrac{255}{8}\Pi_-)L_{aa}\rho_{mm} + (15\Pi_+ + 75\Pi_-)L_{aa}\tau + 600L_{aa}S^2\\
&+(\tfrac{1215}{8}\Pi_+ - \tfrac{315}{8}\Pi_-)L_{aa:bb} - \tfrac{45}{4}\chi L_{ab:ab} + (15\Pi_+ - 30\Pi_-)L_{ab}\rho_{ab} + (-\tfrac{165}{4}\Pi_+\\
&+\tfrac{465}{4}\Pi_-)L_{ab}R_{ammb} + \tfrac{705}{4}L_{aa}L_{bb}S - \tfrac{75}{2}L_{ab}L_{ab}S + (\tfrac{459}{32}\Pi_+ + \tfrac{495}{32}\Pi_-)\\
&\times L_{aa}L_{bb}L_{cc} + (\tfrac{267}{16}\Pi_+ - \tfrac{1485}{16}\Pi_-)L_{cc}L_{ab}L_{ab} + (-54\Pi_+ + \tfrac{225}{2}\Pi_-)L_{ab}L_{bc}L_{ac}\\
&-210\chi_{:a}S_{:a} - \tfrac{165}{16}\chi_{:a}\chi_{:a}L_{cc} - \tfrac{405}{8}\chi_{:a}\chi_{:b}L_{ab} + 135\chi\chi_{:a}\Omega_{am}\} + F_{;mm}\{30L_{aa}S\\
&+(\tfrac{315}{16}\Pi_+ - \tfrac{1215}{16}\Pi_-)L_{aa}L_{bb} + (-\tfrac{645}{8}\Pi_+ + \tfrac{945}{8}\Pi_-)L_{ab}L_{ab} + 60\chi\tau - 90\chi\rho_{mm}\\
&+360\chi E + 360S^2 - 30\chi_{:a}\chi_{:a}\} + F_{;mmm}\{180S + (-30\Pi_+ + 105\Pi_-)L_{aa}\}\\
&+45\chi F_{;mmmm}\}dy.
\end{aligned}
$$

We now consider transmission and transfer boundary conditions. Let M_+ and M_- be two manifolds which are coupled along a common boundary $\Sigma := \partial M_+ = \partial M_-$. We have metrics $g_\pm$ and operators $D_\pm$ of Laplace type on $M_\pm$. We have scalar smearing functions $f_\pm$ over $M_\pm$. Transmission boundary conditions arise in the study of heat conduction problems between closely coupled membranes. We impose the compatibility conditions

$$g_+|_\Sigma = g_-|_\Sigma, \quad V_+|_\Sigma = V_-|_\Sigma = V_\Sigma, \quad f_+|_\Sigma = f_-|_\Sigma\,.$$

No matching condition is assumed on the normal derivatives of f or of g on the interface Σ. Assume given an impedance matching endomorphism U defined on the hypersurface Σ. The *transmission boundary operator* is given by:

$$
\begin{aligned}
&\mathcal{B}_U\phi := \{\phi_+|_\Sigma - \phi_-|_\Sigma\} \quad \oplus \quad \{\nabla_{\nu_+}\phi_+|_\Sigma + \nabla_{\nu_-}\phi_-|_\Sigma - U\phi_+|_\Sigma\}\,, && (5.e)\\
&\omega_a := \nabla_a^+ - \nabla_a^-\,.
\end{aligned}
$$

Since the difference of two connections is tensorial, ω_a is a well defined endomorphism of V_Σ. The tensor ω_a is *chiral*; it changes sign if the roles of $+$ and $-$ are reversed. On the other hand, the tensor field U is *non-chiral* as it is not sensitive to the roles of $+$ and $-$.

The following result is due to Gilkey, Kirsten, and Vassilevich [62]; see also related work by Bordag and Vassilevich [31] and by Moss [95]. Define:

$$\begin{array}{ll}
\mathcal{L}_{ab}^{\text{even}} := L_{ab}^{+} + L_{ab}^{-}, & \mathcal{L}_{ab}^{\text{odd}} := L_{ab}^{+} - L_{ab}^{-}, \\
\mathcal{F}_{;\nu}^{\text{even}} := f_{;\nu^{+}} + f_{;\nu^{-}}, & \mathcal{F}_{;\nu}^{\text{odd}} := f_{;\nu^{+}} - f_{;\nu^{-}}, \\
\mathcal{F}_{;\nu\nu}^{\text{even}} := f_{;\nu^{+}\nu^{+}} + f_{;\nu^{-}\nu^{-}}, & \mathcal{F}_{;\nu\nu}^{\text{odd}} := f_{;\nu^{+}\nu^{+}} - f_{;\nu^{-}\nu^{-}}, \\
\mathcal{E}^{\text{even}} := E^{+} + E^{-}, & \mathcal{E}^{\text{odd}} := E^{+} - E^{-}, \\
\mathcal{E}_{;\nu}^{\text{even}} := E_{;\nu^{+}}^{+} + E_{;\nu^{-}}^{-}, & \mathcal{E}_{;\nu}^{\text{odd}} := E_{;\nu^{+}}^{+} - E_{;\nu^{-}}^{-}, \\
\mathcal{R}_{ijkl}^{\text{even}} := R_{ijkl}^{+} + R_{ijkl}^{-}, & \mathcal{R}_{ijkl}^{\text{odd}} := R_{ijkl}^{+} - R_{ijkl}^{-} \\
\Omega_{ij}^{\text{even}} := \Omega_{ij}^{+} + \Omega_{ij}^{-}, & \Omega_{ij}^{\text{odd}} := \Omega_{ij}^{+} - \Omega_{ij}^{-}\,.
\end{array}$$

Theorem 5.4 *[Transmission boundary conditions]*

(1) $a_0(f,D,\mathcal{B}_U) = (4\pi)^{-m/2}\int_M f\,\mathrm{Tr}(\mathrm{id})dx.$

(2) $a_1(f,D,\mathcal{B}_U) = 0.$

(3) $a_2(f,D,\mathcal{B}_U) = (4\pi)^{-m/2}\frac{1}{6}\int_M f\,\mathrm{Tr}\{\tau\,\mathrm{id} + 6E\}dx$
$\quad +(4\pi)^{-m/2}\frac{1}{6}\int_\Sigma 2f\,\mathrm{Tr}\{\mathcal{L}_{aa}^{\text{even}}\,\mathrm{id} - 6U\}dy.$

(4) $a_3(f,D,\mathcal{B}_U) = (4\pi)^{(1-m)/2}\frac{1}{384}\int_\Sigma \mathrm{Tr}\{f[\frac{3}{2}\mathcal{L}_{aa}^{\text{even}}\mathcal{L}_{bb}^{\text{even}} + 3\mathcal{L}_{ab}^{\text{even}}\mathcal{L}_{ab}^{\text{even}}]\,\mathrm{id}$
$\quad +9\mathcal{L}_{aa}^{\text{even}}\mathcal{F}_{;\nu}^{\text{even}}\,\mathrm{id} + 48fU^2 + 24f\omega_a\omega_a - 24f\mathcal{L}_{aa}^{\text{even}}U - 24\mathcal{F}_{;\nu}^{\text{even}}U\}dy.$

(5) $a_4(f,D,\mathcal{B}_U) = (4\pi)^{-m/2}\frac{1}{360}\int_M f\,\mathrm{Tr}\{60E_{;kk} + 60R_{ijji}E + 180E^2$
$\quad +30\Omega_{ij}\Omega_{ij} + [12\tau_{;kk} + 5\tau^2 - 2|\rho|^2 + 2|R|^2]\,\mathrm{id}\}dx$
$\quad +(4\pi)^{-m/2}\frac{1}{360}\int_\Sigma \mathrm{Tr}\{[-5\mathcal{R}_{ijji}^{\text{odd}}\mathcal{F}_{;\nu}^{\text{odd}} + 2\mathcal{R}_{a\nu a\nu}^{\text{odd}}\mathcal{F}_{;\nu}^{\text{odd}}$
$\quad -5\mathcal{L}_{aa}^{\text{odd}}\mathcal{L}_{bb}^{\text{even}}\mathcal{F}_{;\nu}^{\text{odd}} - \mathcal{L}_{ab}^{\text{odd}}\mathcal{L}_{ab}^{\text{even}}\mathcal{F}_{;\nu}^{\text{odd}} + \frac{18}{7}\mathcal{L}_{ab}^{\text{even}}\mathcal{L}_{ab}^{\text{even}}\mathcal{F}_{;\nu}^{\text{even}}$
$\quad -\frac{12}{7}\mathcal{L}_{aa}^{\text{even}}\mathcal{L}_{bb}^{\text{even}}\mathcal{F}_{;\nu}^{\text{even}} + 12\mathcal{L}_{aa}^{\text{even}}\mathcal{F}_{;\nu\nu}^{\text{even}}]\,\mathrm{id} + f[-\mathcal{L}_{ab}^{\text{odd}}\mathcal{L}_{ab}^{\text{odd}}\mathcal{L}_{cc}^{\text{even}}$
$\quad -\mathcal{L}_{ab}^{\text{even}}\mathcal{L}_{ab}^{\text{odd}}\mathcal{L}_{cc}^{\text{odd}} + 2\mathcal{L}_{ab}^{\text{odd}}\mathcal{L}_{bc}^{\text{odd}}\mathcal{L}_{ac}^{\text{even}} + 2\mathcal{R}_{abcb}^{\text{odd}}\mathcal{L}_{ac}^{\text{odd}} + 12\mathcal{R}_{ijji;\nu}^{\text{even}}$
$\quad +\frac{40}{21}\mathcal{L}_{aa}^{\text{even}}\mathcal{L}_{bb}^{\text{even}}\mathcal{L}_{cc}^{\text{even}} - \frac{4}{7}\mathcal{L}_{ab}^{\text{even}}\mathcal{L}_{ab}^{\text{even}}\mathcal{L}_{cc}^{\text{even}} + \frac{68}{21}\mathcal{L}_{ab}^{\text{even}}\mathcal{L}_{bc}^{\text{even}}\mathcal{L}_{ac}^{\text{even}}$
$\quad +24\mathcal{L}_{aa:bb}^{\text{even}} + 10\mathcal{R}_{ijji}^{\text{even}}\mathcal{L}_{aa}^{\text{even}} + 2\mathcal{R}_{a\nu a\nu}^{\text{even}}\mathcal{L}_{aa}^{\text{even}} - 6\mathcal{R}_{a\nu b\nu}^{\text{even}}\mathcal{L}_{ab}^{\text{even}}$
$\quad +2\mathcal{R}_{abcb}^{\text{even}}\mathcal{L}_{ac}^{\text{even}}]\,\mathrm{id} + 18\omega_a^2\mathcal{F}_{;\nu}^{\text{even}} - 30\mathcal{E}^{\text{odd}}\mathcal{F}_{;\nu}^{\text{odd}} + 15U\mathcal{L}_{aa}^{\text{odd}}\mathcal{F}_{;\nu}^{\text{odd}}$
$\quad -30U\mathcal{F}_{;\nu\nu}^{\text{even}} - 9U\mathcal{L}_{aa}^{\text{even}}\mathcal{F}_{;\nu}^{\text{even}} + 30U^2\mathcal{F}_{;\nu}^{\text{even}} + f[12\omega_a^2\mathcal{L}_{bb}^{\text{even}}$
$\quad +24\omega_a\omega_b\mathcal{L}_{ab}^{\text{even}} + 60\mathcal{E}_{;\nu}^{\text{even}} - 60\omega_a\Omega_{a\nu}^{odd} + 60\mathcal{E}^{\text{even}}\mathcal{L}_{aa}^{\text{even}} - 60U^3$
$\quad -30U\mathcal{R}_{ijji}^{\text{even}} - 180U\mathcal{E}^{\text{even}} - 60U_{:aa} - 18U\mathcal{L}_{aa}^{\text{even}}\mathcal{L}_{bb}^{\text{even}}$
$\quad -6U\mathcal{L}_{ab}^{\text{even}}\mathcal{L}_{ab}^{\text{even}} + 60U^2\mathcal{L}_{aa}^{\text{even}} - 60U\omega_a^2]\}dy.$

We now examine transfer boundary conditions. As previously, we take structures $(M,g,V,D) = ((M_+,g_+,V_+,D_+),(M_-,g_-,V_-,D_-))$. We now assume the compatibility conditions

$$\partial M_+ = \partial M_- = \Sigma \quad \text{and} \quad g_+|_\Sigma = g_-|_\Sigma\,.$$

We no longer assume an identification of $V_+|_\Sigma$ with $V_-|_\Sigma$. Let $F_\pm$ be smooth smearing endomorphisms of $V_\pm$; there is no assumed relation between F_+ and F_-. Let $\mathrm{Tr}_\pm$ denote the fiber trace on $V_\pm$. We suppose given auxiliary impedance matching endomorphisms $\mathfrak{S} := \{S_{\pm\pm}\}$ from $V_\pm$ to $V_\pm$. The *transfer boundary operator* is defined by setting:

$$\mathcal{B}_\mathfrak{S}\phi := \left\{ \begin{pmatrix} \nabla^+_{\nu_+} + S_{++} & S_{+-} \\ S_{-+} & \nabla^-_{\nu_-} + S_{--} \end{pmatrix} \begin{pmatrix} \phi_+ \\ \phi_- \end{pmatrix} \right\} \Bigg|_\Sigma . \tag{5.f}$$

We set $S_{+-} = S_{-+} = 0$ to introduce the associated decoupled Robin boundary conditions

$$\begin{aligned} &\mathcal{B}_{R(S_{++})}\phi_+ := (\nabla^+_{\nu_+} + S_{++})\phi_+|_\Sigma, \quad \text{and} \\ &\mathcal{B}_{R(S_{--})}\phi_- := (\nabla^-_{\nu_-} + S_{--})\phi_-|_\Sigma\,. \end{aligned}$$

Define the correction term $a_n(F, D, S)(y)$ by means of the identity

$$\begin{aligned} a_n(F, D, \mathcal{B}_\mathfrak{S}) &= \int_M a_n(F, D)(x)dx + \int_\Sigma a_n(F_+, D_+, \mathcal{B}_{R(S_{++})})dy \\ &+ \int_\Sigma a_n(F_-, D_-, \mathcal{B}_{R(S_{--})})dy + \int_\Sigma a_n(F, D, S)(y)dy\,. \end{aligned}$$

As the interior invariants $a_n(F, D)$ are discussed in Theorem 3.4 and as the Robin invariants $a_n(F, D, \mathcal{B}_{R(S_{++})})$ and $a_n(F, D, \mathcal{B}_{R(S_{--})})$ are discussed in Theorem 5.2, we must only determine the invariant $a_n(F, D, S)$ which measures the new interactions that arise from S_{+-} and S_{-+}. We refer to [63] for the proof of the following result:

Theorem 5.5 *[Transfer boundary conditions]*

(1) $a_n(F, D, \mathcal{B}_\mathfrak{S})(y) = 0$ *for* $n \le 2$.

(2) $a_3(F, D, \mathcal{B}_\mathfrak{S})(y) = (4\pi)^{(1-m)/2}\frac{1}{2}\big\{ \mathrm{Tr}_+(F_+S_{+-}S_{-+}) + \mathrm{Tr}_-(F_-S_{-+}S_{+-})\big\}$.

(3) $a_4(F, D, \mathcal{B}_\mathfrak{S})(y) = (4\pi)^{-m/2}\frac{1}{360}\big\{ \mathrm{Tr}_+\{480(F_+S_{++} + S_{++}F_+)S_{+-}S_{-+}$
$+480F_+S_{+-}S_{--}S_{-+}$
$+(288F_+L^+_{aa} + 192F_+L^-_{aa} + 240F_{+;\nu_+})S_{+-}S_{-+}\}$
$+\mathrm{Tr}_-\{480(F_-S_{--} + S_{--}F_-)S_{-+}S_{+-} + 480F_-S_{-+}S_{++}S_{+-}$
$+(288F_-L^-_{aa} + 192F_-L^+_{aa} + 240F_{-;\nu_-})S_{-+}S_{+-}\}\big\}$.

We now take up *spectral asymmetry*. We refer to [33, 34] for the material of this section. Let M be a compact Riemannian manifold. Let A be an operator of Dirac type and let $D = A^2$ be the associated operator of Laplace type. Instead of studying $\mathrm{Tr}_{L^2}(e^{-tD})$, we study $\mathrm{Tr}_{L^2}(Ae^{-tD})$; this provides a measure of the spectral asymmetry of A.

Let ∇ be a compatible connection; this means that $\nabla\gamma = 0$ and that if there is a fiber metric on V that ∇ is unitary. Expand $A = \gamma^\nu\nabla_{\partial_{x_\nu}} + \psi_A$. If ∂M is non-empty, we shall use local boundary conditions; we postpone until a subsequent section the question of spectral boundary conditions. Let $\{e_1, ..., e_m\}$ be a local orthonormal frame for the tangent bundle near ∂M which is normalized so e_m is the inward unit geodesic normal vector field. Suppose there exists an endomorphism χ of $V|_{\partial M}$ so that χ is self-adjoint and so that

$$\chi^2 = 1, \quad \chi\gamma_m + \gamma_m\chi = 0, \quad \text{and} \quad \chi\gamma_a = \gamma_a\chi \quad \text{for} \quad 1 \le a \le m-1\,.$$

Such a χ always exists if M is orientable and if m is even as, for example, one could take $\chi = \varepsilon\gamma_1...\gamma_{m-1}$ where ε is a suitable 4^{th} root of unity. There are topological obstructions to the existence of χ if m is odd; if ∂M is empty, χ plays no role. Let $\Pi_\chi^\pm := \frac{1}{2}(\text{id} \pm \chi)$ be orthonormal projection on the ± 1 eigenspaces of χ. We let $\mathcal{B}\phi := \Pi_\chi^- \phi|_{\partial M}$. The associated boundary condition for $D := A^2$ is defined by the operator $\mathcal{B}_1\phi := \mathcal{B}\phi \oplus \mathcal{B}A\phi$ and is equivalent to a mixed boundary operator $\mathcal{B}_{\chi,S}$ where

$$S = \tfrac{1}{2}\Pi_+(\gamma_m\psi_A - \psi_A\gamma_m - L_{aa}\chi)\Pi_+ .$$

As $t \downarrow 0$, there is an asymptotic expansion

$$\text{Tr}_{L^2}(FAe^{-tA_\mathcal{B}^2}) \sim \sum_{n=0}^\infty a_n^\eta(F,A,\mathcal{B})t^{(n-m-1)/2} .$$

These invariants measure the spectral asymmetry of A; $a_n^\eta(F,A,\mathcal{B}) = -a_n^\eta(F,-A,\mathcal{B})$.

Theorem 5.6 *Let* $W_{ij} := \Omega_{ij} - \frac{1}{4}R_{ijkl}\gamma_k\gamma_l$ *where* Ω *is the curvature of* ∇. *Let* $F = f\,\text{id}$ *for* $f \in C^\infty(M)$.

(1) $a_0^\eta(f,A,\mathcal{B}) = 0$.

(2) $a_1^\eta(f,A,\mathcal{B}) = -(4\pi)^{-m/2}(m-1)\int_M f\,\text{Tr}\{\psi_A\}dx$.

(3) $a_2^\eta(f,A,\mathcal{B}) = \frac{1}{4}(4\pi)^{-(m-1)/2}\int_{\partial M}(2-m)f\,\text{Tr}\{\psi_A\chi\}dy$.

(4) $a_3^\eta(f,A,\mathcal{B}) = -\frac{1}{12}(4\pi)^{-m/2}\int_M f\{\,\text{Tr}\{2(m-1)\nabla_{e_i}\psi_A + 3(4-m)\psi_A\gamma_i\psi_A$
$+3\gamma_j\psi_A\gamma_j\gamma_i\psi_A\}_{;i} + (m-3)\,\text{Tr}\{-\tau\psi_A - 6\gamma_i\gamma_jW_{ij}\psi_A + 6\gamma_i\psi_A\nabla_{e_i}\psi_A$
$+(m-4)\psi_A^3 - 3\psi_A^2\gamma_j\psi_A\gamma_j\}\}dx - \frac{1}{12}(4\pi)^{-m/2}\int_{\partial M}\text{Tr}\{6(m-2)f_{;m}\chi\psi_A$
$+f[(6m-18)\chi\nabla_{e_m}\psi_A + 2(m-1)\nabla_{e_m}\psi_A + 6\chi\gamma_m\gamma_a\nabla_{e_a}\psi_A$
$+6(2-m)\chi\psi_AL_{aa} + 2(3-m)\psi_AL_{aa} + 6(3-m)\chi\gamma_m\psi_A^2 + 3\gamma_m\psi_A\gamma_a\psi_A\gamma_a$
$+3(3-m)\chi\gamma_m\psi_A\chi\psi_A + 6\chi\gamma_aW_{am}]\}dy$.

6 Heat trace asymptotics and index theory

We refer to [54] for a more exhaustive treatment; the classical results may be found in [2, 3, 4, 7, 8]. In this section, we only present a brief introduction to the subject as it relates to heat trace asymptotics. Let $P : C^\infty(V_1) \to C^\infty(V_2)$ be a first order partial differential operator on a closed Riemannian manifold M. We assume V_1 and V_2 are equipped with fiber metrics. We say that the triple $\mathcal{C} := (P, V_1, V_2)$ is an *elliptic complex of Dirac type* if the associated second order operators $D_1 := P^*P$ and $D_2 := PP^*$ are of Laplace type. One may then define $\text{Index}(\mathcal{C}) := \dim\ker(D_1) - \dim\ker(D_2)$

Bott noted that $\text{Tr}_{L^2}\{e^{-tD_1}\} - Tr_{L^2}\{e^{-tD_2}\} = \text{Index}(\mathcal{C})$ was independent of the parameter t. He then used the asymptotic expansion of the heat equation to obtain a local formula for the index in terms of heat trace asymptotics. Following the notation of Equation (3.a), one may define the heat trace asymptotics of P by setting:

$$a_n^M(P)(x) := \{\,\text{Tr}\{e_n^M(D_1)\} - \text{Tr}\{e_n^M(D_2)\}\}(x) .$$

One then has a local formula for the index:

Theorem 6.1 *Let $\mathcal{C}$ be an elliptic complex of Dirac type over a closed Riemannian manifold M. Then:*

$$\int_M a_n^M(P)(x)dx = \begin{cases} \text{Index}(\mathcal{C}) & \textit{if} \quad n = m, \\ 0 & \textit{if} \quad n \neq m\,. \end{cases}$$

The critical term $a_m^M(P)(x)dx$ is often referred to as the *index density*. The other terms are in divergence form since they integrate to zero. They need not, however, vanish identically.

The existence of a local formula for the index implies the index is constant under deformations. It also yields, less trivially, that the index is multiplicative under finite coverings and additive with respect to connected sums. In the next section, we shall see that the index of the DeRham complex is the Euler-Poincare characteristic $\chi(M)$ of the manifold. Thus if $F \to M_1 \to M_2$ is a finite covering, then $\chi(M_1) = |F| \cdot \chi(M_2)$. Similarly, if $M = M_1 \# M_2$ is a connected sum, then $\chi(M) + \chi(S^m) = \chi(M_1) + \chi(M_2)$. Analogous formulas hold for the Hirzebruch signature of a manifold.

We define DeRham complex as follows. Let $d : C^\infty(\Lambda^p M) \to C^\infty(\Lambda^{p+1} M)$ be exterior differentiation and let $\delta : C^\infty(\Lambda^p M) \to C^\infty(\Lambda^{p-1} M)$ be the dual, interior multiplication. We may then define a 2-term elliptic complex of Dirac type:

$$\begin{aligned} &(d+\delta) : C^\infty(\Lambda^e M) \to C^\infty(\Lambda^o M) \quad \text{where} \\ &\Lambda^e(M) := \oplus_n \Lambda^{2n}(M) \quad \text{and} \quad \Lambda^o(M) := \oplus_n \Lambda^{2n+1}(M)\,. \end{aligned} \tag{6.a}$$

Let R_{ijkl} be the curvature tensor. Let $m = 2\bar{m}$ be even. Let $\{e_1, ..., e_m\}$ be a local orthonormal frame for the tangent bundle. We sum over repeated indices to define the *Pfaffian*

$$\mathcal{PF}_m : \;=\; \frac{g(e^{i_1} \wedge ... \wedge e^{i_m}, e^{j_1} \wedge ... \wedge e^{j_m})}{\pi^{\bar{m}} 8^{\bar{m}} \bar{m}!} R_{i_1 i_2 j_1 j_2} ... R_{i_{m-1} i_m j_{m-1} j_m}\,.$$

Set $\mathcal{PF}_m = 0$ if m is odd. The following result of Patodi [105] recovers the classical *Gauss-Bonnet theorem* of Chern [43]:

Theorem 6.2 *Let M be a closed even dimensional Riemannian manifold. Then*

(1) $a_n^M(d+\delta)(x) = 0$ *for* $n < m$.

(2) $a_m^M(d+\delta)(x) = \mathcal{PF}_m(x)$.

(3) $\chi(M) = \int_M \mathcal{PF}_m(x)dx$.

One can discuss Gauss-Bonnet theorem for manifolds with boundary similarly. On the boundary, normalize the orthonormal frame so e_m is the inward unit normal and let indices a, b range from 1 to $m-1$ and index the induced frame for the tangent bundle of the boundary. Let L_{ab} be the components of the second fundamental form. Define the *transgression* of the Pfaffian by setting:

$$\begin{aligned} T\mathcal{PF}_m : \;=\; &\sum_k \frac{g(e^{a_1} \wedge ... \wedge e^{a_{m-1}}, e^{b_1} \wedge ... \wedge e^{b_{m-1}})}{\pi^k 8^k k! (m-1-2k)! \operatorname{vol}(S^{m-1-2k})} \\ &\times \; R_{a_1 a_2 b_1 b_2} ... R_{a_{2k-1} a_{2k} b_{2k-1} b_{2k}} L_{a_{2k+1} b_{2k+1}} ... L_{a_{m-1} b_{m-1}}\,. \end{aligned}$$

If we impose absolute boundary conditions as discussed in Equation (5.c) to define the elliptic complex, we recover the Chern-Gauss-Bonnet theorem for manifolds with boundary [44]. Let Δ_M^e and Δ_M^o denote the Laplacians on the space of smooth differential forms of even and odd degrees, respectively. Let

$$a_n^{\partial M}(d+\delta)(y) = \left\{ \operatorname{Tr}\{e_n^{\partial M}(\Delta_M^e, \mathcal{B}_a)\} - \operatorname{Tr}\{e_n^{\partial M}(\Delta_M^0, \mathcal{B}_a)\}\right\}(y).$$

Theorem 6.1 extends to this setting to become:

$$\int_M a_n^M(d+\delta)(x)dx + \int_{\partial M} a_n^{\partial M}(d+\delta)(y)dy = \begin{cases} 0 & \text{if} \quad n \neq m, \\ \chi(M) & \text{if} \quad n = m\,. \end{cases}$$

Theorem 6.2 then extends to this setting to yield:

Theorem 6.3 (1) $a_n^{\partial \mathrm{M}}(d+\delta)(y) = 0$ *for* $n < m$.

(2) $a_m^{\partial \mathrm{M}}(d+\delta)(y) = T\mathcal{PF}_m$.

(3) $\chi(M) = \int_M \mathcal{PF}_m(x)dx + \int_{\partial M} T\mathcal{PF}_m(y)dy$.

The local index invariants $a_{m+2}^M(d+\delta)(x)$ are in divergence form but do not vanish identically. Set

$$\begin{aligned} \Phi_m &= \frac{\bar{m}}{\pi^{\bar{m}} 8^{\bar{m}} \bar{m}!}\{R_{i_1 i_2 j_1 k;k} R_{i_3 i_4 j_3 j_4} \ldots R_{i_{m-1} i_m j_{m-1} j_m}\}_{;j_2} \\ &\times \quad g(e^{i_1} \wedge \ldots \wedge e^{i_m}, e^{j_1} \wedge \ldots \wedge e^{j_m})\,. \end{aligned}$$

Theorem 6.4 *If M is even, then* $a_{m+2}^M(d+\delta) = \frac{1}{12}\mathcal{PF}_{m;kk} + \frac{1}{6}\Phi_m$.

Spectral boundary conditions plan an important role in index theory. We suppose given an elliptic complex of Dirac type $P : C^\infty(V_1) \to C^\infty(V_2)$. Let γ be the leading symbol of P. Then

$$\begin{pmatrix} 0 & \gamma^* \\ \gamma & 0 \end{pmatrix}$$

defines a *unitary Clifford module structure* on $V_1 \oplus V_2$. We may choose a unitary connection ∇ on $V_1 \oplus V_2$ which is compatible with respect to this Clifford module structure and which respects the splitting and induces connections ∇_1 and ∇_2 on the bundles V_1 and V_2, respectively. Decompose $P = \gamma_i \nabla_{e_i} + \psi$. Near the boundary, the structures depend on the normal variable. We set the normal variable x_m to zero to define a tangential operator of Dirac type

$$B(y) := \gamma_m(y,0)^{-1}\left(\gamma_a(y,0)\nabla_{e_a} + \psi(y,0)\right) \text{ on } C^\infty(V_1|_{\partial M})\,.$$

Let B^* be the adjoint of B on $L^2(V_1|_{\partial M})$;

$$B^* = \gamma_m(y,0)^{-1}\gamma_a(y,0)\nabla_{e_a} + \psi_B^*$$

where $\psi_B := \gamma_m(y,0)^{-1}\psi(y,0)$. Let Θ be an auxiliary self-adjoint endomorphism of V_1. We set

$$\begin{aligned} A &:= \tfrac{1}{2}(B + B^*) + \Theta \quad \text{on} \quad C^\infty(V_1|_{\partial M}), \\ A^\# &:= -\gamma^m A(\gamma^m)^{-1} \quad \text{on} \quad C^\infty(V_2|_{\partial M})\,. \end{aligned}$$

The leading symbol of A is then given by $\gamma_a^T := \gamma_m^{-1}\gamma_a$ which is a unitary Clifford module structure on $V_1|_{\partial M}$. Thus A is a self-adjoint operator of Dirac type on $C^\infty(V_1|_{\partial M})$; similarly $A^\#$ is a self-adjoint operator of Dirac type on $C^\infty(V_2|_{\partial M})$.

Let Π_A^+ (resp. $\Pi_{A^\#}^+$) be spectral projection on the eigenspaces of A (resp. $A^\#$) corresponding to the positive (resp. non-negative) eigenvalues; there is always a bit of technical fuss concerning the harmonic eigenspaces that we will ignore as it does not affect the heat trace asymptotic coefficients that we shall be considering. Introduce the associated spectral boundary operators by

$$\begin{aligned} \mathcal{B}_1\phi_1 &:= \Pi_A^+(\phi_1|_{\partial M}) && \text{for} \quad \phi_1 \in C^\infty(V_1), \\ \mathcal{B}_2\phi_2 &:= \Pi_{A^\#}^+(\phi_2|_{\partial M}) && \text{for} \quad \phi_2 \in C^\infty(V_2), \\ \mathcal{B}_\Theta\phi_1 &:= \mathcal{B}_1\phi_1 \oplus \mathcal{B}_2(P\phi_1) && \text{for} \quad \phi_1 \in C^\infty(V_1)\,. \end{aligned}$$

If $P_{\mathcal{B}_1}$, $P_{\mathcal{B}_2}^*$, and $D_{1,\mathcal{B}}$ are the realizations of P, of P^*, and of D_1 with respect to the boundary conditions $\mathcal{B}_1$,$\mathcal{B}_2$, and $\mathcal{B}_\Theta$, respectively, then

$$(P_{\mathcal{B}_1})^* = P_{\mathcal{B}_2}^* \qquad \text{and} \qquad D_{1,\mathcal{B}_\Theta} = P_{\mathcal{B}_1}^* P_{\mathcal{B}_1}\,.$$

We will discuss these boundary conditions in further detail in Section 10.

The local index density for the twisted signature and for the twisted spin complex has been identified using methods of invariance theory; see, for example, the discussion in Atiyah, Bott, and Patodi [5]. This identification of the local index density has been used to give a heat equation proof of the Atiyah-Singer index theorem in complete generality and has led to the proof of the index theorem for manifolds with boundary of Atiyah, Patodi, and Singer [6]. Unlike the DeRham complex, a salient feature of these complexes is the necessity to introduce spectral boundary conditions for the twisted signature and twisted spin complexes – there is a topological obstruction which prevents using local boundary conditions. The eta invariant plays an essential role in this development. We also refer to N. Berline, N. Getzler, and M. Vergne [28], to Bismut [30], and to Melrose [91] for other treatments of the local index theorem.

The Dolbeault complex is a bit different. Patodi [106] showed the heat trace invariants agreed with the classical Riemann-Roch invariant for a Kaehler manifold; it should be noted that this is not the case for an arbitrary Hermitian manifold. The Lefschetz fixed point formulas can also be established using heat equation methods.

7 Heat content asymptotics

We refer to [41, 56] for further details concerning the material of this section; we note that the asymptotic series for the heat content function is established by van den Berg et al [24] in a very general setting. Let D be an operator of Laplace type on a smooth vector bundle V over a smooth Riemannian manifold. Let $\langle\cdot,\cdot\rangle$ denote the natural pairing between V and the dual bundle $\tilde{V}$. Let $\rho \in C^\infty(\tilde{V})$ be the specific heat and let $\phi \in C^\infty(V)$ be the initial heat temperature distribution of the manifold. Impose suitable boundary conditions $\mathcal{B}$; we shall denote the dual boundary conditions for the dual operator $\tilde{D}$ on $C^\infty(\tilde{V})$ by $\tilde{\mathcal{B}}$. Let ∇ be the connection determined by D and E the associated endomorphism. Then the dual connection $\tilde{\nabla}$ and the dual endomorphism $\tilde{E}$ are the connection and the endomorphism determined by $\tilde{D}$.

The total heat energy content of the manifold is given by:

$$\beta(\phi,\rho,D,\mathcal{B})(t)=\beta(\rho,\phi,\tilde{D},\tilde{\mathcal{B}})(t):=\int_M\langle\rho,e^{-tD}\phi\rangle dx\,.$$

As $t\downarrow 0$, there is a complete asymptotic expansion of the form

$$\beta(\phi,\rho,D,\mathcal{B})(t)\sim\sum_{n=0}^{\infty}\beta_n(\phi,\rho,D,\mathcal{B})t^{n/2}\,.$$

There are local interior invariants β_n^M and boundary invariants $\beta_n^{\partial M}$ so that

$$\beta_n(\phi,\rho,D,\mathcal{B})=\int_M\beta_n^M(\phi,\rho,D)(x)dx+\int_{\partial M}\beta_n^{\partial M}(\phi,\rho,D,\mathcal{B})(y)dy\,.$$

These invariants are not uniquely characterized by this formula; divergence terms in the interior can be compensated by corresponding boundary terms.

We now study the heat content asymptotics of the disk D^m in $\mathbb{R}^m$ and the hemisphere H^m in S^m. We let D be the scalar Laplacian, $\phi=\rho=1$, and impose Dirichlet boundary conditions to define $\beta_n(M):=\beta_n(1,1,\Delta_M^0,\mathcal{B}_D)$. One has [16, 17] that:

Theorem 7.1 *Let D^m be the unit disk in $\mathbb{R}^m$. Then:*

(1) $\beta_0(D^m)=\frac{\pi^{m/2}}{\Gamma(\frac{(2+m)}{2})}$.

(2) $\beta_1(D^m)=-4\frac{\pi^{(m-1)/2}}{\Gamma(\frac{m}{2})}$.

(3) $\beta_2(D^m)=\frac{\pi^{m/2}}{\Gamma(\frac{m}{2})}(m-1)$.

(4) $\beta_3(D^m)=-\frac{\pi^{(m-1)/2}}{3\Gamma(\frac{m}{2})}(m-1)(m-3)$.

(5) $\beta_4(D^m)=-\frac{\pi^{m/2}}{8\Gamma(\frac{m}{2})}(m-1)(m-3)$.

(6) $\beta_5(D^m)=\frac{\pi^{(m-1)/2}}{120\Gamma(\frac{m}{2})}(m-1)(m-3)(m+3)(m-7)$.

(7) $\beta_6(D^m)=\frac{\pi^{m/2}}{96\Gamma(\frac{m}{2})}(m-1)(m-3)(m^2-4m-9)$.

(8) $\beta_7(D^m)=-\frac{\pi^{(m-1)/2}}{3360\Gamma(\frac{m}{2})}(m-1)(m-3)(m^4-8m^3-90m^2+424m+633)$.

Theorem 7.2 *Let H^m be the upper hemisphere of the unit sphere S^m. Then*

(1) $\beta_{2k}(H^m)=0$ *for any m if $k>0$.*

(2) $\beta_{2k+1}(H^3)=\frac{8\pi^{1/2}}{k!(2k-1)(2k+1)}$.

(3) $\beta_{2k+1}(H^5)=\frac{\pi^{3/2}2^{2k+3}(2-k)}{3k!(2k-1)(2k+1)}$.

(4) $\beta_{2k+1}(H^7)=\frac{\pi^{5/2}}{30}\left\{\frac{(67-54k)9^k}{k!(2k-1)(2k+1)}+\sum_{\ell=0}^k\frac{3\cdot 2^{3\ell}}{\ell!(k-\ell)!(2k-2\ell+1)}\right\}$.

We now study the heat content asymptotics with Dirichlet boundary conditions. Let $\mathcal{B}_D$ be the Dirichlet boundary operator of Equation (5.a). We refer to [16, 19] for the proof of:

Theorem 7.3 *[Dirichlet boundary conditions]*

(1) $\beta_0(\phi,\rho,D,\mathcal{B}_D)=\int_M\langle\phi,\rho\rangle dx$.

(2) $\beta_1(\phi,\rho,D,\mathcal{B}_D)=-\frac{2}{\sqrt{\pi}}\int_{\partial M}\langle\phi,\rho\rangle dy$.

(3) $\beta_2(\phi,\rho,D,\mathcal{B}_D)=-\int_M\langle D\phi,\rho\rangle dx+\int_{\partial M}\{\langle\frac12 L_{aa}\phi,\rho\rangle-\langle\phi,\rho_{;m}\rangle\}dy$.

(4) $\beta_3(\phi,\rho,D,\mathcal{B}_D)=-\frac{2}{\sqrt{\pi}}\int_{\partial M}\{\frac23\langle\phi_{;mm},\rho\rangle+\frac23\langle\phi,\rho_{;mm}\rangle-\langle\phi_{:a},\rho_{:a}\rangle+\langle E\phi,\rho\rangle$
$-\frac23 L_{aa}\langle\phi,\rho\rangle_{;m}+\langle(\frac{1}{12}L_{aa}L_{bb}-\frac16 L_{ab}L_{ab}-\frac16 R_{amma})\phi,\rho\rangle\}dy$.

(5) $\beta_4(\phi,\rho,D,\mathcal{B}_D)=\frac12\int_M\langle D\phi,\tilde D\rho\rangle dx+\int_{\partial M}\{\frac12\langle(D\phi)_{;m},\rho\rangle+\frac12\langle\phi,(\tilde D\rho)_{;m}\rangle$
$-\frac14\langle L_{aa}D\phi,\rho\rangle-\frac14\langle L_{aa}\phi,\tilde D\rho\rangle+\langle(\frac18 E_{;m}-\frac{1}{16}L_{ab}L_{ab}L_{cc}+\frac18 L_{ab}L_{ac}L_{bc}$
$-\frac{1}{16}R_{ambm}L_{ab}+\frac{1}{16}R_{abcb}L_{ac}+\frac{1}{32}\tau_{;m}+\frac{1}{16}L_{ab:ab})\phi,\rho\rangle$
$-\frac14 L_{ab}\langle\phi_{:a},\rho_{:b}\rangle-\frac18\langle\Omega_{am}\phi_{:a},\rho\rangle+\frac18\langle\Omega_{am}\phi,\rho_{:a}\rangle\}dy$.

We may compute $\beta_n(M)$ for $n\le 4$ by setting $\phi=\rho=1$ and $E=\Omega=0$ in Theorem 7.3. One has a formula [18] for $\beta_5(M)$; $\beta_5(\phi,\rho,D,\mathcal{B}_D)$ is not known in full generality.

Theorem 7.4 $\beta_5(M)=-\frac{1}{240\sqrt{\pi}}\int_{\partial M}\{8\rho_{mm;mm}-8L_{aa}\rho_{mm;m}+16L_{ab}R_{ammb;m}$
$-4\rho_{mm}^2+16R_{ammb}R_{ammb}-4L_{aa}L_{bb}\rho_{mm}-8L_{ab}L_{ab}\rho_{mm}+64L_{ab}L_{ac}R_{mbcm}$
$-16L_{aa}L_{bc}R_{mbcm}-8L_{ab}L_{ac}R_{bddc}-8L_{ab}L_{cd}R_{acbd}+4R_{abcm}R_{abcm}$
$+8R_{abbm}R_{accm}-16L_{aa:b}R_{bccm}-8L_{ab:c}L_{ab:c}+L_{aa}L_{bb}L_{cc}L_{dd}$
$-4L_{aa}L_{bb}L_{cd}L_{cd}+4L_{ab}L_{ab}L_{cd}L_{cd}-24L_{aa}L_{bc}L_{cd}L_{db}+48L_{ab}L_{bc}L_{cd}L_{da}\}dy$.

The invariants $\beta_0(M)$, $\beta_1(M)$, and $\beta_2(M)$ were computed by van den Berg and Davies [20] and by van den Berg and Le Gall [21] for domains in $\mathbb{R}^m$. The invariants $\beta_0(M)$, $\beta_1(M)$, and $\beta_2(M)$ were computed by van den Berg [14] for the upper hemisphere of the unit sphere. The general case where D is an arbitrary operator of Laplace type and where ϕ and ρ are arbitrary was studied in [16, 19]. Savo [112, 113, 114, 115] has given a closed formula for all the heat content asymptotics $\beta_k(M)$ that is combinatorially quite different in nature from the formulas we have presented here. There is also important related work of McAvity [87, 88], of McDonald and Meyers [89], and of Phillips and Jansons [108].

We now study heat content asymptotics for Robin boundary conditions. Let $\mathcal{B}_S$ be the Robin boundary operator of Equation (5.b); the dual boundary condition is then given by $\tilde{\mathcal{B}}_S\rho=\mathcal{B}_{\tilde S}\rho=(\rho_{;m}+\tilde S\rho)|_{\partial M}$ where, of course, we use the dual connection on $\tilde V$ to define $\rho_{;m}$. The following result is proved in [19, 45]:

Theorem 7.5 *[Robin boundary conditions]*

(1) $\beta_0(\phi,\rho,D,\mathcal{B}_S)=\int_M\langle\phi,\rho\rangle dx$.

(2) $\beta_1(\phi,\rho,D,\mathcal{B}_S)=0$.

(3) $\beta_2(\phi,\rho,D,\mathcal{B}_S)=-\int_M\langle D\phi,\rho\rangle dx+\int_{\partial M}\langle\mathcal{B}_S\phi,\rho\rangle dy.$

(4) $\beta_3(\phi,\rho,D,\mathcal{B}_S)=\frac{2}{3}\cdot\frac{2}{\sqrt{\pi}}\int_{\partial M}\langle\mathcal{B}_S\phi,\mathcal{B}_{\tilde S}\rho\rangle dy.$

(5) $\beta_4(\phi,\rho,D,\mathcal{B}_S)=\frac{1}{2}\int_M\langle D\phi,\tilde D\rho\rangle dx+\int_{\partial M}\{-\frac{1}{2}\langle\mathcal{B}_S\phi,\tilde D\rho\rangle-\frac{1}{2}\langle D\phi,\mathcal{B}_{\tilde S}\rho\rangle$
$+\langle(\frac{1}{2}S+\frac{1}{4}L_{aa})\mathcal{B}_S\phi,\mathcal{B}_{\tilde S}\rho\rangle\}dy.$

(6) $\beta_5(\phi,\rho,D,\mathcal{B}_S)=\frac{2}{\sqrt{\pi}}\int_{\partial M}\{-\frac{4}{15}(\langle\mathcal{B}_SD\phi,\mathcal{B}_{\tilde S}\rho\rangle+\langle\mathcal{B}_S\phi,\mathcal{B}_{\tilde S}\tilde D\rho\rangle)$
$-\frac{2}{15}\langle(\mathcal{B}_S\phi)_{:a},(\mathcal{B}_{\tilde S}\rho)_{:a}\rangle+\langle(\frac{2}{15}E+\frac{4}{15}S^2+\frac{4}{15}SL_{aa}+\frac{1}{30}L_{aa}L_{bb}$
$+\frac{1}{15}L_{ab}L_{ab}-\frac{1}{15}R_{amam})\mathcal{B}_S\phi,\mathcal{B}_{\tilde S}\rho\rangle\}dy.$

(7) $\beta_6(\phi,\rho,D,\mathcal{B}_S)=-\frac{1}{6}\int_M\langle D^2\phi,\tilde D\rho\rangle dx+\int_{\partial M}\{\frac{1}{6}\langle\mathcal{B}_SD\phi,\tilde D\rho\rangle+\frac{1}{6}\langle D^2\phi,\tilde{\mathcal{B}}_S\rho\rangle$
$+\frac{1}{6}\langle\mathcal{B}_S\phi,\tilde D^2\rho\rangle-\frac{1}{6}\langle S\mathcal{B}_SD\phi,\tilde{\mathcal{B}}_S\rho\rangle-\frac{1}{6}\langle S\mathcal{B}_S\phi,\tilde{\mathcal{B}}_S\tilde D\rho\rangle-\frac{1}{12}\langle L_{aa}\mathcal{B}_SD\phi,\tilde{\mathcal{B}}_S\rho\rangle$
$-\frac{1}{12}\langle L_{aa}\mathcal{B}_S\phi,\tilde{\mathcal{B}}_S\tilde D\rho\rangle+\langle(\frac{1}{24}E_{;m}+\frac{1}{12}EL_{aa}+\frac{1}{48}L_{ab}L_{ab}L_{cc}+\frac{1}{24}L_{ab}L_{ac}L_{bc}$
$-\frac{1}{48}R_{ambm}L_{ab}+\frac{1}{48}R_{abcb}L_{ac}-\frac{1}{24}R_{amam}L_{bb}+\frac{1}{96}\tau_{;m}+\frac{1}{48}L_{ab:ab}$
$+\frac{1}{12}SL_{aa}L_{bb}+\frac{1}{12}SL_{ab}L_{ab}-\frac{1}{12}SR_{amam}+\frac{1}{12}(SE+ES)+\frac{1}{4}S^2L_{aa}+\frac{1}{6}S^3$
$+\frac{1}{6}S_{:aa})\mathcal{B}_S\phi,\tilde{\mathcal{B}}_S\rho\rangle-\frac{1}{12}L_{aa}\langle(\mathcal{B}_S\phi)_{:b},(\tilde{\mathcal{B}}_S\rho)_{:b}\rangle-\frac{1}{12}L_{ab}\langle(\mathcal{B}_S\phi)_{:a},(\tilde{\mathcal{B}}_S\rho)_{:b}\rangle$
$-\frac{1}{6}\langle S(\mathcal{B}_S\phi)_{:a},(\tilde{\mathcal{B}}_S\rho)_{:a}\rangle-\frac{1}{24}\langle\Omega_{am}(\mathcal{B}_S\phi)_{:a},\tilde{\mathcal{B}}_S\rho\rangle$
$+\frac{1}{24}\langle\Omega_{am}\mathcal{B}_S\phi,(\tilde{\mathcal{B}}_S\rho)_{:a}\rangle\}dy.$

We now turn our attention to mixed boundary conditions. We use Equation (5.d) to defined the mixed boundary operator $\mathcal{B}_{\chi,S}$. The dual boundary operator on $\tilde V$ is given by $\tilde{\mathcal{B}}_{\chi,S}\rho:=\tilde\Pi_+(\rho_{;m}+\tilde S\rho)|_{\partial M}\oplus\tilde\Pi_-\rho|_{\partial M}$. Extend χ to a collared neighborhood of M to be parallel along the inward normal geodesic rays. Then $\chi_{;m}=0$. Let $\phi_\pm:=\Pi_\pm\phi$ and $\rho_\pm:=\Pi_\pm\rho$. Since $\chi_{;m}=0$, $\phi_{\pm;m}=\Pi_\pm(\phi_{;m})$ and $\rho_{\pm;m}=\tilde\Pi_\pm(\phi_{;m})$. As $\chi_{:a}$ need not vanish in general, we need not have equality between $\phi_{\pm:a}$ and $\Pi_\pm(\phi_{:a})$ or between $\rho_{\pm:a}$ and $\tilde\Pi_\pm(\rho_{:a})$. We refer to [45] for the proof of:

Theorem 7.6 *[Mixed boundary conditions]*

(1) $\beta_0(\phi,\rho,D,\mathcal{B}_{\chi,S})=\int_M\langle\phi,\rho\rangle dx.$

(2) $\beta_1(\phi,\rho,D,\mathcal{B}_{\chi,S})=-\frac{2}{\sqrt{\pi}}\int_{\partial M}\langle\phi_-,\rho_-\rangle dy.$

(3) $\beta_2(\phi,\rho,D,\mathcal{B}_{\chi,S})=-\int_M\langle D\phi,\rho\rangle dx+\int_{\partial M}\{\langle\phi_{+;m}+S\phi_+,\rho_+\rangle$
$+\langle\frac{1}{2}L_{aa}\phi_-,\rho_-\rangle-\langle\phi_-,\rho_{-;m}\rangle\}dy.$

(4) $\beta_3(\phi,\rho,D,\mathcal{B}_{\chi,S})=\frac{2}{\sqrt{\pi}}\int_{\partial M}\{-\frac{2}{3}\langle\phi_{-;mm},\rho_-\rangle-\frac{2}{3}\langle\phi_-,\rho_{-;mm}\rangle+\frac{2}{3}L_{aa}\langle\phi_-,\rho_-\rangle_{;m}$
$+\langle(-\frac{1}{12}L_{aa}L_{bb}+\frac{1}{6}L_{ab}L_{ab}+\frac{1}{6}R_{amma})\phi_-,\rho_-\rangle+\frac{2}{3}\langle\phi_{+;m}+S\phi_+,\rho_{+;m}+\tilde S\rho_+\rangle$
$-\langle E\phi_-,\rho_-\rangle+\langle\phi_{-:a},\rho_{-:a}\rangle+\frac{2}{3}\langle\phi_{+:a},\rho_{-:a}\rangle+\frac{2}{3}\langle\phi_{-:a},\rho_{+:a}\rangle$
$-\frac{2}{3}\langle E\phi_-,\rho_+\rangle-\frac{2}{3}\langle E\phi_+,\rho_-\rangle\}dy.$

We adopt the notation of Equation (5.e) to define the transmission boundary operator $\mathcal{B}_U$ and the tensor ω.

Theorem 7.7 *[Transmission boundary conditions]*

(1) $\beta_0(\phi,\rho,D,\mathcal{B}_U)=\int_{M_+}\langle\phi_+,\rho_+\rangle dx_+ + \int_{M_-}\langle\phi_-,\rho_-\rangle dx_-$.

(2) $\beta_1(\phi,\rho,D,\mathcal{B}_U)=-\frac{1}{\sqrt{\pi}}\int_\Sigma\langle\phi_+-\phi_-,\rho_+-\rho_-\rangle dy$.

(3) $\beta_2(\phi,\rho,D,\mathcal{B}_U)=-\int_{M_+}\langle D_+\phi_+,\rho_+\rangle dx_+ - \int_{M_-}\langle D_-\phi_-,\rho_-\rangle dx_-$
$+\int_\Sigma\big\{\frac18(L_{aa}^++L_{aa}^-)(\langle\phi_+,\rho_+\rangle+\langle\phi_-,\rho_-\rangle)$
$-\frac18(L_{aa}^++L_{aa}^-)(\langle\phi_+,\rho_-\rangle+\langle\phi_-,\rho_+\rangle)+\frac12(\langle\phi_{+;\nu},\rho_+\rangle+\langle\phi_{-;\nu}\rho_-\rangle+\langle\phi_{+;\nu},\rho_-\rangle$
$+\langle\phi_{-;\nu},\rho_+\rangle)-\frac12(\langle\phi_+,\rho_{+;\nu}\rangle+\langle\phi_-,\rho_{-;\nu}\rangle)+\frac12(\langle\phi_+,\rho_{-;\nu}\rangle+\langle\phi_-,\rho_{+;\nu}\rangle)$
$-\frac14(\langle U\phi_+,\rho_+\rangle+\langle U\phi_-,\rho_-\rangle+\langle U\phi_+,\rho_-\rangle+\langle U\phi_-,\rho_+\rangle)\big\}dy$.

(4) $\beta_3(\phi,\rho,D,\mathcal{B}_U)=\frac{1}{6\sqrt{\pi}}\int_\Sigma\{4(\langle D_+\phi_+,\rho_+\rangle+\langle\phi_+,\tilde D_+\rho_+\rangle+\langle D_-\phi_-,\rho_-\rangle$
$+\langle\phi_-,\tilde D_-\rho_-\rangle)-4(\langle D_+\phi_+,\rho_-\rangle+\langle\phi_+,\tilde D_-\rho_-\rangle+\langle D_-\phi_-,\rho_+\rangle+\langle\phi_-,\tilde D_+\rho_+\rangle)$
$-(\langle\omega_a\phi_{+;a},\rho_+\rangle-\langle\omega_a\phi_{-;a},\rho_-\rangle-\langle\omega_a\phi_+,\rho_{+;a}\rangle+\langle\omega_a\phi_-,\rho_{-;a}\rangle)$
$-(\langle\omega_a\phi_{+;a},\rho_-\rangle-\langle\omega_a\phi_{-;a},\rho_+\rangle+\langle\omega_a\phi_+,\rho_{-;a}\rangle-\langle\omega_a\phi_-,\rho_{+;a}\rangle)$
$+4(\langle\phi_{+;\nu},\rho_{+;\nu}\rangle+\langle\phi_{-;\nu},\rho_{-;\nu}\rangle+\langle\phi_{+;\nu},\rho_{-;\nu}\rangle+\langle\phi_{-;\nu},\rho_{+;\nu}\rangle)-2(\langle\phi_{+;a},\rho_{+;a}\rangle$
$+\langle\phi_{-;a},\rho_{-;a}\rangle)+2(\langle\phi_{+;a},\rho_{-;a}\rangle+\langle\phi_{-;a},\rho_{+;a}\rangle)-2(\langle U\phi_{+;\nu},\rho_+\rangle$
$+\langle U\phi_+,\rho_{+;\nu}\rangle+\langle U\phi_{-;\nu},\rho_-\rangle+\langle U\phi_-,\rho_{-;\nu}\rangle)-2(\langle U\phi_{-;\nu},\rho_+\rangle+\langle U\phi_-,\rho_{+;\nu}\rangle$
$+\langle U\phi_{+;\nu},\rho_-\rangle+\langle U\phi_+,\rho_{-;\nu}\rangle)+(L_{aa}^--L_{aa}^+)(\nu_+\langle\phi_+,\rho_+\rangle-\nu_-\langle\phi_-,\rho_-\rangle)$
$+L_{aa}^+(\langle\phi_{+;\nu},\rho_-\rangle+\langle\phi_-,\rho_{+;\nu}\rangle)+L_{aa}^-(\langle\phi_{-;\nu},\rho_+\rangle+\langle\phi_+,\rho_{-;\nu}\rangle)$
$-(L_{aa}^-(\langle\phi_{+;\nu},\rho_-\rangle+\langle\phi_-,\rho_{+;\nu}\rangle)+L_{aa}^+(\langle\phi_{-;\nu},\rho_+\rangle+\langle\phi_+,\rho_{-;\nu}\rangle))$
$+\langle\omega_a\omega_a\phi_+,\rho_+\rangle+\langle\omega_a\omega_a\phi_-,\rho_-\rangle-\frac12L_{aa}^+L_{bb}^-(\langle\phi_+,\rho_+\rangle+\langle\phi_-,\rho_-\rangle)$
$+\frac12L_{aa}^+L_{bb}^-(\langle\phi_+,\rho_-\rangle+\langle\phi_-,\rho_+\rangle)+\frac12(L_{ab}^+L_{ab}^+\langle\phi_+,\rho_+\rangle+L_{ab}^-L_{ab}^-\langle\phi_-,\rho_-\rangle)$
$+\frac12(L_{ab}^-L_{ab}^-\langle\phi_+,\rho_+\rangle+L_{ab}^+L_{ab}^+\langle\phi_-,\rho_-\rangle)-\frac12(L_{ab}^+L_{ab}^++L_{ab}^-L_{ab}^-)(\langle\phi_+,\rho_-\rangle$
$+\langle\phi_-,\rho_+\rangle)+L_{aa}^+\langle U\phi_+,\rho_+\rangle+L_{aa}^-\langle U\phi_-,\rho_-\rangle-L_{aa}^-\langle U\phi_+,\rho_+\rangle$
$-L_{aa}^+\langle U\phi_-,\rho_-\rangle+\langle U^2\phi_+,\rho_+\rangle+\langle U^2\phi_-,\rho_-\rangle+\langle U^2\phi_+,\rho_-\rangle+\langle U^2\phi_-,\rho_+\rangle$
$+\langle E_+\phi_+,\rho_+\rangle+\langle E_-\phi_-,\rho_-\rangle+\langle E_-\phi_+,\rho_+\rangle+\langle E_+\phi_-,\rho_-\rangle$
$-\langle(E_++E_-)\phi_+,\rho_-\rangle-\langle(E_++E_-)\phi_-,\rho_+\rangle+\frac12(R_{amma}^++R_{amma}^-)(\langle\phi_+,\rho_+\rangle$
$+\langle\phi_-,\rho_-\rangle)-\frac12(R_{amma}^++R_{amma}^-)(\langle\phi_+,\rho_-\rangle+\langle\phi_-,\rho_+\rangle)\}dy$.

We continue our studies by examining the heat content asymptotics for transfer boundary conditions Adopt the Equation (5.f) to define the transfer boundary operator $\mathcal{B}_{\mathfrak{S}}$. Let $\tilde{\mathcal{B}_{\mathfrak{S}}}$ be the dual boundary operator

$$\tilde{\mathcal{B}_{\mathfrak{S}}}\rho:=\left\{\left(\begin{array}{cc}\tilde\nabla_{\nu_+}^++\tilde S_{++} & \tilde S_{-+}\\ \tilde S_{+-} & \tilde\nabla_{\nu_-}^-+\tilde S_{--}\end{array}\right)\left(\begin{array}{c}\rho_+\\ \rho_-\end{array}\right)\right\}\Big|_\Sigma.$$

We refer to [57] for the proof of the following result:

Theorem 7.8 *[Transfer boundary conditions]*

(1) $\beta_0(\phi,\rho,D,\mathcal{B}_{\mathfrak{S}}) = \int_{M_+}\langle\phi_+,\rho_+\rangle dx_+ + \int_{M_-}\langle\phi_-,\rho_-\rangle dx_-$.

(2) $\beta_1(\phi,\rho,D,\mathcal{B}_{\mathfrak{S}}) = 0$.

(3) $\beta_2(\phi,\rho,D,\mathcal{B}_{\mathfrak{S}}) = -\int_{M_+}\langle D_+\phi_+,\rho_+\rangle dx_+ - \int_{M_-}\langle D_-\phi_-,\rho_-\rangle dx_-$
$+\int_\Sigma\langle\mathcal{B}_{\mathfrak{S}}\phi,\rho\rangle dy$.

(4) $\beta_3(\phi,\rho,D,\mathcal{B}_{\mathfrak{S}}) = \frac{4}{3\sqrt{\pi}}\int_\Sigma\langle\mathcal{B}_{\mathfrak{S}}\phi,\tilde{\mathcal{B}_{\mathfrak{S}}}\rho\rangle)dy$.

Oblique boundary conditions are of particular interest. Let D be an operator of Laplace type on a bundle V over M. Let $\mathcal{B}_T$ be a tangential first order partial differential operator on $V|_{\partial M}$ and let $\tilde{B}_T$ be the dual operator on $\tilde{V}|_{\partial M}$. The associated *oblique boundary conditions* on V and dual boundary conditions on $\tilde{V}$ are defined by:

$$\mathcal{B}_{\mathcal{O}}\phi := (\phi_{;m} + \mathcal{B}_T\phi)|_{\partial M} \quad \text{and} \quad \tilde{\mathcal{B}}_{\mathcal{O}}\rho := (\rho_{;m} + \tilde{\mathcal{B}}_T\rho)|_{\partial M}\,.$$

Note that we recover Robin boundary conditions by taking $\mathcal{B}_T$ to be a 0^{th} order operator. We refer to [59] for the proof of the following result:

Theorem 7.9 *[Oblique boundary conditions]*

(1) $\beta_0(\phi,\rho,D,\mathcal{B}_{\mathcal{O}}) = \int_M\langle\phi,\rho\rangle dx$.

(2) $\beta_1(\phi,\rho,D,\mathcal{B}_{\mathcal{O}}) = 0$.

(3) $\beta_2(\phi,\rho,D,\mathcal{B}_{\mathcal{O}}) = -\int_M\langle D\phi,\rho\rangle dx + \int_{\partial M}\langle\mathcal{B}_{\mathcal{O}}\phi,\rho\rangle dy$.

(4) $\beta_3(\phi,\rho,D,\mathcal{B}_{\mathcal{O}}) = \frac{4}{3\sqrt{\pi}}\int_{\partial M}\langle\mathcal{B}_{\mathcal{O}}\phi,\tilde{\mathcal{B}}_{\mathcal{O}}\rho\rangle dy$.

(5) $\beta_4(\phi,\rho,D,\mathcal{B}_{\mathcal{O}}) = \frac12\int_M\langle D\phi,\tilde{D}\rho\rangle dx + \int_{\partial M}\{-\frac12\langle\mathcal{B}_{\mathcal{O}}\phi,\tilde{D}\rho\rangle$
$-\frac12\langle D\phi,\tilde{\mathcal{B}}\rho\rangle + \langle(\frac12\mathcal{B}_T + \frac14 L_{aa})\mathcal{B}_{\mathcal{O}}\phi,\tilde{\mathcal{B}}_{\mathcal{O}}\rho\rangle\}dy$.

We refer to [24] for further details concerning *Zaremba boundary conditions*. We assume given a decomposition $\partial M = C_R \cup C_D$ as the union of two closed submanifolds with common smooth boundary $C_R \cap C_D = \Sigma$. Let $\phi_{;m}$ denote the covariant derivative of ϕ with respect to the inward unit normal on ∂M. Let S be an auxiliary endomorphism of $V|_{C_R}$. We take Robin boundary conditions on C_R and Dirichlet boundary conditions on C_D arising from the boundary operator:

$$\mathcal{B}_Z\phi := (\phi_{;m} + S\phi)|_{\{C_R-\Sigma\}} \oplus \phi|_{C_D}\,.$$

We refer to related work of Avramidi [11], of Dowker [46, 47], and of Jakobson et al. [80] concerning the heat trace asymptotics.

There is some additional technical fuss concerned with choosing a boundary condition on the interface $C_D \cap C_R$ that we will suppress in the interests of brevity. Instead, we shall simply give a classical formulation of the problem. Suppose $D = \Delta$ is the Laplacian and that $S = 0$. Let $W^{1,2}(M)$ be the closure of $C^\infty(M)$ with respect to the Sobolev norm

$$||\phi||_1^2 = \int_M \{|\nabla\phi|^2 + |\phi|^2\}dx.$$

Let $W^{1,2}_{0,C_D}(M)$ be the closure of the set $\{\phi \in W^{1,2}(M) : \text{supp}(\phi) \cap C_D = \emptyset\}$. Let

$$N(M, C_D, \lambda) = \sup(\dim E_\lambda) \quad \text{for} \quad \lambda > 0$$

where the supremum is taken over all subspaces $E_\lambda \subset W^{1,2}_{0,C_D}(M)$ such that

$$||\nabla\phi||_{L^2(M)} < \lambda||\phi||_{L^2(M)}, \quad \forall \phi \in E_\lambda .$$

This is the spectral counting function for the Zaremba problem described above.

On Σ, we choose an orthonormal frame so e_m is the inward unit normal of ∂M in M and so that e_{m-1} is the inward unit normal of Σ in C_D.

Theorem 7.10 *[Zaremba boundary conditions] There exist universal constants c_1 and c_2 so that:*

(1) $\beta_0(\phi, \rho, D, \mathcal{B}) = \int_M \langle\phi, \rho\rangle dx.$

(2) $\beta_1(\phi, \rho, D, \mathcal{B}) = -\frac{2}{\sqrt{\pi}} \int_{C_D} \langle\phi, \rho\rangle dy.$

(3) $\beta_2(\phi, \rho, D, \mathcal{B}) = -\int_M \langle D\phi, \rho\rangle dx + \int_{C_R} \{\langle\phi_{;m} + S\phi, \rho\rangle\} dy$
$+ \int_{C_D} \{\frac{1}{2} L_{aa}\langle\phi, \rho\rangle - \langle\phi, \rho_{;m}\rangle\} dy - \frac{1}{2} \int_\Sigma \langle\phi, \rho\rangle dz.$

(4) $\beta_3(\phi, \rho, D, \mathcal{B}) = \frac{4}{3\sqrt{\pi}} \int_{C_R} \langle\phi_{;m} + S\phi, \rho_{;m} + \tilde{S}\rho\rangle dy - \frac{2}{\sqrt{\pi}} \int_{C_D} \{\frac{2}{3}\langle\phi_{;mm}, \rho\rangle$
$+\frac{2}{3}\langle\phi, \rho_{;mm}\rangle - \langle\phi_{:a}, \rho_{:a}\rangle + \langle E\phi, \rho\rangle - \frac{2}{3} L_{aa}\langle\phi, \rho\rangle_{;m} + \langle(\frac{1}{12} L_{aa}L_{bb}$
$-\frac{1}{6} L_{ab}L_{ab} + \frac{1}{6} R_{amam})\phi, \rho\rangle\} dy + \int_\Sigma \{\langle(c_1 L_{m-1,m-1} + (\frac{1}{2} c_2 + \frac{2}{3\sqrt{\pi}}) L_{uu}$
$+\frac{1}{2\sqrt{\pi}} \tilde{L}_{uu} + c_2 S)\phi, \rho\rangle + \frac{1}{2\sqrt{\pi}} \langle\phi, \rho\rangle_{;m-1} - \frac{2}{3\sqrt{\pi}} \langle\phi, \rho\rangle_{;m}\} dz.$

We conclude this section with a brief description of the non-smooth setting. We refer to van den Berg and Srisatkunarajah [25] for a discussion of the heat content asymptotics of polygonal regions in the plane. The fractal setting also an important one and we refer to van den Berg [15], to Fleckinger et al. [51], to Griffith and Lapidus [70], to Lapidus and Pang [85], and to Neuberger et al. [100] for a discussion of some asymptotic results for heat problems on the von Koch snowflake.

8 Heat content with source terms

We follow the discussion in [18, 22, 23, 56] throughout this section. Let D be an operator of Laplace type. Assume $\partial M = C_D \cup C_R$ decomposes as a disjoint union of two closed, possibly empty, disjoint subsets; in contrast to the Zaremba problem, we emphasize that $C_D \cap C_R$ is empty. Let $\mathcal{B}$ be the Dirichlet boundary operator on C_D and the Robin boundary operator on C_R. Let ϕ be the initial temperature of the manifold, let $\rho = \rho(x; t)$ be a variable specific heat, let $p = p(x; t)$ be an auxiliary smooth internal heat source and let $\psi = \psi(y; t)$ be the temperature of the boundary. We assume, for the sake of simplicity, that the underlying geometry is fixed. Let $u(x; t) = u_{\phi,p,\psi}(x; t)$ be the subsequent temperature distribution which is defined by the relations:

$$(\partial_t + D)u(x; t) = p(x; t) \quad \text{for} \quad t > 0,$$

$$\mathcal{B}u(y;t) = \psi(y;t) \quad \text{for} \quad t > 0, y \in \partial M,$$
$$\lim_{t \downarrow 0} u(\cdot;t) = \phi(\cdot) \text{ in } L^2 .$$

The associated heat content function has a complete asymptotic series as $t \downarrow 0$:

$$\begin{aligned} \beta(\phi, \rho, D, \mathcal{B}, p, \psi)(t) : &= \int_M \langle u_{\phi,p,\psi}(x;t), \rho(x;t)\rangle dx \\ &\sim \sum_{n=0}^{\infty} \beta_n(\phi, \rho, D, \mathcal{B}, p, \psi) t^{n/2} . \end{aligned}$$

Assertions (1)-(4) in the following result are valid for quite general boundary conditions. Assertion (5) refers to the particular problem under consideration. This result when combined with the results of Theorems 7.3 and 7.4 permits evaluation of this invariant for $n \leq 4$. Assertion (1) reduces to the case ρ is static and Assertion (2) decouples the invariants as a sum of 3 different invariants. Assertion (3) evaluates the invariant which is independent of $\{p, \psi\}$, Assertion (4) evaluates invariant which depends on p, and Assertion (5) evaluates the invariant which depends on ψ.

Theorem 8.1 (1) *Expand the specific heat* $\rho(x;t) \sim \sum_{k \geq 0} t^k \rho_k(x)$ *in a Taylor series. Then* $\beta_n(\phi, \rho, D, \mathcal{B}, p, \psi) = \sum_{2k \leq n} \beta_{n-2k}(\phi, \rho_k, D, \mathcal{B}, p, \psi)$.

(2) *If the specific heat* ρ *is static, then* $\beta_n(\phi, \rho, D, \mathcal{B}, p, \psi) = \beta_n(\phi, \rho, D, \mathcal{B}, 0, 0) + \beta_n(0, \rho, D, \mathcal{B}, p, 0) + \beta_n(0, \rho, D, \mathcal{B}, 0, \psi)$.

(3) *If the specific heat* ρ *is static, then* $\beta_n(\phi, \rho, D, \mathcal{B}, 0, 0) = \beta_n(\phi, \rho, D, \mathcal{B})$.

(4) *Let* $c_{ij} := \int_0^1 (1-s)^i s^{j/2} ds$. *Expand* $p(x;t) \sim \sum_{k \geq 0} t^k p_k(x)$ *in a Taylor series. If the specific heat is static, then:*

a) $\beta_0(0, \rho, D, \mathcal{B}, p, 0) = 0$.

b) If $n > 0$, *then* $\beta_n(0, \rho, D, \mathcal{B}, p, 0) = \sum_{2i+j+2=n} c_{ij} \beta_j(p_i, \rho, D, \mathcal{B})$.

(5) *Expand the boundary source term* $\psi(x,t) \sim \sum_{k \geq 0} t^k \psi_k(x)$ *in a Taylor series. Assume the specific heat* ρ *is static. Then:*

a) $\beta_0(0, \rho, D, \mathcal{B}, 0, \psi) = 0$.

b) $\beta_1(0, \rho, D, \mathcal{B}, 0, \psi) = \frac{2}{\sqrt{\pi}} \int_{C_D} \langle \psi_0, \rho \rangle dy$.

c) $\beta_2(0, \rho, D, \mathcal{B}, 0, \psi) = -\int_{C_D} \{\langle \frac{1}{2} L_{aa} \psi_0, \rho \rangle - \langle \psi_0, \rho_{;m} \rangle\} dy - \int_{C_R} \langle \psi_0, \rho \rangle dy$.

d) $\beta_3(0, \rho, D, \mathcal{B}, 0, \psi) = \frac{2}{\sqrt{\pi}} \int_{C_D} \{\frac{2}{3} \langle \psi_0, \rho_{;mm} \rangle + \frac{1}{3} \langle \psi_0, \rho_{:aa} \rangle + \langle \frac{1}{3} E\psi, \rho \rangle$
$- \frac{2}{3} L_{aa} \langle \psi_0, \rho_{;m} \rangle + \langle (\frac{1}{12} L_{aa} L_{bb} - \frac{1}{6} L_{ab} L_{ab} - \frac{1}{6} R_{amma}) \psi_0, \rho \rangle\} dy$
$- \frac{4}{3\sqrt{\pi}} \int_{C_R} \langle \psi_0, \tilde{\mathcal{B}} \rho \rangle dy + \frac{4}{3\sqrt{\pi}} \int_{C_D} \langle \psi_1, \rho \rangle dy$.

e) $\beta_4(0, \rho, D, \mathcal{B}, 0, \psi) = -\int_{C_D} \{\frac{1}{2} \langle \psi_0, (\tilde{D}\rho)_{;m} \rangle - \frac{1}{4} \langle L_{aa} \psi_0, \tilde{D}\rho \rangle + \langle (\frac{1}{8} E_{;m}$
$- \frac{1}{16} L_{ab} L_{ab} L_{cc} + \frac{1}{8} L_{ab} L_{ac} L_{bc} - \frac{1}{16} R_{ambm} L_{ab} + \frac{1}{16} R_{abcb} L_{ac}$
$+ \frac{1}{32} \tau_{;m} + \frac{1}{16} L_{ab:ab}) \psi_0, \rho \rangle - \frac{1}{4} L_{ab} \langle \psi_{0:a}, \rho_{:b} \rangle - \frac{1}{8} \langle \Omega_{am} \psi_{0:a}, \rho \rangle$
$+ \frac{1}{8} \langle \Omega_{am} \psi_0, \rho_{:a} \rangle + \frac{1}{4} L_{aa} \langle \psi_1, \rho \rangle - \frac{1}{2} \langle \psi_1, \rho_{;m} \rangle\} dy$
$- \int_{C_R} \{-\frac{1}{2} \langle \psi_0, \tilde{D}\rho \rangle + \langle (\frac{1}{2} S + \frac{1}{4} L_{aa}) \psi_0, \tilde{\mathcal{B}}\rho \rangle + \frac{1}{2} \langle \psi_1, \rho \rangle\} dy$.

9 Time dependent phenomena

We refer to [56] for proofs of the assertions in this section and also for a more complete historical discussion. Let $\mathfrak{D} = \{D_t\}$ be a time-dependent family of operators of Laplace type. We expand $\mathfrak{D}$ in a Taylor series expansion

$$D_t u := Du + \sum_{r=1}^{\infty} t^r \Big\{ \mathcal{G}_{r,ij} u_{;ij} + \mathcal{F}_{r,i} u_{;i} + \mathcal{E}_r u \Big\} .$$

We use the initial operator $D := D_0$ to define a reference metric g_0. Choose local frames $\{e_i\}$ for the tangent bundle of M and local frames $\{e_a\}$ for the tangent bundle of the boundary which are orthonormal with respect to the initial metric g_0. Use g_0 to define the measures dx on M and dy on ∂M. The metric g_0 defines the curvature tensor R and the second fundamental form L. We also use D to define a background connection ∇_0 that we use to multiply covariantly differentiate tensors of all types and to define the endomorphism E.

As in Section 8, we again assume $\partial M = C_D \cup C_R$ decomposes as a disjoint union of two closed, possibly empty, disjoint subsets. We consider a 1 parameter family $\mathfrak{B} = \{\mathcal{B}_t\}$ of boundary operators which we expand formally in a Taylor series

$$\mathcal{B}_t \phi := \phi \Big|_{C_D} \oplus \Big\{ \phi_{;m} + S\phi + \sum_{r>0} t^r (\Gamma_{r,a} \phi_{;a} + S_r \phi) \Big\} \Big|_{C_R} .$$

The reason for including a dependence on time in the boundary condition comes, for example, by considering the dynamical Casimir effect. Slowly moving boundaries give rise to such boundary conditions. We let u be the solution of the time-dependent heat equation

$$(\partial_t + D_t) u = 0, \quad \mathcal{B}_t u = 0, \quad \lim_{t \downarrow 0} u(\cdot; t) = \phi(\cdot) \text{ in } L^2 .$$

There is a smooth kernel function so that $u(x;t) = \int_M K(t, x, \bar{x}, \mathfrak{D}, \mathfrak{B}) \phi(\bar{x}) d\bar{x}$. The analogue of the heat trace expansion in this setting and of the heat content asymptotic expansion are given, respectively, by

$$\int_M f(x) \operatorname{Tr}_{V_x} \Big\{ K(t, x, x, \mathfrak{D}, \mathfrak{B}) \Big\} dx \sim \sum_{n=0}^{\infty} a_n(f, \mathfrak{D}, \mathfrak{B}) t^{(n-m)/2},$$

$$\int_M \langle K(t, x, \bar{x}, \mathfrak{D}, \mathfrak{B}) \phi(x), \rho(\bar{x}) \rangle dx d\bar{x} \sim \sum_{n=0}^{\infty} \beta_n(\phi, \rho, \mathfrak{D}, \mathfrak{B}) t^{n/2} .$$

By assumption, the operators $\mathcal{G}_{r,ij}$ are scalar. The following theorem describes the additional terms in the heat trace asymptotics which arise from the structures described by $\mathcal{G}_{r,ij}$, $\mathcal{F}_{r,i}$, $\mathcal{E}_r$, $\Gamma_{r,a}$, and S_r given above.

Theorem 9.1 *[Varying geometries]*

(1) $a_0(F, \mathfrak{D}, \mathfrak{B}) = a_0(F, D, \mathcal{B})$.

(2) $a_1(F, \mathfrak{D}, \mathfrak{B}) = a_1(F, D, \mathcal{B})$.

(3) $a_2(F, \mathfrak{D}, \mathfrak{B}) = a_2(F, D, \mathcal{B}) + (4\pi)^{-m/2} \frac{1}{6} \int_M \operatorname{Tr}\{ \frac{3}{2} F \mathcal{G}_{1,ii} \} dx$.

(4) $a_3(F,\mathfrak{D},\mathfrak{B}) = a_3(F,D,\mathcal{B}) + (4\pi)^{(1-m)/2}\frac{1}{384}\int_{C_D}\operatorname{Tr}\{-24F\mathcal{G}_{1,aa}\}dy$
$+(4\pi)^{(1-m)/2}\frac{1}{384}\int_{C_R}\operatorname{Tr}\{24F\mathcal{G}_{1,aa}\}dy.$

(5) $a_4(F,\mathfrak{D},\mathfrak{B}) = a_4(F,D,\mathcal{B}) + (4\pi)^{-m/2}\frac{1}{360}\int_M \operatorname{Tr}\{F(\frac{45}{4}\mathcal{G}_{1,ii}\mathcal{G}_{1,jj} + \frac{45}{2}\mathcal{G}_{1,ij}\mathcal{G}_{1,ij}$
$+60\mathcal{G}_{2,ii} - 180\mathcal{E}_1 + 15\mathcal{G}_{1,ii}\tau - 30\mathcal{G}_{1,ij}\rho_{ij} + 90\mathcal{G}_{1,ii}E + 60\mathcal{F}_{1,i;i} + 15\mathcal{G}_{1,ii;jj}$
$-30\mathcal{G}_{1,ij;ij})\}dx + (4\pi)^{-m/2}\frac{1}{360}\int_{C_D}\operatorname{Tr}\{f(30\mathcal{G}_{1,aa}L_{bb} - 60\mathcal{G}_{1,mm}L_{bb}$
$+30\mathcal{G}_{1,ab}L_{ab} + 30\mathcal{G}_{1,mm;m} - 30\mathcal{G}_{1,aa;m} - 30\mathcal{F}_{1,m}) + F_{;m}(-45\mathcal{G}_{1,aa}$
$+45\mathcal{G}_{1,mm})\}dy + (4\pi)^{-m/2}\frac{1}{360}\int_{C_R}\operatorname{Tr}\{F(30\mathcal{G}_{1,aa}L_{bb} + 120\mathcal{G}_{1,mm}L_{bb}$
$-150\mathcal{G}_{1,ab}L_{ab} - 60\mathcal{G}_{1,mm;m} + 60\mathcal{G}_{1,aa;m} + 150\mathcal{F}_{1,m} + 180S\mathcal{G}_{1,aa} - 180S\mathcal{G}_{1,mm}$
$+360S_1) + F_{;m}(45\mathcal{G}_{1,aa} - 45\mathcal{G}_{1,mm})\}dy.$

Next we study the heat content asymptotics for variable geometries. We have the following formulas for Dirichlet and for Robin boundary conditions. Let $\mathcal{B} := \mathfrak{B}_0$.

Theorem 9.2 *[Dirichlet boundary conditions]*

(1) $\beta_n(\phi,\rho,\mathfrak{D},\mathcal{B}) = \beta_n(\phi,\rho,D_0,\mathcal{B})$ *for* $n = 0,1,2$.

(2) $\beta_3(\phi,\rho,\mathfrak{D},\mathcal{B}) = \beta_3(\phi,\rho,D_0,\mathcal{B}) + \frac{1}{2\sqrt{\pi}}\int_{C_D}\langle\mathcal{G}_{1,mm}\phi,\rho\rangle dy.$

(3) $\beta_4(\phi,\rho,\mathfrak{D},\mathcal{B}) = \beta_4(\phi,\rho,D_0,\mathcal{B}) - \frac{1}{2}\int_M\langle\mathcal{G}_{1,ij}\phi_{;ij} + \mathcal{F}_{1,i}\phi_{;i} + \mathcal{E}_1\phi,\rho\rangle dx$
$+\int_{C_D}\{\frac{7}{16}\langle\mathcal{G}_{1,mm;m}\phi,\rho\rangle - \frac{9}{16}L_{aa}\langle\mathcal{G}_{1,mm}\phi,\rho\rangle - \frac{5}{16}\langle\mathcal{F}_{1,m}\phi,\rho\rangle$
$+\frac{5}{16}L_{ab}\langle\mathcal{G}_{1,ab}\phi,\rho\rangle - \frac{5}{8}\langle\mathcal{G}_{1,am}\phi_{:a},\rho\rangle + \frac{1}{2}\langle\mathcal{G}_{1,mm}\phi,\rho_{;m}\rangle\}dy$
$+\int_{C_R}\{-\frac{1}{2}\langle\mathcal{G}_{1,mm}\mathcal{B}_0\phi,\rho\rangle + \frac{1}{2}\langle(S_1 + \Gamma_a\nabla_{e_a})\phi,\rho\rangle\}dy.$

10 Spectral boundary conditions

We adopt the notation used to discuss spectral boundary conditions in Section 6. Let $P : C^\infty(V_1) \to C^\infty(V_2)$ be an elliptic complex of Dirac type. Let $D = P^*P$ and let $\mathcal{B}_\Theta$ be the spectral boundary conditions defined by the auxiliary self-adjoint endomorphism Θ of V_1. Let ∇ be a compatible connection. Expand $P = \gamma_i\nabla_{e_i} + \psi$.

We begin by studying the heat trace asymptotics with spectral boundary conditions. There is an asymptotic series

$$\operatorname{Tr}_{L^2}(fe^{-tD_{\mathcal{B}_\Theta}}) \sim \sum_{k=0}^{m-1} a_k(f,D_{\mathcal{B}_\Theta},\mathcal{B}_\Theta)t^{(k-m)/2} + O(t^{-1/8})\,.$$

Continuing further introduces non-local terms; we refer to Atiyah et al. [6], to Grubb [71, 72], and to Grubb and Seeley [74, 75] for further details. Define $\gamma_a^T := \gamma_m^{-1}\gamma_a$, $\hat\psi := \gamma_m^{-1}\psi$, and $\beta(m) := \Gamma(\frac{m}{2})\Gamma(\frac{1}{2})^{-1}\Gamma(\frac{m+1}{2})^{-1}$. We refer to [48] for the proof of the following result:

Theorem 10.1 *[Spectral boundary conditions] Let* $f \in C^\infty(M)$. *Then:*

(1) $a_0(f,D,\mathcal{B}_\Theta) = (4\pi)^{-m/2}\int_M \operatorname{Tr}(f\,\mathrm{id})dx.$

(2) *If* $m \geq 2$, *then* $a_1(f, D, \mathcal{B}_\Theta) = \frac{1}{4}[\beta(m) - 1](4\pi)^{-(m-1)/2} \int_{\partial M} \mathrm{Tr}(f \,\mathrm{id}) dy$.

(3) *If* $m \geq 3$, *then* $a_2(f, D, \mathcal{B}_\Theta) = (4\pi)^{-m/2} \int_M \frac{1}{6} \mathrm{Tr}\{f(\tau \,\mathrm{id} + 6E)\} dx$
$+(4\pi)^{-m/2} \int_{\partial M} \mathrm{Tr}\{\frac{1}{2}[\hat{\psi} + \hat{\psi}^*] f + \frac{1}{3}[1 - \frac{3}{4}\pi\beta(m)] L_{aa} f \,\mathrm{id}$
$-\frac{m-1}{2(m-2)}[1 - \frac{1}{2}\pi\beta(m)] f_{;m} \,\mathrm{id}\} dy$.

(4) *If* $m \geq 4$, *then* $a_3(f, D, \mathcal{B}_\Theta) = (4\pi)^{-(m-1)/2} \int_{\partial M} \mathrm{Tr}\{\frac{1}{32}(1 - \frac{\beta(m)}{m-2}) f(\hat{\psi}\hat{\psi} + \hat{\psi}^*\hat{\psi}^*)$
$+\frac{1}{16}(5 - 2m + \frac{7-8m+2m^2}{m-2}\beta(m)) f\hat{\psi}\hat{\psi}^* - \frac{1}{48}(\frac{m-1}{m-2}\beta(m) - 1) f\tau \,\mathrm{id}$
$+\frac{1}{32(m-1)}(2m - 3 - \frac{2m^2-6m+5}{m-2}\beta(m)) f(\gamma_a^T\hat{\psi}\gamma_a^T\hat{\psi} + \gamma_a^T\hat{\psi}^*\gamma_a^T\hat{\psi}^*)$
$+\frac{1}{16(m-1)}(1 + \frac{3-2m}{m-2}\beta(m)) f\gamma_a^T\hat{\psi}\gamma_a^T\hat{\psi}^* + \frac{1}{48}(1 - \frac{4m-10}{m-2}\beta(m)) f\rho_{mm} \,\mathrm{id}$
$+\frac{1}{48(m+1)}(\frac{17+5m}{4} + \frac{23-2m-4m^2}{m-2}\beta(m)) f L_{ab} L_{ab} \,\mathrm{id}$
$+\frac{1}{48(m^2-1)}(-\frac{17+7m^2}{8} + \frac{4m^3-11m^2+5m-1}{m-2}\beta(m)) f L_{aa} L_{bb} \,\mathrm{id}$
$+\frac{1}{8(m-2)}\beta(m) f(\Theta\Theta + \frac{1}{m-1}\gamma_a^T\Theta\gamma_a^T\Theta)\} + \frac{m-1}{16(m-3)}(2\beta(m) - 1) f_{;mm} \,\mathrm{id}$
$+\frac{1}{8(m-3)}(\frac{5m-7}{8} - \frac{5m-9}{3}\beta(m)) L_{aa} f_{;m} \,\mathrm{id}\} dy$.

We now study heat content asymptotics with spectral boundary conditions. To simplify the discussion, we suppose P is formally self-adjoint. We refer to [60, 61] for the proof of:

Theorem 10.2 (1) $\beta_0(\phi, \rho, D, \mathcal{B}_\Theta) = \int_M \langle\phi, \rho\rangle dx$.

(2) $\beta_1(\phi, \rho, D, \mathcal{B}_\Theta) = -\frac{2}{\sqrt{\pi}} \int_{\partial M} \langle \Pi_A^+ \phi, \Pi_{A\#}^+ \rho\rangle dy$.

(3) $\beta_2(\phi, \rho, D, \mathcal{B}_\Theta) = -\int_M \langle D\phi, \rho\rangle dx + \int_{\partial M}\{-\langle\gamma_m \Pi_A^+ P\phi, \rho\rangle - \langle\gamma_m \Pi_A^+\phi, \tilde{P}\rho\rangle$
$+\frac{1}{2}\langle(L_{aa} + A + \tilde{A}^\# - \gamma_m\psi_P + \psi_P\gamma_m - \psi_A - \tilde{\psi}_A^\#)\Pi_A^+\phi, \Pi_{A\#}^+\rho\rangle\} dy$.

11 Operators which are not of Laplace type

We follow Avramidi and Branson [12], Branson et al. [32], Fulling [52], Gusynin [77], and Ørsted and Pierzchalski [101] to discuss the heat trace asymptotics of *non-minimal operators*. Let M be a compact Riemannian manifold with smooth boundary and let $\mathcal{B}$ define either absolute or relative boundary conditions. Let $E \in C^\infty(\mathrm{End}(\Lambda^p M))$ be an auxiliary endomorphism and let A and B be positive constants. Let

$$D_E^p := Ad\delta + B\delta d - E \quad \text{on} \quad C^\infty(\Lambda^p(M)),$$
$$c_{m,p}(A, B) := B^{-m} + (B^{-m} - A^{-m}) \sum_{k<p} (-1)^{k+p} \binom{m}{p}^{-1} \binom{m}{k}.$$

Theorem 11.1 (1) *If* $E = 0$, *then* $a_n(1, D^p, \mathcal{B}) = B^{(n-m)/2} a_n(1, \Delta_M^p, \mathcal{B})$
$+(B^{(n-m)/2} - A^{(n-m)/2}) \sum_{k<p} (-1)^{k+p} a_n(1, \Delta_M^p, \mathcal{B})$.

(2) *For general* E *one has:*

a) $a_0(1, D_E^p, \mathcal{B}) = a_0(1, D^p, \mathcal{B})$.

b) $a_1(1, D_E^p, \mathcal{B}) = a_1(1, D^p, \mathcal{B})$.

c) $a_2(1, D_E^p, \mathcal{B}) = a_2(1, D^p, \mathcal{B}) + (4\pi)^{-m/2} c_{m,p}(A, B) \int_M \operatorname{Tr}(E) dx$.

We follow the discussion in [56] to study the heat content asymptotics of the non-minimal operator $D := Ad\delta + B\delta d - E$ on $C^\infty(\Lambda^1(M))$. Let ϕ and ρ be smooth 1 forms; expand $\phi = \phi_i e_i$ and $\rho = \rho_i e_i$ where e_m is the inward geodesic normal.

Theorem 11.2 (1) *Let $\mathcal{B}$ define absolute boundary conditions. Then:*

a) $\beta_0(\phi, \rho, D, \mathcal{B}) = \int_M (\phi, \rho) dx$.

b) $\beta_1(\phi, \rho, D, \mathcal{B}) = -\frac{2}{\sqrt{\pi}} \sqrt{A} \int_{\partial M} \phi_m \rho_m dy$.

c) $$\beta_2(\phi, \rho, D, \mathcal{B}) = -\int_M \{A(\delta\phi, \delta\rho) + B(d\phi, d\rho) - E(\phi, \rho)\} dx + \int_{\partial M} A\{-\phi_m \rho_{a:a} - \phi_{a:a}\rho_m - \phi_{m;m}\rho_m - \phi_m \rho_{m;m} + \tfrac{3}{2} L_{aa} \phi_m \rho_m\} dy.$$

(2) *Let $\mathcal{B}$ define relative boundary conditions. Then*

a) $\beta_0(\phi, \rho, D, \mathcal{B}) = \int_M (\phi, \rho) dx$.

b) $\beta_1(\phi, \rho, D, \mathcal{B}) = -\frac{2}{\sqrt{\pi}} \sqrt{B} \int_{\partial M} \phi_a \rho_a dy$.

c) $$\beta_2(\phi, \rho, D, \mathcal{B}) = -\int_M \{A(\delta\phi, \delta\rho) + B(d\phi, d\rho) - E(\phi, \rho)\} dx + \int_{\partial M} B\{-\phi_{a:a}\rho_m - \phi_m \rho_{a:a} - \phi_{a;m}\rho_a - \phi_a \rho_{a;m} + L_{ab}\phi_b \rho_a + \tfrac{1}{2} L_{aa} \phi_b \rho_b\} dy.$$

We now turn our attention to fourth order operators. Let M be a closed Riemannian manifold. Let ∇ be a connection on a vector bundle V over a closed Riemannian manifold M. Set

$$\Gamma(\tfrac{m-n}{2})^{-1}\Gamma(\tfrac{m-n}{4}) := \lim_{s \to n} \{\Gamma(\tfrac{m-s}{2})^{-1}\Gamma(\tfrac{m-s}{4})\}.$$

Theorem 11.3 *Let $Pu = u_{;iijj} + p_{2,ij}u_{;ij} + p_{1,i}u_{;i} + p_0$ on a closed Riemannian manifold where $p_{2,ij} = p_{2,ji}$ and where $\{p_{2,ij}, p_{1,i}, p_0\}$ are endomorphism valued. . Then:*

(1) $a_0(1, P) = \frac{1}{2}(4\pi)^{-m/2}\Gamma(\frac{m}{2})^{-1}\Gamma(\frac{m}{4}) \int_M \operatorname{Tr}(\mathrm{id}) dx$.

(2) $a_2(1, P) = \frac{1}{2}(4\pi)^{-m/2}\Gamma(\frac{m-2}{2})^{-1}\Gamma(\frac{m-2}{4}) \frac{1}{6} \int_M \operatorname{Tr}\{\tau \,\mathrm{id} + \frac{3}{m} p_{2,ii}\} dx$.

(3) $$a_4(1, P) = \tfrac{1}{2}(4\pi)^{-m/2}\Gamma(\tfrac{m}{2})^{-1}\Gamma(\tfrac{m}{4}) \tfrac{1}{360} \int_M \operatorname{Tr}\{\tfrac{90}{m+2} p_{2,ij}p_{2,ij} + \tfrac{45}{m+2} p_{2,ii}p_{2,jj} + (m-2)(5\tau^2 \,\mathrm{id} - 2|\rho|^2 \,\mathrm{id} + 2|R|^2 \,\mathrm{id} + 30\Omega_{ij}\Omega_{ij}) + 30\tau p_{2,ii} - 60\rho_{ij}p_{2,ij} - 360 p_0\} dx.$$

12 The spectral geometry of Riemannian submersions

We refer to [64] for further details concerning the material of this section; additionally see Bergery and Bourguignon[13], Besson and Bordoni [29], Goldberg and Ishihara [65] and Watson [126]. Let $\pi : Z \to Y$ be a smooth map where Z and Y are connected closed Riemannian manifolds. We say that π is a submersion if π is surjective and if $\pi_* : T_z Z \to T_{\pi z} Y$ is surjective for every $z \in Z$.

Submersions are fiber bundles. Let $\mathcal{F} := \pi^{-1}(y_0)$ be the fiber over some point $y_0 \in Y$. If $\mathcal{O}$ is a contractable open subset of Y, then $\pi^{-1}(\mathcal{O})$ is homeomorphic to $\mathcal{O} \times F$ and under this homeomorphism, π is projection on the first factor. The vertical distribution $\mathcal{V} := \ker(\pi_*)$ is a smooth subbundle of TZ. The horizontal distribution is defined by $\mathcal{H} := \mathcal{V}^\perp$. One says that π is a Riemannian submersion if $\pi_* : \mathcal{H}_z \to T_{\pi z} Y$ is an isometry for every point z in Z.

The fundamental tensors may be introduced as follows. Let $\pi : Z \to Y$ be a Riemannian submersion. We use indices a, b, c to index local orthonormal frames $\{f_a\}$, $\{f^a\}$, $\{F_a\}$, and $\{F^a\}$ for $\mathcal{H}$, $\mathcal{H}^*$, TY, and T^*, respectively. We use indices i, j, k to index local orthonormal frames $\{e_i\}$ and $\{e^i\}$ for $\mathcal{V}$ and $\mathcal{V}^*$, respectively. There are two fundamental tensors which arise naturally in this setting. The unnormalized mean curvature vector θ and the integrability tensor ω are defined by:

$$\begin{aligned}
&\theta := -g_Z([e_i, f_a], e_i) f^a = {}^Z\Gamma_{iia} f^a \in C^\infty(\mathcal{H}),\\
&\omega := \omega_{abi} = \tfrac{1}{2} g_Z(e_i, [f_a, f_b]) = \tfrac{1}{2}({}^Z\Gamma_{abi} - {}^Z\Gamma_{bai}) .
\end{aligned}$$

Lemma 12.1 *Let $\pi : Z \to Y$ be a Riemannian submersion.*

(1) *The following assertions are equivalent:*
a) The fibers of π are minimal. b) π is a harmonic map. c) $\theta = 0$.

(2) *The following assertions are equivalent:*
a) The distribution $\mathcal{H}$ is integrable. b) $\omega = 0$.

(3) *Let $\Theta := \pi_* \theta$ be the integration of θ along the fiber, and let $V(y)$ be the volume of the fiber. Then $\Theta = -d_Y \ln(V)$. Thus in particular, if $\theta = 0$, then the fibers have constant volume.*

By naturality $\pi^* d_Y = d_Z \pi^*$. The intertwining formulas for the coderivatives and for the Laplacians are more complicated. Let $\mathcal{E} := \omega_{abi} \operatorname{ext}_Z(e^i) \operatorname{int}_Z(f^a) \operatorname{int}_Z(f^b)$ and let $\Xi := \operatorname{int}_Z(\theta) + \mathcal{E}$.

Lemma 12.2 *Let $\pi : Z \to Y$ be a Riemannian submersion. Then $\delta_Z \pi^* - \pi^* \delta_Y = \Xi \pi^*$ and $\Delta_Z^p \pi^* - \pi^* \Delta_Y^p = \{\Xi d_Z + d_Z \Xi\} \pi^*$.*

One is interested in relating the spectrum on the base to the spectrum on the total space. The situation is particularly simple if $p = 0$:

Theorem 12.3 *Let $\pi : Z \to Y$ be a Riemannian submersion.*

(1) *If $\Phi \in E(\lambda, \Delta_Y^0)$ is nontrivial and if $\pi^* \Phi \in E(\mu, \Delta_Z^0)$, then $\lambda = \mu$.*

(2) *The following conditions are equivalent:*
a) $\Delta_Z^0 \pi^* = \pi^* \Delta_Y^0$. b) For all λ, $\pi^* E(\lambda, \Delta_Y^0) \subset E(\lambda, \Delta_Z^0)$. c) $\theta = 0$.

Muto [97, 98, 99] has given examples of Riemannian principal S^1 bundles where eigenvalues can change. The study of homogeneous space also provides examples. This leads to the result:

Theorem 12.4 (1) *Let Y be a homogeneous manifold with $H^2(Y;\mathbb{R}) \neq 0$. There exists a complex line bundle L over Y with associated circle fibration $\pi_S : S(L) \to Y$, and there exists a unitary connection ${}^L\nabla$ on L so that the curvature $\mathcal{F}$ of ${}^L\nabla$ is harmonic and has constant norm $\epsilon \neq 0$ and so that $\pi_S^*\mathcal{F} \in E(\epsilon, \Delta_S^2)$.*

(2) *Let $0 \leq \lambda \leq \mu$ and let $p \geq 2$ be given. There exists a principal circle bundle $\pi : P \to Y$ over some manifold Y, and there exists $0 \neq \Phi \in E(\lambda, \Delta_Y^p)$ so that $\pi^*\Phi \in E(\mu, \Delta_Z^p)$.*

The case $p = 1$ is unsettled; it is not known if eigenvalues can change if $p = 1$. On the other hand, one can show that eigenvalue can never decrease.

Theorem 12.5 *Let $\pi : Z \to Y$ be a Riemannian submersion of closed smooth manifolds. Let $1 \leq p \leq \dim(Y)$. If $0 \neq \Phi \in E(\lambda, \Delta_Y^p)$ and if $\pi^*\Phi \in E(\mu, \Delta_Z^p)$, then $\lambda \leq \mu$. The following conditions are equivalent:*
a) We have $\Delta_Z^p\pi^* = \pi^*\Delta_Y^p$.
b) For all λ, we have $\pi^*E(\lambda, \Delta_Y^p) \subset E(\lambda, \Delta_Z^p)$.
c) For all λ, there exists $\mu = \mu(\lambda)$ so $\pi^*E(\lambda, \Delta_Y^p) \subset E(\mu, \Delta_Z^p)$.
d) We have $\theta = 0$ and $\omega = 0$.

Results of Park [103] show this if Neumann boundary conditions are imposed on a manifolds with boundary, then eigenvalues can decrease.

There are results related to finite Fourier series. We have $L^2(\Lambda^p M) = \oplus_\lambda E(\lambda, \Delta_M^p)$. Thus if ϕ is a smooth p-form, we may decompose $\phi = \sum_\lambda \phi_\lambda$ for $\phi_\lambda \in E(\lambda, \Delta_M^p)$. Let $\nu(\phi)$ be the number of λ so that $\phi_\lambda \neq 0$. We say that ϕ has *finite Fourier series* if $\nu(\phi) < \infty$. For example, if $M = S^1$, then ϕ has finite Fourier series if and only if ϕ is a trignometric polynomial. The first assertion in the following result is an immediate consequence of the Peter-Weyl theorem; the second result follows from [49].

Theorem 12.6 (1) *Let $\pi : G \to G/H$ be a homogeneous space where G/H is equipped with a G invariant metric and where G is equipped with a left invariant metric. If ϕ is a smooth p-form on G/H with finite Fourier series, then $\pi^*\phi$ has finite Fourier series on G.*

(2) *Let $1 \leq p$, $0 < \lambda$, and $2 \leq \mu_0$ be given. There exists $\pi : G \to G/H$ and there exists $\phi \in E(\lambda, \Delta_{G/H}^p)$ so that $\mu_G(\pi^*\phi) = \nu_0$.*

In general, there is no relation between the heat trace asymptotics on the base, fiber, and total space of a Riemannian submersion. McKean and Singer [90] have determined the heat equation asymptotics for the sphere S^n. Let

$$Z(M,t) := \frac{(4\pi t)^{m/2}}{\mathrm{Vol}(M)} \mathrm{Tr}_{L^2} e^{-t\Delta_M^0} \sim \sum_{n\geq 0} \frac{(4\pi t)^{m/2}}{\mathrm{Vol}(M)} a_n(\Delta_M^0) t^{n/2}$$

be the normalized heat trace; with this normalization, $Z(M,t)$ is regular at the origin and has leading coefficient 1. Their results (see page 63 of McKean and Singer [90]) show that

$$Z(S^1,t) = 1 + O(t^k) \text{ for any k}$$
$$Z(S^2,t) = \frac{e^{t/4}}{\sqrt{\pi t}} \int_0^1 \frac{e^{-x/t}}{\sin\sqrt{x}} dx = 1 + \frac{t}{3} + \frac{t^2}{15} + \ldots$$
$$Z(S^1 \times S^2, t) = Z(S^2,t)Z(S^1,t) = 1 + \frac{t}{3} + \frac{t^2}{15} + \ldots$$

$$Z(S^3, t) = e^t = 1 + t + \tfrac{1}{2}t^2 + \dots.$$

The two fibrations $\pi : S^1 \times S^2 \to S^2$ and $\pi : S^3 \to S^2$ have base S^2 and minimal fibers S^1. However, the heat trace asymptotics are entirely different.

On the other hand, the following result shows that the heat content asymptotics on Z *are* determined by the heat content asymptotics of the base and by the volume of the fiber if $\theta = 0$; Lemma 12.1 shows the volume V of the fiber is independent of the point in question in this setting.

Theorem 12.7 *Let $\pi : Z \to Y$ be a Riemannian submersion of compact manifolds with smooth boundary. Let $\rho_Z := \pi^*\rho_Y$ and let $\phi_Z := \pi^*\phi_Y$. If $\theta = 0$ and if $\mathcal{B} = \mathcal{B}_D$ or $\mathcal{B} = \mathcal{B}_N$, then $\beta_n(\rho_Z, \phi_Z, \Delta_Z^0, \mathcal{B}) = \beta_n(\rho_Y, \phi_Y, \Delta_Y^0, \mathcal{B}) \cdot V$.*

Acknowledgments

Research of P. Gilkey was partially supported by the Max Planck Institute in the Mathematical Sciences (Leipzig, Germany) and by Project MTM2006-01432 (Spain). It is a pleasure to acknowledge helpful conversations with E. Puffini and L. Loro concerning these matters.

Bibliography

[1] P. Amsterdamski, A. Berkin and D. O'Connor: b_8 Hamidew coefficient for a scalar field *Classical Quantum Gravity* **6** (1989) 1981–1991

[2] M. F. Atiyah and R. Bott: The index problem for manifolds with boundary *In: Differential Analysis - papers presented at the Bombay Coll. (1964)* (Oxford Univ. Press, 1965) 337–351

[3] M. F. Atiyah and R. Bott: A Lefschetz fixed point formula for elliptic differential operators *Bull Amer. Math. Soc.* **72** (1966) 245–256

[4] M. F. Atiyah and R. Bott: A Lefschetz fixed point formula for elliptic complexes (I,II) *Ann. of Math.* **86** (1967) 374–407; *Ann. of Math.* **88** (1968) 451–491

[5] M. F. Atiyah, R. Bott and V. K. Patodi: On the heat equation and the index theorem *Invent. Math.* **19** (1973) 279–330; *Errata* **28** (1975) 277–280

[6] M. F. Atiyah, V. K. Patodi and I. M. Singer: Spectral asymmetry and Riemannian geometry I, II, III *Math. Proc. Cambridge Philos. Soc.* **77** (1975) 43–69; *Math. Proc. Cambridge Philos. Soc.* **78** (1975) 405–432; *Math. Proc. Cambridge Philos. Soc.* **79** (1976) 71–99

[7] M. F. Atiyah and G. B. Segal: The index of elliptic operators III *Ann. of Math.* **87** (1968) 531–545

[8] M. F. Atiyah and I. M. Singer: The index of elliptic operators I, II, IV, V *Ann. of Math.* **87** (1968) 484–530; *Ann. of Math.* **87** (1968) 546–604; *Ann. of Math.* **93** (1971) 119–138; *Ann. of Math.* **93** (1971) 139–149

[9] I. G. Avramidi: The covariant technique for the calculation of the heat kernel asymptotic expansion *Phys. Lett. B* **238** (1990) 92–97

[10] I. G. Avramidi: Covariant methods of studying the nonlocal structure of an effective action *Sov. J. Nucl. Phys.* **49** (1989) 735–739

[11] I. G. Avramidi: Heat kernel asymptotics of Zaremba boundary value problem *Mathematical Physics, Analysis and Geometry* **7** (2004) 9-46

[12] I. G. Avramidi and T. Branson: Heat kernel asymptotics of operators with non-Laplace principal part *Rev. Math. Phys.* **13** (2001) 847–890

[13] L. Bérard Bergery and J. P. Bourguignon: Laplacians and Riemannian submersions with totally geodesic fibers *Illinois J. Math.* **26** (1982) 181–200

[14] M. van den Berg: Heat equation on a hemisphere *Proc. Edinb. Math. Soc.* **118A** (1991) 5–12

[15] M. van den Berg: Heat equation on the arithmetic von Koch snowflake *Probab. Theory Related Fields* **118** (2000) 17–36

[16] M. van den Berg and P. Gilkey: Heat content asymptotics of a Riemannian manifold with boundary *J. Funct. Anal.* **120** (1994) 48–71

[17] M. van den Berg and P. Gilkey: Heat invariants for odd dimensional hemispheres *Proc. Edinb. Math. Soc.* **126A** (1996) 187–193

[18] M. van den Berg and P. Gilkey: The heat equation with inhomogeneous Dirichlet boundary conditions *Comm. Anal. Geom.* **7** (1999) 279–294

[19] M. van den Berg, S. Desjardins and P. Gilkey: Functorality and heat content asymptotics for operators of Laplace type *Topol. Methods Nonlinear Anal.* **2** (1993) 147–162

[20] M. van den Berg and E. B. Davies: Heat flow out of regions in $\mathbb{R}^m$ *Math. Z.* **202** (1989) 463–482

[21] M. van den Berg and J-F Le Gall: Mean curvature and the heat equation *Math. Z.* **215** (1994) 437–464

[22] M. van den Berg and P. Gilkey: The heat content asymptotics of a time dependent process *Proc. Edinb. Math. Soc.* **130A** (2000) 307–312

[23] M. van den Berg and P. Gilkey: Heat content asymptotics with inhomogeneous Neumann and Dirichlet boundary conditions *Potential Anal.* **14** (2001) 269–274

[24] M. van den Berg, P. Gilkey, K. Kirsten and V. A. Kozlov: Heat Content Asymptotics for Riemannian manifolds with Zaremba boundary conditions *Potential Analysis* **26** (2007) 225–254; `arXiv:math-ph/0506076`

[25] M. van den Berg and S. Srisatkunarajah: Heat flow and Brownian motion for a region in $\mathbb{R}^2$ with a polygonal boundary *Probab. Th. Rel. Fields* **86** (1990) 41–52

[26] P. Bérard: Variétés Riemanniennes isospectrales non isométriques *In: Séminaire Bourbaki* (41, 1989)

[27] M. Berger: Sur le spectre des variétés Riemanniennes *Rev. Roum. Math. Pure Appl.* **13** (1969) 915–931

[28] N. Berline, N. Getzler and M. Vergne: *Heat kernels and Dirac operators* (Grundlehren Band 298, Springer-Verlag, Berlin, 1991)

[29] G. Besson and M. Bordoni: On the spectrum of Riemannian submersions with totally geodesic fibers *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. (9) Mat. Appl.* **1** (1990) 335–340

[30] J.-M. Bismut: Index theory and the heat equation *In: Proc. ICM Berkeley 1986* (Vol. I, 1987) 491–504

[31] M. Bordag and D. V. Vassilevich: Heat kernel expansion for semitransparent boundaries *J. Phys. A* **32** (1999) 8247–8259

[32] T. Branson, S. Fulling and P. Gilkey: Heat equation asymptotics of 'nonminimal' operators on differential forms *J. Math. Phys.* **32** (1991) 2089–2091

[33] T. Branson and P. Gilkey: Residues of the eta function for an operator of Dirac Type *J. Funct. Anal.* **108** (1992) 47–87

[34] T. Branson and P. Gilkey: Residues of the eta function for an operator of Dirac type with local boundary conditions *Differential Geom. Appl.* **2** (1992) 249–267

[35] T. Branson, P. Gilkey, K. Kirsten and D. Vassilevich: Heat kernel asymptotics with mixed boundary conditions *Nuclear Phys. B* **563** (1999) 603–626

[36] T. Branson, P. Gilkey and B. Ørsted: Leading terms in the heat invariants *Proc. Amer. Math. Soc.* **109** (1990) 437–450

[37] R. Brooks and C. Gordon: Isospectral families of conformally equivalent Riemannian metrics *Bull. Amer. Math. Soc.* **23** (1990) 433–436

[38] R. Brooks, P. Perry and P. Petersen: Finiteness of diffeomorphism types of isospectral manifolds *In: Proc. Symp. Pure Math.* (54, Part 3 1993) 89–94

[39] R. Brooks, P. Perry and P. Yang: Isospectral sets of conformally equivalent metrics *Duke Math. J.* **58** (1989) 131–150

[40] T. Carleman: Propriétés asymptotiques des fonctions fondamentales des membranes vibrantes *In: Åttonde Skand Matematikerkongressen i Stockholm, 1934* (, 1935) 34–44

[41] H. S. Carslaw and J. C. Jaeger: *Conduction of Heat in Solids* (Oxford Science Publications, Oxford Univ. Press, 1946)

[42] A. Chang and P. Yang: Isospectral conformal metrics on 3 manifolds *J. Amer. Math. Soc.* **3** (1990) 117–145

[43] S. Chern: A simple intrinsic proof of the Gauss-Bonnet formula for closed Riemannian manifolds *Ann. of Math.* **45** (1944) 741–752

[44] S. Chern: On the curvature integra in a Riemannian manifold *Ann. of Math.* **46** (1945) 674–684

[45] S. Desjardins and P. Gilkey: Heat content asymptotics for operators of Laplace type with Neumann boundary conditions *Math. Z.* **215** (1994) 251–268

[46] J. S. Dowker: The hybrid spectral problem and Robin boundary conditions *J. Phys. A* **38** (2005) 4735–4754

[47] J. S. Dowker: The $N \cup D$ problem *Preprint* `arXiv:hep-th/0007129` (2000)

[48] J. S. Dowker, P. Gilkey and K. Kirsten: Heat asymptotics with spectral boundary conditions *Contemp. Math.* **242** (1999) 107–124

[49] C. Dunn, P. Gilkey and J. H. Park: The spectral geometry of the canonical Riemannian submersion of a compact Lie Group *to appear Journal of Geometry and Physics* (2007)

[50] H. Fegan and P. Gilkey: Invariants of the heat equation *Pacific Journal of Mathematics* **117** (1985) 233–254

[51] J. Fleckinger, M. Levitin and D. Vassiliev: Heat content of the triadic von Koch snowflake *Internat. J. Appl. Sci. Comput.* **2** (1995) 289–305

[52] S. Fulling: Kernel asymptotics of exotic second-order operators *In: Proceedings of the Third International Colloquium on Differential Equations* (Plovdiv, 1992) 63–76

[53] P. Gilkey: The spectral geometry of a Riemannian manifold *J. Differential Geom.* **10** (1975) 601–618

[54] P. Gilkey: *Invariance Theory, the Heat Equation, and the Atiyah-Singer index theorem* (Second edition. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1995)

[55] P. Gilkey: The heat content asymptotics for variable geometries *J. Phys. A* **32** (1999) 2825–2834

[56] P. Gilkey: *Asymptotic Formulas in Spectral Geometry* (Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 2004)

[57] P. Gilkey and K. Kirsten: Heat content asymptotics with transmittal and transmission boundary *J. London Math Soc.* **68** (2003) 431–443

[58] P. Gilkey, K. Kirsten and J. H. Park: Heat trace asymptotics of a time dependent process *J. Phys. A* **34** (2001) 1153–1168

[59] P. Gilkey, K. Kirsten and J. H. Park: Heat content asymptotics for oblique boundary conditions *Lett. Math. Phys.* **59** (2002) 269–276

[60] P. Gilkey, K. Kirsten and J. H. Park: Heat content asymptotics for spectral boundary conditions *Trends in Math.* **5** (2002) 49–58

[61] P. Gilkey, K. Kirsten and J. H. Park: Heat Content asymptotics for operators of Laplace type with spectral boundary conditions *Lett. Math. Phys.* **68** (2004) 67–76

[62] P. Gilkey, K. Kirsten and D. Vassilevich: Heat trace asymptotics with transmittal boundary conditions and quantum brane-world scenario *Nuclear Phys. B* **601** (2001) 125–148

[63] P. Gilkey, K. Kirsten and D. Vassilevich: Heat trace asymptotics defined by transfer boundary conditions *Lett. Math. Phys.* **63** (2003) 29–37

[64] P. Gilkey, J. V. Leahy and J. H. Park: *Spectral Geometry, Riemannian Submersions, and the Gromov-Lawson Conjecture* (CRC Press, 1999)

[65] S. I. Goldberg and T. Ishihara: Riemannian submersions commuting with the Laplacian *J. Differential Geom.* **13** (1978) 139–144

[66] C. Gordon: Survey of isospectral manifolds *In: Handbook of differential geometry Vol I* (North-Holland, Amsterdam, 2000) 747–778

[67] C. Gordon and E. Wilson: Isospectral deformations of compact solvmanifolds *J. Differential Geom.* **19** (1984) 241–256

[68] P. Greiner: An asymptotic expansion for the heat equation *In: Global Analysis, Proc. Symp. Pure Math., Berkeley 1968* (Amer. Math. Soc., Providence, R.I, 1970) 133–137

[69] P. Greiner: An asymptotic expansion for the heat equation *Arch. Rat. Mech. Anal.* **41** (1971) 163–218

[70] C. Griffith and M. L. Lapidus: Computer graphics and the eigenfunctions for the Koch snowflake drum *In: Progress in inverse spectral geometry* (Trends Math., 1997) 95–113

[71] G. Grubb: Heat operator trace expansions and index for general Atiyah Patodi Singer boundary problems *Comm. in PDE* **17** (1992) 2031–2077

[72] G. Grubb: Complete heat trace, resolvent and zeta expansions for general Atiyah Patodi Singer problems *Nonlinear Diff. Eq. Appl.* **21** (1996) 158–167

[73] G. Grubb: *Functional calculus of pseudodifferential boundary problems* (Prog. in Math., Birkhaeuser, 1996)

[74] G. Grubb and R. Seeley: Weakly parametric pseudodifferential operators and Atiyah Patodi Singer boundary problems *Invent. Math.* **121** (1995) 481–529

[75] G. Grubb and R. Seeley: Zeta and eta functions for Atiyah-Patodi-Singer operators *J. Geom. Anal.* **6** (1996) 31–77

[76] P. Günther: *Huygen's principle and hyperbolic equations* (Perspectives in Mathematics, 5. Academic Press, Inc., Boston, MA, 1988)

[77] V. P. Gusynin, E. V. Gorbar and V. V. Romankov: Heat kernel expansion for non minimal differential operators and manifolds with torsion *Nuclear Phys. B* **362** (1991) 449–471

[78] J. Hadamard: *Le problème de Cauchy et les èquations aux dérivées partielles linéaires hyperboliques* (Hermann et Cie, Paris, 1932)

[79] A. Ikeda: Riemannian manifolds p isospectral but not $p+1$ isospectral *In: Geometry of manifolds* (Coll. Pap. 35th Symp. Differ. Geom., Matsumoto/Japan 1988, Perspect. Math. 8, 1989) 383–417

[80] D. Jakobson, M. Levitin, N. Nadirashvili and I. Polterovich: Spectral problems with mixed Dirichlet-Neumann boundary conditions: isospectrality and beyond *J. Comput. Appl. Math.* **194** (2006) 141–155

[81] M. Kac: Can one hear the shape of a drum? *Amer. Math. Monthly* **783** (1966) 1–23

[82] G. Kennedy, R. Critchley and J. S. Dowker: Finite temperature field theory with boundaries: stress tensor and surface action renormalization *Ann. Physics* **125** (1980) 346–400

[83] K. Kirsten: The a_5 heat kernel coefficient on a manifold with boundary *Class. Quantum Grav.* **15** (1998) L5–L12

[84] K. Kirsten: *Spectral functions in mathematics and physics* (Chapman and Hall/CRC, Boca Raton, FL, 2001)

[85] M. Lapidus and M. Pang: Eigenfunctions of the Koch snowflake domain *Comm. Math. Phys.* **172** (1995) 359–376

[86] A. Lichnerowicz: Spineurs harmoniques *C. R. Acad. Sci. Paris Sér. I Math.* **257** (1963) 7–9

[87] D. M. McAvity: Heat kernel asymptotics for mixed boundary conditions *Classical Quantum Gravity* **9** (1992) 1983–1997

[88] D. M. McAvity: Surface energy from heat content asymptotics *J. Phys. A* **26** (1993) 823–830

[89] P. McDonald and R. Meyers: Diffusions on graphs, Poisson problems and spectral geometry *Trans. Amer. Math. Soc.* **354** (2002) 5111–5136

[90] H. P. McKean and I. M. Singer: Curvature and the eigenvalues of the Laplacian *J. Differential Geom.* **1** (1967) 43–69

[91] R. B. Melrose: *The Atiyah-Patodi-Singer Index Theorem* (Research Notes in Mathematics, 4, A K Peters, Ltd., Wellesley, MA, 1993)

[92] J. Milnor: Eigenvalues of the Laplace operator on certain manifolds *Proc. Nat. Acad. Sci. U.S.A.* **51** (1964) 542

[93] S. Minakshisundaram: Eigenfunctions on Riemannian manifolds *J. Indian Math. Soc.* **17** (1953) 158–165

[94] S. Minakshisundaram and Å. Pleijel: Some properties of the eigenfunctions of the Laplace operator on Riemannian manifolds *Canad. J. Math.* **1** (1949) 242–256

[95] I. G. Moss: Heat kernel expansions for distributional backgrounds *Phys. Lett. B* **491** (2000) 203–206

[96] I. Moss and J. S. Dowker: The correct B_4 coefficient *Phys. Lett. B* **229** (1989) 261–263

[97] Y. Muto: Riemannian submersions and critical Riemannian metrics *J. Math. Soc. Japan* **29** (1977) 493–511

[98] Y. Muto: Some eigenforms of the Laplace-Beltrami operators in a Riemannian submersion *J. Korean Math. Soc.* **15** (1978) 39–57

[99] Y. Muto: Riemannian submersion and the Laplace-Beltrami operator *Kodai Math. J.* **1** (1978) 329–338

[100] J. Neuberger, N. Sieben and J. Swift: Computing eigenfunctions on the Koch snowflake: a new grid and symmetry *J. Comput. Appl. Math.* **191** (2006) 126–142

[101] B. Ørsted and A. Pierzchalski: The Ahlfors Laplacian on a Riemannian manifold with boundary *Michigan Math. J.* **43** (1996) 99–122

[102] B. Osgood, R. Phillips and P. Sarnak: Compact isospectral sets of surfaces *J. Funct. Anal.* **80** (1988) 212–234

[103] J. H. Park: The spectral geometry of Riemannian submersions for manifolds with boundary *Rocky Mountain J. Math.* **30** (2000) 353–369

[104] J. H. Park: The spectral geometry of Einstein manifolds with boundary *J. Korean Math. Soc.* **41** (2004) 875–882

[105] V. K. Patodi: Curvature and the fundamental solution of the heat operator *J. Indian Math. Soc.* **34** (1970) 269–285

[106] V. K. Patodi: Curvature and the eigenforms of the Laplace operators *J. Differential Geom.* **5** (1971) 233–249

[107] V. K. Patodi: An analytic proof of the Riemann Roch Hirzebruch theorem for Kaehler manifolds *J. Differential Geom.* **5** (1971) 251–283

[108] C. G. Phillips and K. M. Jansons: The short-time transient of diffusion outside a conducting body *Proc. London Math. Soc.* **428** (1990) 431–449

[109] I. Polterovich: A commutator method for computation of heat invariants *Indag. Math. (N.S.)* **11** (2000) 139–149

[110] I. Polterovich: Heat invariants of Riemannian manifolds *Israel J. Math.* **119** (2000) 239–252

[111] T. Sakai: On eigen-values of Laplacian and curvature of Riemannian manifolds *Tohoku Math. J.* **23** (1971) 589–603

[112] A. Savo: Une méthode de symétrization et quelques applications *C. R. Acad. Sci. Paris Sér. I Math.* **322** (1996) 861–864

[113] A. Savo: Uniform estimates and the whole asymptotic series of the heat content on manifolds *Geom. Dedicata* **73** (1998) 181–214

[114] A. Savo: Heat content and mean curvature *Rend. Mat. Appl.* **18** (1998) 197–219

[115] A. Savo: On the asymptotic series of the heat content *Contemp. Math.* **288** (2001) 428–432

[116] R. Seeley: *Topics in pseudo-differential operators* (1969 Pseudo-Diff. Operators (C.I.M.E., Stresa, 1968), 167–305. Edizioni Cremonese, Rome; republished in Pitman Research Notes in Mathematics Series 359. Longman, Harlow, 1996)

[117] R. Seeley: Complex powers of an elliptic operator *In: Singular Integrals, Proc. Sympos. Pure Math., Chicago, Ill. 1966* (Amer. Math. Soc., Providence, R.I, 1967) 288–307

[118] R. Seeley: The resolvent of an elliptic boundary value problem *Amer. J. Math.* **91** (1969) 889–920

[119] R. Seeley: Analytic extension of the trace associated with elliptic boundary problems *Amer. J. Math.* **91** (1969) 963–983

[120] K. Stewartson and R. T. Waechter: On hearing the shape of a drum: further results *Proc. Cambridge Philos. Soc.* **69** (1971) 353–363

[121] T. Sunada: Riemannian coverings and isospectral manifolds *Ann. of Math.* **121** (1985) 169–186

[122] V. Tanno: A characterization of the canonical spheres by the spectrum *Math. Z.* **175** (1980) 267–274

[123] H. Urakawa: Bounded domains which are isospectral but not congruent *Ann. Scient. École Norm. Sup.* **15** (1982) 441–456

[124] A. E. van de Ven: Index-free heat kernel coefficients *Classical Quantum Gravity* **15** (1998) 2311–2344

[125] M. Vigneras: Variete Riemanniennes isospectrales et non isometriques *Ann. of Math.* **112** (1980) 21–32

[126] B. Watson: Manifold maps commuting with the Laplacian *J. Diff. Geo.* **8** (1973) 85–94

[127] H. Weyl: Das asymptotische Verteilungsgesetz der Eigenschwingungen eines beliebig gestalten elastischen Körpers *Rendiconti del Circolo Mat. di Palermo* **39** (1915) 1–49

PETER B. GILKEY

Mathematics Department, University of Oregon, Eugene Or 97403 USA

http://www.uoregon.edu/~gilkey

E-mail: gilkey@uoregon.edu.

Handbook of Global Analysis
Demeter Krupka and David Saunders

Lagrangian formalism on Grassmann manifolds

D. R. Grigore

Contents

1 Introduction

The Lagrangian formalism is usually based, from the kinematical point of view, on a fiber bundle with the base manifold and the fibres interpreted as the "space-time" variables and the field variables respectively. As proved in the literature, the natural object associated to such a fiber bundle (from the Lagrangian formalism point of view) is the so-called variational exact sequence [24] - [26]. This sequence contains as distinguished terms the Lagrange, Euler-Lagrange and Helmholtz-Sonin forms, which are the main ingredients of the Lagrangian formalism. Using the exactness property one can obtain, beside a intrinsic geometrical formulation of the Lagrangian formalism, the most general expression of a variationally trivial Lagrangian and the generic form of a locally variational differential equation [13] - [14].

However, many interesting physical applications, such as the Minkowski space describing the relativistic particle, do not have a fiber bundle structure; one says that "there is no absolute time". One needs a generalization of this formalism to this more general case. For first-order Lagrangian systems a generalization of the Lagrangian formalism covering this case was proposed in [18] and for arbitrary order in particle mechanics in [12] (see also [11] for a review); the key notion is the so-called Lagrange-Souriau form which is a special type of Lepage form. For higher-order Lagrangian systems the general construction of the corresponding Grassmann manifold is more subtle and was performed in [15]. The idea is to start with the manifold of jets of immersions in the kinematical manifold of the problem. This manifold is usually called the velocity manifold and physically corresponds to parametrised evolutions. There exists a natural action of the so-called differential group on this manifold which physically corresponds to changing the parametrisation. One considers the submanifold of the regular velocities and takes its factorization to the differential group. This is exactly the Grassmann manifold associated to the kinematical manifold of the problem. The main combinatorial difficulty consists in establishing a convenient chart system on this factor manifold.

After the basic kinematical construction of Lagrangian systems is done one can proceed to the construction of the corresponding Lagrange, Euler-Lagrange and Helmholtz-Sonin form [16] and the Poincaré-Cartan form [17]. The expressions from the fibrating case are no longer well defined geometrical objects so one must find out proper substitutes for them. The idea is to construct these kind of objects first on the velocity manifold and impose some homogeneity properties. One discovers that these globally defined objects are inducing *locally* defined expressions on the Grassmann manifold which have convenient transformation properties with respect to a change of charts and formally coincide with the desired expressions of the usual Lagrangian formalism. In this way one is able to define on the Grassmann manifold the classes (modulo contact forms) of the Lagrange, Euler-Lagrange and Helmholtz-Sonin forms; the same is true for the Poincaré-Cartan form.

The paper is organized as follows. In Section 2 we remind the basic construction of a Grassmann manifold following essentially [15] but also providing some new results and we will also sketch the proof of the main results. In Section 3 we define the main objects of the Lagrangian formalism in the non-fibrating case. In Section 4 we give new proofs for the construction of the Lagrange-Souriau form in the case of first-order Lagrangian systems. In Section 5 we present some physical applications.

2 Grassmann manifolds

2.1 The manifold of (r,n)-velocities

Let us consider $N, n \geq 1$ and $r \geq 0$ integers such that $n \leq N$, and let X be a smooth manifold of dimension N describing the kinematical degrees of freedom of a certain physical problem.

We will consider $U \subset \mathbb{R}^n$ a neighborhood of the point $0 \in \mathbb{R}^n$, $x \in X$ and let $\Gamma_{(0,x)}$ be the set of smooth immersions $\gamma : U \to X$ such that $\gamma(0) = x$. As usual, we consider on $\Gamma_{(0,x)}$ the relation "$\gamma \sim \delta$" *iff* there exists a chart $(V,\psi)\quad \psi = (x^A), \quad A = 1,\dots,N$ on X such that the functions $\psi \circ \gamma, \psi \circ \delta : \mathbb{R}^n \to \mathbb{R}^N$ have the same partial derivatives up to order r in the point 0. The relation $\sim$ is a (chart independent) equivalence relation. By an (r,n)-*velocity* at a point $x \in X$ we mean such an equivalence class of the type $\Gamma_{(0,x)}/\sim$. The equivalence class of γ will be denoted by $j^r_0\gamma$. The set of (r,n)-velocities at x is denoted by $T^r_{(0,x)}(\mathbb{R}^n, Y) \equiv \Gamma_{(0,x)}/\sim$.

Further, we denote

$$T^r_n X = \bigcup_{x \in X} T^r_{(0,x)}(\mathbb{R}^n, X),$$

and define surjective mappings $\tau^{r,s}_n : T^r_n X \to T^s_n X$, where $0 < s \leq r$, by $\tau^{r,s}_n(j^r_0\gamma) = j^s_0\gamma$ and $\tau^{r,0}_n : T^r_n X \to X$, where $1 \leq r$, by $\tau^{r,0}_n(j^r_0\gamma) = \gamma(0)$.

In the conditions above let $(V,\psi), \quad \psi = (x^A)$, be a chart on X. Then we define the couple (V^r_n, ψ^r_n) where $V^r_n = (\pi^{r,0}_n)^{-1}(V), \quad \psi^r_n = (x^A, x^A_j, \cdots, x^A_{j_1,j_2,\dots,j_r})$, where $1 \leq j_1 \leq j_2 \leq \cdots \leq j_r \leq n$, and

$$x^A_{j_1,\dots,j_k}(j^r_0\gamma) \equiv \frac{\partial^k}{\partial t^{j_1}\dots\partial t^{j_k}} x^A \circ \gamma\Big|_0, \quad 0 \leq k \leq r. \tag{2.1.1}$$

Remark 2.1 Let us note that the expressions $x^A_{j_1,\cdots j_k}(j^r_0\gamma)$ are defined for all indices $j_1,\dots,j_r$ in the set $\{1,\dots,n\}$ but because of the symmetry property

$$x^A_{j_{P(1)},\dots,j_{P(k)}}(j^r_0\gamma) = x^A_{j_1,\dots,j_k}(j^r_0\gamma) \quad (k = 2,...,n) \tag{2.1.2}$$

for all permutations $P \in \mathcal{P}_k$ of the numbers $1,\dots,k$ we consider only the independent components given by the restrictions $1 \leq j_1 \leq j_2 \leq \cdots \leq j_r \leq n$. Taking this into account one can use multi-index notations i.e. $\psi^r_n = (x^A_J), \quad |J| = 0,...,r$ where by definition $x^A_\emptyset \equiv x^A$. The same comment is true for the partial derivatives $\frac{\partial}{\partial x^A_{j_1,\dots,j_k}}$.

The couple (V^r_n, ψ^r_n) is a chart on $T^r_n X$ called the *associated chart* of the chart (V,ψ) and the set $T^r_n X$ has a smooth structure defined by the system of charts (V^r_n, ψ^r_n); moreover $T^r_n X$ is a fiber bundle over X with the canonical projection $\tau^{r,0}$. The set $T^r_n Y$ endowed with the smooth structure defined by the associated charts defined above is called the *manifold of* (r,n)*-velocities* over X. The equations of the mapping $\tau^{r,s}_n : T^r_n X \to T^s_n X$ in terms of the associated charts are given by $x^A_{j_1,\dots,j_k} \circ \tau^{r,s}_n(j^r_0\gamma) = x^A_{j_1,\dots,j_k}(j^r_0\gamma)$, where $0 \leq k \leq s$. These mappings are all submersions.

2.2 Formal derivatives

Like in [2], [3], [13], we consider in the chart (V^r_n, ψ^r_n) the following differential operators:

$$\Delta^{j_1,\dots,j_k}_A \equiv \frac{r_1!\dots r_n!}{k!} \frac{\partial}{\partial x^A_{j_1,\dots,j_k}}, \qquad j_1,\dots,j_k \in \{1,\dots,n\} \tag{2.2.1}$$

where r_k is the number of times the index k shows up in the sequence $j_1,\dots j_k$.

The combinatorial coefficients are chosen in such a way that the following relation is true:

$$\Delta^{i_1,\dots,i_k}_A x^B_{j_1,\dots,j_l} = \begin{cases} \delta^B_A \mathcal{S}^+_{j_1,\dots,j_k} \delta^{i_1}_{j_1} \dots \delta^{i_k}_{j_k} & \text{if } k = l \\ 0 & \text{if } k \neq l. \end{cases} \tag{2.2.2}$$

Here we use the notations from [11], namely $\mathcal{S}^\pm_{j_1,\dots,j_k}$ are the symmetrization (for the sign $+$) and respectively the antisymmetrization (for the sign $-$) projector operators defined by

$$\mathcal{S}^\pm_{j_1,\dots,j_k} f_{j_1,\dots,j_k} \equiv \frac{1}{k!} \sum_{P \in \mathcal{P}_k} \epsilon_\pm(P) f_{j_{P(1)},\dots,j_{P(k)}} \tag{2.2.3}$$

where the sum runs over the permutation group $\mathcal{P}_k$ of the numbers $1,\dots,k$ and

$$\epsilon_+(P) \equiv 1, \quad \epsilon_-(P) \equiv (-1)^{|P|}, \quad \forall P \in \mathcal{P}_k;$$

here $|P|$ is the signature of the permutation P. In this way we avoid overcounting the indices. More precisely, for any smooth function on V^r, the following formula is true:

$$df = \sum_{k=0}^{r} (\Delta_A^{j_1,\dots,j_k} f) dx_{j_1,\dots,j_k}^A = \sum_{|I|\leq r} (\Delta_A^I f) dx_I^A \tag{2.2.4}$$

where we have also used the convenient multi-index notation.

We define now in the chart (V_n^r, ψ_n^r) the *formal derivatives* by the expressions

$$D_i^r \equiv \sum_{k=0}^{r-1} x_{i,j_1,\dots,j_k}^A \Delta_A^{j_1,\dots,j_k} = \sum_{|J|\leq r-1} x_{iJ}^A \Delta_A^J. \tag{2.2.5}$$

The last expression uses the multi-index notation; if I and J are two such multi-indices we mean by IJ the reunion of the two sets I, J. We note that the preceding formula does not define a vector field on $T_n^r X$. When no danger of confusion exists we simplify the notation putting simply $D_i = D_i^r$. One can easily verify that the following formulas follow directly from the definition:

$$D_i x_{j_1,\dots,j_k}^A = \begin{cases} x_{i,j_1,\dots,j_k}^A & \text{if } k \leq r-1 \\ 0 & \text{if } k = r, \end{cases} \tag{2.2.6}$$

$$\left[\Delta_A^{j_1,\dots,j_k}, D_i\right] = \mathcal{S}_{j_1,\dots,j_k}^+ \delta_i^{j_1} \Delta_A^{j_2,\dots,j_k}, \quad k = 0,\dots,r \tag{2.2.7}$$

and

$$[D_i, D_j] = 0. \tag{2.2.8}$$

The formal derivatives can be used to conveniently express the change of charts on the velocity manifold induced by a change of charts on X. Let (V, ψ) and $(\bar{V}, \bar{\psi})$ two charts on X such that $V \cap \bar{V} \neq \emptyset$ and let (V^r, ψ^r) and $(\bar{V}^r, \bar{\psi}^r)$ the corresponding attached charts from $T_n^r X$. The change of charts on X is $F : \mathbb{R}^N \to \mathbb{R}^N$ given by: $F \equiv \bar{\psi} \circ \psi^{-1}$. It is convenient to denote by $F^A : \mathbb{R}^N \to \mathbb{R}$ the components of F given by $F^A \equiv \bar{x}^A \circ \psi^{-1}$. We now consider the change of charts on $T_n^r X$ given by $F^r \equiv \bar{\psi}^r \circ (\psi^r)^{-1}$. One notes that $V^r \cap \bar{V}^r \neq \emptyset$; we need the explicit formulas for the components of F^r, namely for the functions

$$F_{j_1,\dots,j_k}^A \equiv \bar{x}_{j_1,\dots,j_k}^A \circ (\psi^r)^{-1}, \quad j_1 \leq j_2 \cdots \leq j_k, \quad k = 1,\dots,r$$

defined on the overlap: $V^r \cap \overline{V}^r$. First one notes the following relation:

$$\overline{D}_i = D_i. \tag{2.2.9}$$

Indeed, one defines for any immersion $\gamma \in \Gamma_{(0,x)}$ the map $j^r\gamma$ from $\mathbb{R}^n$ into $T_n^r X$ given by

$$x_{j_1,\dots,j_k}^A \circ j^r\gamma(t) \equiv \frac{\partial^k x^A \circ \gamma}{\partial t^{j_1} \dots \partial t^{j_k}}(t) \quad 0 \leq k \leq r \tag{2.2.10}$$

and easily discovers that

$$(j^r\gamma)_{*0}\frac{\partial}{\partial t^i} = D_i = \bar{D}_i. \tag{2.2.11}$$

Using (2.2.9) one easily finds out that the functions $F^A_{j_1,\dots,j_k}$ are given recursively by the following relation:

$$F^A_{jI} = D_j F^A_I \quad |I| \le r-1; \tag{2.2.12}$$

(compare with (2.2.6).) This relation can be "solved" explicitly according to

Lemma 2.2 *The following formula holds*

$$F^A_I = \sum_{p=1}^{|I|}\sum_{(I_1,\dots,I_p)} x^{B_1}_{I_1}\cdots x^{B_p}_{I_p}(\Delta_{B_1}\cdots\Delta_{B_p}F^A), \quad 1 \le |I| \le r \tag{2.2.13}$$

where the second sum denotes summation over all partitions $\mathcal{P}(I)$ of the set I and two partitions are considered identical if they differ only by a permutation of the subsets.

Proof We sketch the proof because the argument will be used repeatedly in this paper. It is natural to use complete induction on $|I|$. For $I = \{j\}$ the formula from the statement coincides with (2.2.12) for $I = \emptyset$. We suppose the formula true for any multi-index I with $|I| = s < r$ and prove it for the multi-index jI. If we use (2.2.12) we get:

$$\begin{aligned}
F^A_{jI} &= \sum_{p=1}^{|I|}\sum_{(I_1,\dots,I_p)}\left[\sum_{l=1}^{p} x^{B_1}_{I_1}\cdots(D_j x^{B_l}_{I_l})\cdots x^{B_p}_{I_p}(\Delta_{B_1}\cdots\Delta_{B_p}F^A)\right.\\
&\quad\left.+x^{B_1}_{I_1}\cdots x^{B_p}_{I_p}D_j(\Delta_{B_1}\cdots\Delta_{B_p}F^A)\right]\\
&= \sum_{p=1}^{|I|}\sum_{(I_1,\dots,I_p)}\left[\sum_{l=1}^{p} x^{B_1}_{I_1}\cdots x^{B_l}_{jI_l}\cdots x^{B_p}_{I_p}(\Delta_{B_1}\cdots\Delta_{B_p}F^A)\right.\\
&\quad\left.+x^{B_1}_{I_1}\cdots x^{B_p}_{I_p}x^{B_{p+1}}_{j}(\Delta_{B_1}\cdots\Delta_{B_{p+1}}F^A)\right].
\end{aligned}$$

We now note that the partitions $\mathcal{P}(jI)$ of the set jI can be obtained in two distinct ways:

- by taking a partition $(I_1,\dots,I_p) \in \mathcal{P}(I)$ and adjoining the index j to $I_1, I_2,\dots,I_p$;
- by taking a partition $(I_1,\dots,I_p) \in \mathcal{P}(I)$ and constructing the associated partition $(I_1,\dots,I_p,j) \in \mathcal{P}(jI)$.

We get the two types of contributions in the formula above and this finishes the proof. □

Remark 2.3 The combinatorial argument above will be called **the partition argument**.

Remark 2.4 From the formula derived above it immediately follows that we have:

$$\Delta^J_B F^A_I = 0, \quad 0 \le |I| < |J| \le r \tag{2.2.14}$$

i.e. the functions F^A_I depend only of the variables x^B_J with the restrictions specified above.

2.3 The differential group

By definition the *differential group of* order r is the set

$$L_n^r \equiv \{j_0^r\alpha \in J_{0,0}^r(\mathbb{R}^n, \mathbb{R}^n) | \alpha \in \mathit{Diff}(\mathbb{R}^n)\} \tag{2.3.1}$$

i.e. the group of invertible r-jets with source and target at $0 \in \mathbb{R}^n$. The group multiplication in L_n^r is defined by the jet composition $L_n^r \times L_n^r \ni (j_0^r\alpha, j_0^r\beta) \mapsto j_0^r(\alpha \circ \beta) \in L_n^r$.

The *canonical* (global) *coordinates* on L_n^r are defined by

$$a^i_{j_1,\dots,j_k}(j_0^r\alpha) = \left.\frac{\partial^k \alpha^i}{\partial t^{j_1} \dots \partial t^{j_k}}\right|_0, \quad j_1 \le j_2 \le \dots \le j_k, \quad k = 0, ..., r \tag{2.3.2}$$

where α^i are the components of a representative α of $j_0^r\alpha$.

We denote

$$a \equiv (a^i_j, a^i_{j_1,j_2}, \dots, a^i_{j_1,\dots,j_k}) = (a^i_J)_{1 \le |J| \le r}$$

and notice that one has

$$\det(a^i_j) \neq 0. \tag{2.3.3}$$

To obtain the composition law for the differential group we need a combinatorial result following easily by induction with the partition argument:

Lemma 2.5 *Let $U, V \in \mathbb{R}^n$ be open sets, $\alpha : U \to V$ and $f : V \to \mathbb{R}$ smooth functions. Then the following formula is true:*

$$\partial_I(f \circ \alpha) = \sum_{p=1}^{|I|} \sum_{(I_1,\dots,I_p)} (\partial_{I_1}\alpha^{i_1}) \dots (\partial_{I_p}\alpha^{i_p})(\partial_{i_1,\dots,i_p} f) \circ \alpha \tag{2.3.4}$$

where we have denoted for any multi-index $I = \{i_1, \dots, i_s\}$

$$\partial_I f \equiv \frac{\partial^s f}{\partial t^{i_1} \dots \partial t^{i_s}}.$$

We now have:

Lemma 2.6 *The group multiplication in L_n^r is expressed in the canonical coordinates by the equations*

$$(a \cdot b)^k_I = \sum_{p=1}^{|I|} \sum_{(I_1,\dots,I_p)} b^{j_1}_{I_1} \dots b^{j_p}_{I_p} a^k_{j_1,\dots j_p}, \quad |I| = 1, \dots r. \tag{2.3.5}$$

The group L_n^r is a Lie group.

Proof (i) We start from the defining formula:

$$(a \cdot b)^k_{j_1,\dots,j_l} = \left.\frac{\partial^l \alpha^k \circ \beta}{\partial t^{j_1} \dots \partial t^{j_l}}\right|_0$$

and apply the lemma above. One obtains the composition formula. It is clear that the composition formula (2.3.5) is a smooth function. The identity is evidently: $e \equiv (\delta^i_j, 0, \ldots, 0)$ and it remains to prove that the map $a \to a^{-1}$ is smooth; it follows immediately by induction that $(a^{-1})^k_I = \left(\det(a^i_j)\right)^{-|I|} \times P^k_I(a)$ where P^k_I is a polynomial in the variables $a^i_I, \quad |I| = 0, \ldots, r.$ □

The manifolds of (r,n)-velocities $T^r_n X$ admits a (natural) smooth right action of the differential group L^r_n, defined by the jet composition

$$(x \cdot a)^A_I \equiv x^A_I(j^r_0(\gamma \circ \alpha)) \tag{2.3.6}$$

where the connection between x^A_I and γ is given by (2.1.1) and the connection between a^i_I and α is given by (2.3.2). We determine the chart expression of this action.

Proposition 2.7 *The group action (2.3.6) is expressed by the equations*

$$(x \cdot a)^A = x^A, \quad (x \cdot a)^A_I = \sum_{p=1}^{|I|} \sum_{(I_1,\ldots,I_p) \in \mathcal{P}(\mathcal{I})} a^{j_1}_{I_1} \ldots a^{j_p}_{I_p} x^A_{j_1,\ldots,j_p}, \quad |I| \geq 1 \tag{2.3.7}$$

and it is smooth.

Proof One applies the definitions (2.1.1) and (2.3.2) together with the lemma 2.5. The smoothness is obvious from the explicit action formula given above. □

The group L^r_n has a natural smooth left action on the set of smooth real functions defined on $T^r_n X$, namely for any such function f we have:

$$(a \cdot f)(x) \equiv f(x \cdot a). \tag{2.3.8}$$

2.4 Higher order regular velocities

We say that a (r,n)-velocity $j^r_0\gamma \in T^r_n X$ is *regular*, if γ (or any other representative) is an immersion. If (V, ψ), $\psi = (x^A)$, is a chart, and the target $\gamma(0)$ of an element $j^r_0\gamma \in T^r_n X$ belongs to V, then $j^r_0\gamma$ is regular *iff* there exists a subsequence $\mathbf{I} \equiv (i_1, \ldots, i_n)$ of the sequence $(1, 2, \ldots, n, n+1, \ldots, n+m)$ such that

$$\det(x^{i_k}_j) \neq 0; \tag{2.4.1}$$

(here $x^{i_k}_j$ is a $n \times n$ real matrix.)

The associated charts have the form

$$(V^{\mathbf{I},r}, \psi^{\mathbf{I},r}), \quad \psi^{\mathbf{I},r} = (x^k_I, x^\sigma_I), \quad k = 1, \ldots, n, \quad \sigma = 1, \ldots m \equiv N - n, \quad |I| \leq r$$

where

$$x^k_I \equiv x^{i_k}_I, \quad k = 1, \ldots n$$

and $\sigma \in \{1, \ldots, N\} - \{i_1, \ldots, i_n\}$. The set of regular (r,n)-velocities is an open, L^r_n-invariant subset of $T^r_n X$, which is called the *manifold of regular* (r,n)*-velocities*, and is denoted by $\mathrm{Imm} T^r_n X$.

One can determine a complete system of L^r_n-invariants (in the sense of Weyl) of the action (2.3.7) on $\mathrm{Imm} T^r_n X$. We consider, for simplicity a chart for which one has $\{i_1, \ldots, i_n\} = \{1, \ldots, n\}$ and denote $x^\sigma_I \equiv x^{n+\sigma}_I, \quad \sigma = 1, \ldots m, \quad |I| \leq r$. First we have:

Proposition 2.8 *Let* (x^σ_I, x^i_I) *be the coordinates of a point in* $\mathrm{Imm}T^r_n X$. *Then*

$$\mathbf{x} \equiv (x^i_I)_{1\leq |I| \leq r} \tag{2.4.2}$$

is a element from L^r_n. *We denote its inverse by*

$$\mathbf{z} \equiv (z^i_I)_{1\leq |I| \leq r}. \tag{2.4.3}$$

Then z^i_j *is the inverse of the matrix* x^l_p:

$$z^i_j x^j_p = \delta^i_p \tag{2.4.4}$$

and the functions $z^i_{j_1,\dots,j_k}, k = 2,\dots r$ *can be determined recursively from the equations:*

$$z^i_{j_1,\dots,j_k} = z^p_{j_1} D_p z^i_{j_2,\dots,j_k}, \quad k = 2,\dots,r \tag{2.4.5}$$

Proof For the first assertion one uses (2.3.3) and (2.4.1). For the relation (2.4.5) one starts from the definition $\mathbf{z}\cdot\mathbf{x} = e$ or, in detail

$$\sum_{k=1}^{|I|} \sum_{(I_1,\dots,I_k)} x^{j_1}_{I_1} \dots x^{j_k}_{I_k} z^i_{j_1,\dots,j_k} = \begin{cases} \delta^i_I & \text{for } |I| = 1 \\ 0 & \text{for } |I| = 2,\dots r. \end{cases}$$

One performs two distinct operations on this relation: (a) we apply the operator D_p; (b) we make $I \mapsto Ip$. Next one subtracts the two results and uses the partition argument; the following relation follows:

$$\sum_{k=1}^{|I|} \sum_{(I_1,\dots,I_k)} x^{j_1}_{I_1} \dots x^{j_k}_{I_k} (D_p z^i_{j_1,\dots,j_k} - x^{j_0}_p z^i_{j_0,\dots,j_k}) = 0.$$

From this relation, we obtain, by induction the formula from the statement. □

The formula (2.4.5) suggests the following result:

Proposition 2.9 *Let* $(V,\psi), \psi = (x^A)$, *be a chart on* X *and let* (V^r_n, ψ^r_n) *be the associated chart on* $\mathrm{Imm}T^r_n X$. *We define recursively on this chart the following functions*

$$y^\sigma \equiv x^\sigma, \quad y^\sigma_{i_1,\dots,i_k} = z^j_{i_1} D_j y^\sigma_{i_2,\dots,i_k}, \quad k = 1,\dots,r; \tag{2.4.6}$$

(here z^j_i *are the first entries of the element* $\mathbf{z} \in L^r_n$.*) Then the functions* $y^\sigma_{i_1,\dots,i_k}$ *so defined depend smoothly only on* $x^A_J, \quad |J| \leq k$ *and are completely symmetric in all indices* $i_1,\dots,i_k, \quad k = 1,\dots r$.

Proof The first assertion follows immediately by induction. Next, one derives directly from the formula (2.4.6) that

$$y^\sigma_{i_1,\dots,i_k} = z^{j_1}_{i_1} z^{j_2}_{i_2} \left(D_{j_1} D_{j_2} y^\sigma_{i_3,\dots,i_k} - x^p_{i_1,i_2} z^j_p D_j y^\sigma_{i_3,\dots,i_k} \right), \quad k = 2,\dots r.$$

In particular we see that for $k = 2$ the symmetry property is true. One can proceed now by induction. If $y^\sigma_{i_1,\dots,i_{k-1}}$ is completely symmetric then the formula above shows that we have the symmetry property in the indices i_1 and i_2; moreover the recurrence relation (2.4.6) shows that we have the symmetry property in the indices $i_2,\dots,i_k$. So we obtain the desired property in all indices. □

As a result of the symmetry property just proved we can use the convenient multi-index notation $y^\sigma_I, \quad |I| \leq r$. Now we have an explicit formula for these functions.

Proposition 2.10 *The functions* $y_I^\sigma, \quad 1 \leq |I| \leq r$ *are* ***uniquely*** *determined by the recurrence relations:*

$$x_I^\sigma = \sum_{p=1}^{|I|} \sum_{(I_1,\ldots,I_p)} x_{I_1}^{j_1} \ldots x_{I_p}^{j_p} y_{j_1,\ldots,j_p}^\sigma. \tag{2.4.7}$$

Using the notation $\mathbf{x} \in L_n^r$ *one can compactly write the relation above as*

$$x_I^\sigma = (y \cdot \mathbf{x})_I^\sigma, \quad 1 \leq |I| \leq r. \tag{2.4.8}$$

Proof Goes by induction on $|I|$. The formula above is obvious for $I = \{j\}$. If it is valid for $|I| < r$ we apply to the relation above the operator d_j and use (2.2.6) and the partition argument. One obtains in this way the formula from the statement for Ij. The unicity also follows by induction. The last assertion is a consequence of the first formula and of the expression of the group action (2.3.7). □

Let us note that one can "invert" the formulas from the statement. Indeed, (2.4.8) is equivalent to

$$y_I^\sigma = (x \cdot \mathbf{z})_I^\sigma, \quad 1 \leq |I| \leq r \tag{2.4.9}$$

or explicitly

$$y_I^\sigma = \sum_{p=1}^{|I|} \sum_{(I_1,\ldots,I_p)} z_{I_1}^{j_1} \ldots z_{I_p}^{j_p} x_{j_1,\ldots,j_p}^\sigma. \tag{2.4.10}$$

Corollary 2.11 *One can use on* V^r *the new coordinates* $(y_I^\sigma, x_I^i), \quad |I| \leq r.$

Now we have the following result

Proposition 2.12 *The functions* $y_I^\sigma, \quad |I| \leq r$ *are* L_n^r*-invariants with respect to the natural action (2.3.8).*

Proof Let $a \in L_n^r$ be arbitrary. We start from (2.4.8) and use the associativity of the group composition law of L_n^r; we get:

$$(x \cdot a)_I^\sigma = ((y \cdot \mathbf{x}) \cdot a)_I^\sigma = (y \cdot (\mathbf{x} \cdot a))_I^\sigma.$$

On the other hand if we make in (2.4.8) the substitution $x \mapsto x \cdot a$ we get:

$$(x \cdot a)_I^\sigma = ((a \cdot y) \cdot (x \cdot a))_I^\sigma.$$

(here $a \cdot y$ denotes the action of the differential group on the functions y according to (2.3.8).)

By comparing the two formulas and using of the unicity statement from the preceding proposition we get the desired result. □

Moreover, we can prove:

Theorem 2.13 *The functions* $y_I^\sigma, \quad |I| \leq r$ *are a complete system of invariants in the sense of Weyl.*

Proof The fact that the functions $y^\sigma_I, \quad |I| \leq r$ are functionally independent follows by *reductio ad absurdum*. If they would be functionally dependent, then from (2.4.10) it would follow that the expressions $x^\sigma_I, \quad |I| \leq r$ also are functionally dependent. We still must show that there are no other invariants beside $y^\sigma_I, \quad |I| \leq r$. We proceed as follows. From corollary 2.11 it follows that one can use on V^r the coordinates $(y^\sigma_I, x^i_I), \quad |I| \leq r$. In these coordinates the action of the group L^r_n is:

$$(y \cdot a)^\sigma_I = y^\sigma_I, \quad |I| \leq r,$$

$$(x \cdot a)^i = x^i, \quad (x \cdot a)^i_I = \sum_{p=1}^{|I|} \sum_{(I_1,\dots,I_p)} a^{j_1}_{I_1} \dots a^{j_p}_{I_p} x^i_{j_1,\dots,j_p}, \quad 1 \leq |I| \leq r. \tag{2.4.11}$$

One can prove now by induction that this action is transitive. This shows that the system of invariants from the statement is complete. □

2.5 Higher order Grassmann bundles

The formalism presented above can be implemented in an arbitrary chart system $(V^{\mathbf{I},r}, \psi^{\mathbf{I},r})$ on $\mathrm{Imm}T^r_n X$ (see the beginning of the preceding subsection). In this context we have the central result [15]:

Theorem 2.14 *The set $P^r_n X \equiv \mathrm{Imm}T^r_n X / L^r_n$ has a unique differential manifold structure such that the canonical projection ρ^r_n is a submersion. The group action (2.3.7) defines on $\mathrm{Imm}T^r_n X$ the structure of a right principal L^r_n-bundle. A chart system on $P^r_n X$ adapted to this fiber bundle structure is formed from couples $(W^{\mathbf{I},r}, \Phi^{\mathbf{I},r})$ where:*

$$W^{\mathbf{I},r} = \left\{ j^r_0\gamma \in V^r | \det(x^{i_k}_j(j^r_0\gamma)) \neq 0 \right\} \tag{2.5.1}$$

and

$$\Phi^{\mathbf{I},r} = (x^i_I, y^\sigma_I), \quad |I| \leq r. \tag{2.5.2}$$

In this case the local expression of the canonical projection is

$$\rho^r_n(x^i_I, y^\sigma_I) = (x^i, y^\sigma_I).$$

Proof We define on $\mathrm{Imm}T^r_n X \times \mathrm{Imm}T^r_n X$ the equivalence relation

$$x \sim \bar{x} \quad \text{iff} \quad \exists a \in L^r_n \quad s.t. \quad \bar{x} = x \cdot a.$$

To prove the first assertion from the statement is sufficient (according to [7], par. 16.19.3) to prove that the graph of $\sim$ is a closed submanifold of the product manifold. We will look for a convenient system of coordinates on $\mathrm{Imm}T^r_n X$.

The first step is to take x and $\bar{x}$ such that $x \sim \bar{x}$ and to solve the system of equations

$$\bar{x} = x \cdot a_{x,\bar{x}}$$

for the unknown functions $a_{x,\bar{x}} \in L^r_n$.

One easily gets

$$(a_{x,\bar{x}})^i_j = z^i_p \bar{x}^p_j$$

and then shows by induction that $(a_{x,\bar{x}})^i_I$ are uniquely determined smooth functions of the variables $x^i_J, \bar{x}^i_J, \quad |J| \leq |I| \leq r$.

We now define the (local) smooth functions on $\mathrm{Imm}T^r_n X \times \mathrm{Imm}T^r_n X$:

$$\Phi^\sigma_I(x,\bar{x}) \equiv \bar{x}^\sigma_I - (x \cdot a_{x,\bar{x}})^\sigma_I, \quad |I| \leq r.$$

It is clear that on can take on $\mathrm{Imm}T^r_n X \times \mathrm{Imm}T^r_n X$ the (local) coordinates $(x^A_I, \Phi^\sigma_I, \bar{x}^i_I), \quad |I| \leq r$; it follows that the graph of $\sim$ is given by $\Phi^\sigma_I = 0, \quad |I| \leq r$ i.e. it is a closed submanifold. To prove the fiber bundle structure it is sufficient to show (see also [7]) that the action of L^r_n is free i.e. $x \cdot a = x \Longrightarrow a = e$. This fact follows elementary by induction. Finally, we have remarked before (see the preceding theorem) that one can take on $\mathrm{Imm}T^r_n X$ the coordinates $(y^\sigma_I, x^i_I), \quad |I| \leq r$ with the action given by (2.4.11). The last assertion about the expression of the canonical projection follows. □

A point of $P^r_n X$ containing a regular (r,n)-velocity $j^r_0\gamma$ is called an (r,n)-*contact element*, or an r-*contact element* of an n-dimensional submanifold of X, and is denoted by $[j^r_0\gamma]$. As in the case of r-jets, the point $0 \in \mathbb{R}^n$ (resp. $\gamma(0) \in X$) is called the *source* (resp. the *target*) of $[j^r_0\gamma]$. The manifold P^r_n is called the (r,n)-*Grassmannian bundle*, or simply a *higher order Grassmannian bundle* over X.

Besides the quotient projection $\rho^r_n : \mathrm{Imm}T^r_n X \to P^r_n$ we have for every $1 \leq s \leq r$, the *canonical projection* of $P^r_n X$ onto $P^s_n X$ defined by $\rho^{r,s}_n([j^r_0\gamma]) = [j^s_0\gamma]$ and the *canonical projection* of $P^r_n X$ onto X defined by $\rho^r_n([j^r_0\gamma]) = \gamma(0)$.

Remark 2.15 When the manifold X is fibred over a manifold M of dimension n one can also construct the jet extension $J^r X$ (see [24], [13]). One can establish a canonical isomorphism between $P^s_n X$ and $J^r X$ as follows: let $x \in M, x = \gamma(m), \gamma \in \Gamma_{(m,x)}$; and $\phi : \mathbb{R}^n \to M$ a (local) diffeomorphism such that $\phi(0) = m$. We can define $\tilde{\gamma} \in \Gamma_{(0,x)}$ by the formula $\tilde{\gamma} \equiv \gamma \circ \phi$. One notices that $\tilde{\gamma}_1 \sim \tilde{\gamma}_2$ *iff* there exists $\alpha \in \mathit{Diff}(\mathbb{R}^n)$ such that $\gamma_1 = \gamma_2 \circ \alpha$. This means that the map $j^r_n\gamma \mapsto j^r\tilde{\gamma}$ can be factorized to a map from $P^r_n X \to J^r X$ which is proved to be an isomorphism.

We also note the following result:

Proposition 2.16 *The following formula is true*

$$\Phi^\sigma_I = \sum_{p=1}^{|I|} \sum_{(I_1,\dots,I_p)} \bar{x}^{j_1}_{I_1} \dots \bar{x}^{j_p}_{I_p} (\bar{y}^\sigma_{j_1,\dots,j_p} - y^\sigma_{j_1,\dots,j_p}), \quad 1 \leq |I| \leq r. \tag{2.5.3}$$

or, in compact notations

$$\Phi^\sigma_I = ((\bar{y} - y) \cdot \bar{\mathbf{x}})^\sigma_I. \tag{2.5.4}$$

In particular, the equation

$$\Phi^\sigma_I = 0, \quad 1 \leq |I| \leq r$$

is equivalent to

$$\bar{y}^\sigma_I = y^\sigma_I, \quad |I| \leq r.$$

Proof The proof relies heavily on induction. Firstly, we define on $\mathrm{Imm}T^r_n X \times \mathrm{Imm}T^r_n X$ the expressions

$$V_i \equiv \bar{D}_i - (a_{x,\bar{x}})^i_j D_j$$

(where $a_{x,\bar{x}}$ have been defined previously) and we prove by induction the following formula:

$$V_i(a_{x,\bar{x}})^j_I = (a_{x,\bar{x}})^j_{iI}, \quad 1 \leq |I| \leq r-1.$$

Next, one uses this formula to prove by direct computation that

$$V_i \Phi^\sigma_I = \Phi^\sigma_{iI}.$$

Finally, one uses the preceding formula to prove by induction the formula (2.5.3) from the statement. □

Remark 2.17 The dimension of the factor manifold $P^r_n X$ is

$$\dim P^r_n X = m \binom{n+r}{n} + n.$$

Now we try to define on $P^r_n X$ the analogue of the total differential operators.

Proposition 2.18 *Let us consider on the regular velocities manifold* $\mathrm{Imm}T^r_n X$ *the coordinates* $(y^\sigma_I, x^i_I), \quad |I| \leq r$ *and define the operators*

$$\widetilde{\partial}^{j_1,\dots j_k}_\sigma \equiv \frac{r_1! \dots r_n!}{k!} \frac{\partial}{\partial y^\sigma_{j_1,\dots,j_k}} \tag{2.5.5}$$

(where we use the same conventions as in (2.2.1)).

We also define, by analogy to (2.2.5)

$$\tilde{d}^r_i \equiv \frac{\partial}{\partial x^i} + \sum_{k=0}^{r-1} y^\sigma_{i,j_1,\dots,j_k} \widetilde{\partial}^{j_1,\dots,j_k}_\sigma = \sum_{|J| \leq r-1} y^\sigma_{iJ} \widetilde{\partial}^J_\sigma. \tag{2.5.6}$$

Then the following formula is true

$$D_i = \sum_{p=1}^{r-1} x^l_{i,j_1,\dots,j_k} \Delta^{j_1,\dots,j_k}_l + x^p_i \tilde{d}_p \tag{2.5.7}$$

where, as usual, $\tilde{d}_i = \tilde{d}^r_i$ *when no danger of confusion arises.*

The proof goes by direct computation. Now we define on $P^r_n X$, in the chart $\rho^r_n(W^{\mathbf{I},r})$ some operators which are the analogues of (2.5.5) and (2.2.5), namely

$$\partial^{j_1,\dots j_k}_\sigma \equiv \frac{r_1! \dots r_n!}{k!} \frac{\partial}{\partial y^\sigma_{j_1,\dots,j_k}} \tag{2.5.8}$$

and

$$d_i \equiv \frac{\partial}{\partial x^i} + \sum_{k=0}^{r-1} y^\sigma_{i,j_1,\dots,j_k} \partial^{j_1,\dots,j_k}_\sigma = \frac{\partial}{\partial x^i} + \sum_{|J| \leq r-1} y^\sigma_{iJ} \partial^J_\sigma. \tag{2.5.9}$$

These operators are also called *total derivatives*. A formula similar to (2.2.2) is valid; moreover, if we use (2.2.9) the preceding proposition has the following consequence:

Proposition 2.19 *The following formula is true:*

$$(\rho^r_n)_*(z^j_i D_j) = d_i. \tag{2.5.10}$$

In particular, we have for any smooth function f on $\rho^r_n(W^r)$ the following formula:

$$D_i(f \circ \rho^r_n) = x^j_i (d_j f) \circ \rho^r_n. \tag{2.5.11}$$

Therefore, if (V, ψ) and $(\bar{V}, \bar{\psi})$ are two charts on X such that $V \cap \bar{V} \neq \emptyset$ and $d_i, \bar{d}_i, \quad i = 1, \dots n$ are the corresponding operators defined on $\rho^r_n(V^r)$ and respectively on $\rho^r_n(\bar{V}^r)$, then we have on $\rho^r_n(V^r \cap \bar{V}^r)$:

$$\mathrm{Span}(d_1, \dots, d_n) = \mathrm{Span}(\bar{d}_1, \dots, \bar{d}_n). \tag{2.5.12}$$

Finally we can give the formula for the chart change on $P^r_n X$.

Proposition 2.20 *In the conditions of the preceding proposition, let $(\rho^r_n(V^r), (x^i, y^\sigma))$ and respectively $(\rho^r_n(\bar{V}^r), (\bar{x}^i, \bar{y}^\sigma))$ be the two (overlapping charts); then the change of charts on $\rho^r_n(V^r) \cap \rho(\bar{V}^r)$ is given by:*

$$\bar{y}^\sigma_{iI} = P^j_i d_j \bar{y}^\sigma_I, \quad |I| \leq r - 1 \tag{2.5.13}$$

where P is the inverse of the matrix Q:

$$Q^l_p \equiv d_p \bar{x}^l, \quad P^j_i Q^l_j = \delta^l_i. \tag{2.5.14}$$

Proof We have from (2.4.6)

$$\bar{x}^i_j \bar{y}^\sigma_{iI} = \bar{D}_j \bar{y}^\sigma_I, \quad |I| \leq r - 1$$

with $\bar{y}^\sigma_I$ functions of $\bar{x}^A_J$. We will consider this relation on the overlap $V^r \cap \bar{V}^r$ such that $\bar{y}^\sigma_I$ can be considered as functions of x^A_J through the chart transformation formulas. Using also (2.2.9) one gets:

$$\bar{x}^i_j \bar{y}^\sigma_{iI} = D_j \bar{y}^\sigma_I. \quad |I| \leq r - 1$$

We rewrite this relation in the new coordinates (x^i, y^σ_I, x^i_I) (see corollary 2.11) and also use (2.5.7); as a result one finds out:

$$z^j_p \bar{x}^i_j \bar{y}^\sigma_{iI} = \tilde{d}_p \bar{y}^\sigma_I, \quad |I| \leq r - 1.$$

It remains to prove using also (2.5.7) that

$$Q^j_p = z^j_p \bar{x}^i_j \tag{2.5.15}$$

and the change transformation formula from the statement follows. □

We now note two other properties of the total differential operators d_i. The first one follows immediately from (2.5.10) and (2.5.15):

$$Q^j_i \bar{d}_j = d_i. \tag{2.5.16}$$

The second one is the analogue of (2.2.8):

$$[d_i, d_j] = 0. \tag{2.5.17}$$

So, for every multi-index I, the following expression makes sense:

$$d_I \equiv \prod_{i \in I} d_i. \tag{2.5.18}$$

We close this subsection with a result which will be useful later.

Proposition 2.21 *The following formula is true on the overlap of two charts:*

$$\partial_\sigma^{i_1,\dots,i_k} \bar{y}^\nu_{j_1,\dots,j_k} = \mathcal{S}^+_{j_1,\dots,j_k} P^{i_1}_{j_1} \dots P^{i_k}_{j_k} Q^\nu_\sigma, \quad k = 1, ..., r. \tag{2.5.19}$$

Here we have defined:

$$Q^\sigma_\nu \equiv \partial_\nu \bar{y}^\sigma - \bar{y}^\sigma_i (\partial_\nu \bar{x}^i). \tag{2.5.20}$$

Proof It is done by recurrence. First one proves directly from the definitions that:

$$\partial_\sigma^{i_1,\dots,i_k} \bar{y}^\nu_{j_1,\dots,j_k} = \mathcal{S}^+_{j_1,\dots,j_k} P^{i_1}_{j_1} \partial_\sigma^{i_2,\dots,i_k} \bar{y}^\nu_{j_2,\dots,j_k}, \quad k = 2, \dots, r \tag{2.5.21}$$

and then we obtain by recurrence:

$$\partial_\sigma^{i_1,\dots,i_k} \bar{y}^\nu_{j_1,\dots,j_k} = \mathcal{S}^+_{j_1,\dots,j_k} P^{i_1}_{j_1} \dots P^{i_{k-1}}_{j_{k-1}} \partial_\sigma^{i_k} \bar{y}^\nu_{j_k}, \quad k = 1, \dots, r. \tag{2.5.22}$$

Finally one establishes by direct computation that

$$\partial^i_\sigma \bar{y}^\nu_j = P^i_j Q^\nu_\sigma \tag{2.5.23}$$

and the formula from the statement follows. □

As a corollary we have the following fact:

Corollary 2.22 *Let us denote by* $\Omega^r_q(PX), \quad q \geq 0$ *the modulus of differential forms of order* q *on* P^r_n. *Then the subspace*

$$\Omega^r_{q,hor}(PX) \equiv \{\alpha \in \Omega^r_q(PX) | i_{\partial_\sigma^{j_1,\dots,j_r}} \alpha = 0\}$$

is globally well defined.

Proof One has, according to the chain rule on the overlap of two charts:

$$\partial_\sigma^{i_1,\dots,i_r} = \left(\partial_\sigma^{i_1,\dots,i_r} \bar{y}^\nu_{j_1,\dots,j_r}\right) \bar{\partial}_\nu^{i_1,\dots,i_r} = P^{i_1}_{j_1} \dots P^{i_r}_{j_r} Q^\nu_\sigma \bar{\partial}_\nu^{i_1,\dots,i_r}$$

and a similar formula for the corresponding inner contractions. It follows that the relation

$$i_{\partial_\sigma^{j_1,\dots,j_r}} \alpha = 0$$

is chart independent. □

2.6 Contact forms on Grassmann manifolds

In this subsection we give some new material about the possibility of defining the contact forms on the factor manifold $P^r_n X$. Fortunately, most of the definitions and properties from [24]-[26] and [13] can be adapted to this more general situation.

By a *contact form* on $P^r_n X$ we mean any form $\rho \in \Omega^r_q(PX)$ verifying

$$[j^r\gamma]^* \rho = 0 \tag{2.6.1}$$

for any immersion $\gamma : \mathbb{R}^n \to X$. We denote by $\Omega^r_{q(c)}(PX)$ the set of contact forms of degree $q \leq n$. Here $[j^r\gamma] : \mathbb{R}^n \to P^r_n$ is given by (see def. (2.2.10)) $[j^r\gamma](t) \equiv [j^r_t\gamma]$. Now, many results from [13] are practically unchanged. We mention some of them.

If one considers only the contact forms on an open set $\rho^r_n(V^r) \subset P^r_n X$ then we emphasize this by writing $\Omega^r_{q(c)}(V)$. One immediately notes that $\Omega^r_{0(c)} = 0$ and that for $q > n$ any q-form is contact. It is also elementary to see that the set of all contact forms is an ideal, denoted by $\mathcal{C}(\Omega^r)$, with respect to the operation $\wedge$. Because the operations of pull-back and of differentiation are commuting this ideal is left invariant by exterior differentiation:

$$d\mathcal{C}(\Omega^r) \subset \mathcal{C}(\Omega^r). \tag{2.6.2}$$

By elementary computations one finds out that, as in the case of a fiber bundle, for any chart (V, ψ) on X, every element of the set $\Omega^r_{1(c)}(V)$ is a linear combination of the following expressions:

$$\omega^\sigma_{j_1,\dots,j_k} \equiv dy^\sigma_{j_1,\dots,j_k} - y^\sigma_{i,j_1,\dots,j_k} dx^i, \quad k = 0, \dots, r-1 \tag{2.6.3}$$

or, in multi-index notations

$$\omega^\sigma_J \equiv dy^\sigma_J - y^\sigma_{iJ} dx^i, \quad |J| \leq r-1. \tag{2.6.4}$$

From the definition above it is clear that the linear subspace of the 1-forms on $P^r_n X$ is generated by $dx^i, \quad \omega^\sigma_J, \quad (|J| \leq r-1)$ and $dy^\sigma_I, |I| = r$. For any smooth function on $\rho(V^r)$ we have

$$df = (d_i f) dx^i + \sum_{|J| \leq r-1} (\partial^J_\sigma f) \omega^\sigma_J + \sum_{|I|=r} (\partial^I_\nu f) dy^\nu_I. \tag{2.6.5}$$

We also have the formula

$$d\omega^\sigma_J = -\omega^\sigma_{Ji} \wedge dx^i, \quad |J| \leq r-2. \tag{2.6.6}$$

The structure theorem from [26], [13] stays true, i.e. any $\rho \in \Omega^r_q(PX), \quad q = 2, \dots, n$ is contact *iff* it has the following expression in the associated chart:

$$\rho = \sum_{|J| \leq r-1} \omega^\sigma_J \wedge \Phi^J_\sigma + \sum_{|I|=r-1} d\omega^\sigma_I \wedge \Psi^I_\sigma \tag{2.6.7}$$

where $\Phi^J_\sigma \in \Omega^r_{q-1}$ and $\Psi^I_\sigma \in \Omega^r_{q-2}$ can be arbitrary forms. (We adopt the convention that $\Omega^r_q \equiv 0, \forall q < 0$).

We will need in the following the transformation formula relevant for change of charts. It is to be expected that there will be some modifications of the corresponding formula from the fiber bundle case. Namely, we have by elementary computations

Proposition 2.23 *Let (V, ψ) and $(\bar{V}, \bar{\psi})$ two overlapping charts on X and let (W^r, Φ^r), $\Phi^r = (x^i, y^\sigma_I, x^i_I)$ and $(\bar{W}^r, \bar{\Phi}^r)$, $\bar{\Phi}^r = (\bar{x}^i, \bar{y}^\sigma_I, \bar{x}^i_I)$ the corresponding charts on $T^r_n X$. Then the following formulas are true on $\rho^r_n(W^r \cap \bar{W}^r)$:*

$$\bar{\omega}^\sigma_I = \sum_{|J|=1}^{|I|} (\partial^J_\nu \bar{y}^\sigma_I)\omega^\nu_J - Q^\sigma_{I,\nu}\omega^\nu, \qquad 1 \le |I| \le r-1. \tag{2.6.8}$$

and

$$\bar{\omega}^\sigma = Q^\sigma_\nu \omega^\nu \tag{2.6.9}$$

where we have defined:

$$Q^\sigma_{I,\nu} \equiv \partial_\nu \bar{y}^\sigma_I - \bar{y}^\sigma_{jI}(\partial_\nu \bar{x}^j), \qquad 0 \le |I| \le r-1. \tag{2.6.10}$$

Remark 2.24 Let us note that the notations are consistent in the sense that $Q^\sigma_\nu = Q^\sigma_{\emptyset,\nu}$ where Q^σ_ν is given by the formula (2.5.20)

As a consequence of the preceding proposition we have:

Corollary 2.25 *If a q-form has the expression*

$$\rho = \sum_{p+s=q-n+1} \sum_{|J_1|,\ldots,|J_p| \le r-1} \sum_{|I_1|,\ldots,|I_s|=r-1} \omega^{\sigma_1}_{J_1} \cdots \wedge \omega^{\sigma_p}_{J_p} \wedge \\ \wedge d\omega^{\nu_1}_{I_1} \cdots \wedge d\omega^{\nu_s}_{I_s} \wedge \Phi^{J_1,\ldots,J_p,I_1,\ldots,I_s}_{\sigma_1,\ldots,\sigma_p,\nu_1,\ldots,\nu_s} \tag{2.6.11}$$

valid in one chart, then it is valid in any other chart.

This corollary allows us to define for any $q = n+1, \ldots, dim(J^r Y) = m\binom{n+r}{n}$ a *strongly contact form* to be any $\rho \in \Omega^r_q$ such that it has in one chart (thereafter in any other chart) the expression above. For a certain uniformity of notations, we denote these forms by $\Omega^r_{q(c)}$. Now it follows that one can define the variational sequence and prove its exactness as in the fiber bundle case. We also mention the fact that one can define a global operator on the linear space $\Omega^r_{q,hor}X$ defined at the end of the preceding subsection, at least in the case $r = 2$. In fact we have [18]

Proposition 2.26 *Let $r = 2$. Then, the operator locally defined on any differential form by:*

$$K\alpha \equiv i_{d_j} i_{\partial^j_\sigma} (\omega^\sigma \wedge \alpha) \tag{2.6.12}$$

is globally defined on the subspace $\Omega^2_{q,hor}X$.

Proof One works on the overlap of two charts and starts from the definition above trying to transform everything into the other set of coordinates. It is quite elementary to use corollary 2.22 to find

$$K\alpha \equiv i_{\bar{d}_j} i_{\bar{\partial}^j_\nu} (Q^\nu_\sigma \omega^\sigma \wedge \alpha)$$

Now one uses (2.5.19) and the transformation formula (2.6.9) for the contact forms to obtain

$$K\alpha = \bar{K}\bar{\alpha}.$$

that it, K is well defined globally. □

2.7 Morphisms of Grassmannian manifolds

Let $X_i, \quad i = 1, 2$ be two differential manifolds and $\phi : X_1 \to X_2$ a smooth map. We define the new map $j^r\phi : T^r_n X_1 \to T^r_n X_2$ according to

$$j^r\phi(j^r_0\gamma) \equiv j^r_0\phi \circ \gamma \tag{2.7.1}$$

for any immersion γ. If γ is a regular immersion, then one can see that the map $j^r\phi$ maps $\mathrm{Imm}T^r_n(X_1)$ into $\mathrm{Imm}T^r_n(X_2)$ and so, it factorizes to a map $J^r\phi : P^r_n X_1 \to P^r_n X_2$ given by

$$J^r\phi([j^r_0\gamma]) \equiv [j^r_0\phi \circ \gamma]. \tag{2.7.2}$$

The map $J^r\phi$ is called the *extension of order r of the map* ϕ. One can show that the contact ideal behaves naturally with respect to prolongations i.e.

$$(J^r\phi)^*\mathcal{C}(\Omega^r(PX_2)) \subset \mathcal{C}(\Omega^r(PX_1)). \tag{2.7.3}$$

The proof follows directly from the definition of a contact form. If ξ is a vector field on X we define its *extension of order r* on $T^r_n X$ and on $P^r_n X$ the vector fields $j^r\xi$ and $J^r\xi$ respectively given by the following formulas:

$$j^r\xi_{j^r_0\gamma}f \equiv \frac{d}{dt} f \circ j^r e^{t\xi}(j^r_0\gamma)\Big|_{t=0} \tag{2.7.4}$$

(for any smooth real function f on $T^r_n X$) and

$$J^r\xi \equiv \frac{d}{dt} J^r e^{t\xi}\Big|_{t=0}; \tag{2.7.5}$$

here $e^{t\xi}$ is, as usual, the flow associated to ξ. One will need the explicit formula of $j^r\xi$. If in the chart (V, ψ) we have

$$\xi = \xi^A \Delta_A \tag{2.7.6}$$

with ξ^A smooth function, then $j^r\xi$ has the following expression in the associated chart (V^r, ψ^r):

$$j^r\xi = \sum_{|I|\leq r} (D_J\xi^A)\Delta^J_A. \tag{2.7.7}$$

The proof of this fact follows by direct computation from the definition above. We call *evolutions* these type of vector fields on $T^r_n X$ and denote the set of evolutions by $\mathcal{E}(T^r_n X)$. As a consequence of (2.7.3), if ξ is a vector field on X, then

$$L_{J^r\xi}\mathcal{C}(\Omega^r(X)) \subset \mathcal{C}(\Omega^r(X)). \tag{2.7.8}$$

Now, as in [13] we have the following results. Suppose that in local coordinates we have

$$\xi = a^i(x, y)\frac{\partial}{\partial x^i} + b^\sigma(x, y)\partial_\sigma \tag{2.7.9}$$

with a^i and b^σ smooth function; then $J^r\xi$ must have the following expression in the associated chart:

$$J^r\xi = a^i \frac{\partial}{\partial x^i} + \sum_{|J|\le r} b^\sigma_J \partial^J_\sigma. \tag{2.7.10}$$

where

$$b^\sigma_J = d_I(b^\sigma - y^\sigma_j a^j) + y^\sigma_{jI} a^j, \quad |I| \le r-1, \qquad b^\sigma_I = d_I(b^\sigma - y^\sigma_j a^j), \quad |I| = r. \tag{2.7.11}$$

Finally we give the expression of the prolongation $J^r\phi$ where ϕ is a bundle morphism of the X. If ϕ has the local expression

$$\phi(x^i, y^\sigma) = (f^i, F^\sigma) \tag{2.7.12}$$

then we must have in the associated chart:

$$J^r\phi(x^i, y^\sigma, y^\sigma_j, ..., y^\sigma_{j_1,...,j_r}) = (f^i, F^\sigma, F^\sigma_j, ..., F^\sigma_{j_1,...,j_r}) \tag{2.7.13}$$

where $F^\sigma_{j_1,...,j_k}, \quad j_1 \le j_2 \le \cdots \le j_k, \quad k = 1, ..., r$ are smooth local functions given recursively by:

$$F^\sigma_{Ji} = P^l_i d_l F^\sigma_J \quad |J| \le r-1; \tag{2.7.14}$$

we also have

$$\partial^I_\nu F^\sigma_J = 0 \quad |I| = r. \tag{2.7.15}$$

3 The Lagrangian formalism on a Grassmann manifold

3.1 Euler-Lagrange forms

We outline a construction from [2] which is the main combinatorial trick in the study of globalisation of the Lagrangian formalism. We call any map $P : \mathcal{E}(T^r_n X) \to \Omega^s(T^s_n X)$, $s \ge r$ covering the identity map: $id : T^r_n X \to T^r_n X$ a *total differential operator*. In local coordinates such an operator has the following expression: if ξ has the local expression (2.7.6), then:

$$P(\xi) = \sum_{|I|\le r} (D_I \xi^A) P^I_A = \sum_{k=0}^{r} \left(D_{j_1} \ldots D_{j_k} \xi^A \right) P^{j_1,...,j_k}_A \tag{3.1.1}$$

where $P^{j_1,...,j_k}_A$ are differential forms on $T^s_n X$ and $D_j = D^s_j$.

Then, as in [2] and [13] one has the following combinatorial lemma:

Lemma 3.1 *In the conditions above, the following formula is true:*

$$P(\xi) = \sum_{|I|\le r} D_I(\xi^A Q^I_A) \tag{3.1.2}$$

where

$$Q_A^I \equiv \sum_{|J| \leq r-|I|} (-1)^{|J|} \binom{|I|+|J|}{|I|} D_J P_A^{IJ} \tag{3.1.3}$$

and one assumes that the action of a formal derivative D_j on a form is realized by through the Lie derivatives. Moreover, the relation (3.1.2) ***uniquely*** *determines the forms Q_A^I.*

The proof is identical with the one presented in [13]. We also have

Proposition 3.2 *In the conditions above one has on the overlap $V^s \cap \bar{V}^s$ the following formula:*

$$Q_A = (\Delta_A \bar{x}^B) \bar{Q}_B. \tag{3.1.4}$$

In particular, there exists a globally defined form, denoted by $E(P)(\xi)$ with the local expression

$$E(P)(\xi) = Q_A \xi^A. \tag{3.1.5}$$

Proof From the formula (3.1.2) we have

$$P(\xi) = \xi^A Q_A + \sum_{k=1}^{r} D_{j_1} \ldots D_{j_k} \left(\xi^A Q_A^{j_1,\ldots,j_k} \right)$$

So, in the overlap $V^s \cap \bar{V}^s$ we have

$$\xi^A Q_A - \bar{\xi}^A \bar{Q}_A = \sum_{k=1}^{r} \left[\bar{D}_{j_1} \ldots \bar{D}_{j_k} \left(\bar{\xi}^A \bar{Q}_A^{j_1,\ldots,j_k} \right) - D_{j_1} \ldots D_{j_k} \left(\xi^A Q_A^{j_1,\ldots,j_k} \right) \right].$$

But because of the relation (2.2.9) we can simplify considerably this formula, namely we get:

$$\xi^A Q_A - \bar{\xi}^A \bar{Q}_A = \sum_{k=1}^{r} D_{j_1} \ldots D_{j_k} \left(\bar{\xi}^A \bar{Q}_A^{j_1,\ldots,j_k} - \xi^A Q_A^{j_1,\ldots,j_k} \right)$$

Now one proves that both sides are zero as in [2], [13] making use of Stokes theorem. □

The operator $E(P)$ defined by (3.1.5) is called the *Euler operator* associated to the total differential operator P; it has the local expression:

$$E(P)(\xi) = \xi^A E_A(P) \tag{3.1.6}$$

where

$$E_A(P) = \sum_{|I|=0}^{r} (-1)^{|I|} D_I P_A^I. \tag{3.1.7}$$

Now one takes $\mathcal{L} \in \mathcal{F}(T_n^r)$ and constructs the total differential operator $P_{\mathcal{L}}$ according to:

$$P_{\mathcal{L}}(\xi) \equiv L_{pr(\xi)} \mathcal{L}. \tag{3.1.8}$$

Lemma 3.1 can be applied and immediately gives the following local formula:

$$P_{\mathcal{L}}(\xi) = \sum_{|I|=0}^{r} D_I \left(\xi^A \mathcal{E}_A^I(\mathcal{L})\right) \tag{3.1.9}$$

where

$$\mathcal{E}_A^I(L) \equiv \sum_{|J|\leq r-|I|} (-1)^{|J|} \binom{|I|+|J|}{|I|} D_J \Delta_A^{IJ} \mathcal{L} \tag{3.1.10}$$

are the so-called *Lie-Euler operators*; the Euler operator associated to $P_{\mathcal{L}}$ has the following expression:

$$E(P_{\mathcal{L}}) = \xi^A \mathcal{E}_A(L) \tag{3.1.11}$$

where

$$\mathcal{E}_A(L) \equiv \sum_{|J|\leq r} (-1)^{|J|} D_J \Delta_A^J \mathcal{L} \tag{3.1.12}$$

are the *Euler-Lagrange expressions* associated to $\mathcal{L}$. The proposition above leads to

Proposition 3.3 *If $\mathcal{L} \in \mathcal{F}(T_n^r)$, then there exists a globally defined 1-form, denoted by $\mathcal{E}(\mathcal{L})$ such that we have in the chart $V^s, \quad s \geq 2r$:*

$$\mathcal{E}(\mathcal{L}) = \mathcal{E}_A(\mathcal{L}) dx^A. \tag{3.1.13}$$

Proof By construction $Q_A = \mathcal{E}_A(\mathcal{L})$. Now, one has from (3.1.4) :

$$\mathcal{E}_A(\mathcal{L}) = (\Delta_A \bar{x}^B) \bar{\mathcal{E}}_B(\bar{\mathcal{L}}). \tag{3.1.14}$$

Combining with the transformation property $d\bar{x}^A = (\Delta_B \bar{x}^A) dx^B$ and obtains that the formula (3.1.13) has a global meaning. □

One calls this form the *Euler-Lagrange form* associated to $\mathcal{L}$. All the properties of this form listed in [13] are true in this case also. We insist only on the so-called *product rule* for the Euler-Lagrange expressions [1] which will be repeatedly used in the following.

Proposition 3.4 *If f and g are smooth functions on V^r then one has in $V^s, \quad s \geq 2r$ the following formula:*

$$\mathcal{E}_A^I(fg) = \sum_{|J|\leq r-|I|} \binom{|I|+|J|}{|J|} \left[(D_J f)\mathcal{E}_A^{IJ}(g) + (D_J g)\mathcal{E}_A^{IJ}(f)\right], \quad |I| \leq r. \tag{3.1.15}$$

The proof goes by direct computation, directly from the definition of the Lie-Euler operators combined with Leibnitz rule of differentiation of a product. We now come to the main definition. A smooth real function $\mathcal{L}$ on $\mathrm{Imm}T_n^r$ is called a *homogeneous Lagrangian* if it verifies the relation:

$$\mathcal{L}(x \cdot a) = \det(a)\mathcal{L}(x), \quad \forall a \in L_n^r; \tag{3.1.16}$$

here by $\det(a)$ we mean $\det(a_j^i)$. Such an object induces on the factor manifold $P_n^r X$ an non-homogeneous object.

Proposition 3.5 *Let $\mathcal{L}$ be a homogeneous Lagrangian. Then for every chart (W^r, Φ^r) on $\mathrm{Imm}T^r_n X$ there exists a smooth real function on $\rho^r_n(W^r)$ such that:*

$$\mathcal{L} = \det(\mathbf{x}) L \circ \rho^r_n \tag{3.1.17}$$

and conversely, if $\mathcal{L}$ is locally defined by this relation, then it verifies (3.1.16).

Proof One chooses in (3.1.16) $a = \mathbf{x}$ (see (2.4.2)). □

The function L is called the *non-homogeneous (local) Lagrangian* associated to $\mathcal{L}$. As a consequence of the connection (3.1.17) we have

Proposition 3.6 *Let $\mathcal{L}$ a homogeneous Lagrangian and (V, ψ), $(\bar{V}, \bar{\psi})$ two overlapping charts on X. We consider on the associated charts (W^r, Φ^r) and $(\bar{W}^r, \bar{\Phi}^r)$ the corresponding non-homogeneous Lagrangians L and respectively $\bar{L}$. Then we have on the overlap $\rho^r_n(W^r \cap \bar{W}^r)$ the following formula:*

$$L = \mathcal{J} \bar{L} \tag{3.1.18}$$

where

$$\mathcal{J} \equiv \det(Q) = \det(d_i \bar{x}^j). \tag{3.1.19}$$

Proof One writes (3.1.17) for both charts and gets

$$\mathcal{L} = \det(\mathbf{x}) L \circ \rho^r_n = \det(\bar{\mathbf{x}}) \bar{L} \circ \rho^r_n;$$

as a consequence

$$L \circ \rho^r_n = \det(\bar{\mathbf{x}}\mathbf{z}) \bar{L} \circ \rho^r_n.$$

One now uses the relation (2.5.15) and obtains the relation from the statement. □

It follows that we have

Theorem 3.7 *One can globally define the equivalence class*

$$[\lambda] \in \Omega^r_n(PX)/\Omega^r_{n(c)}(PX)$$

such that the local expression of λ is

$$\lambda = L\theta_0; \tag{3.1.20}$$

here, as usual

$$\theta_0 \equiv dx^1 \wedge \cdots \wedge dx^n. \tag{3.1.21}$$

Proof One proves immediately that on the overlap of the associated charts from the proposition above one has:

$$\bar{\theta}_0 = \mathcal{J}\theta_0 + \text{contact terms}. \tag{3.1.22}$$

This result must be combined with (3.1.18) to obtain on $\rho_n^r(W^r \cap \bar{W}^r)$:

$$\lambda - \bar{\lambda} \in \Omega^r_{n(c)}(PX);$$

this proves the theorem. □

It is natural to ask what is the connection between the Euler-Lagrange expression of the homogeneous and the corresponding non-homogeneous Lagrangian. The answer is contained in:

Theorem 3.8 *Suppose $\mathcal{L}$ is a homogeneous Lagrangian defined on $\mathrm{Imm}T^r_n X$ and L is the associated non-homogeneous Lagrangian. Then the following relations are valid on the chart $W^s, \quad s \geq 2r$:*

$$\mathcal{E}_\sigma(\mathcal{L}) = \det(\mathbf{x}) E_\sigma(L) \circ \rho_n^s, \tag{3.1.23}$$

$$\mathcal{E}_\sigma^{j_1,\dots,j_k}(\mathcal{L}) = (-1)^k \det(\mathbf{x}) \sum_{|I| \geq k} (-1)^{|I|} \sum_{(I_1,\dots,I_k)} \mathcal{S}^+_{j_1,\dots,j_k} z^{j_1}_{I_1} \dots z^{j_k}_{I_k} E^I_\sigma(T) \circ \rho_n^s, \tag{3.1.24}$$

and

$$\mathcal{E}_p(\mathcal{L}) = -\det(\mathbf{x}) y^\sigma_p E_\sigma(T) \circ \rho_n^s, \tag{3.1.25}$$

$$\mathcal{E}_p^j(\mathcal{T}) = \det(\mathbf{x}) \left[z^j_p L + y^\sigma_p \sum_{I \leq r} z^j_I E^I_\sigma(L) \right] \circ \rho_n^s, \tag{3.1.26}$$

$$\mathcal{E}_p^{j_1,\dots,j_k}(\mathcal{L}) = (-1)^{k+1} \det(\mathbf{x}) y^\sigma_p \sum_{|I| \geq k} (-1)^{|I|} \sum_{(I_1,\dots,I_k)} \mathcal{S}^+_{j_1,\dots,j_k} z^{j_1}_{I_1} \dots z^{j_k}_{I_k} E^I_\sigma(L) \circ \rho_n^s. \tag{3.1.27}$$

Proof (i) As a general strategy of the proof, we will try to transform the expression of the total differential operator $P_\mathcal{L}(\xi)$ in terms of L; we have by definition

$$P_\mathcal{L}(\xi) = \sum_{|I| \leq r} (D_I \xi^A) \Delta^I_A \left[\det(\mathbf{x}) L \circ \rho \right].$$

Using the chain rule one obtains rather easily from here:

$$\begin{aligned} P_\mathcal{L}(\xi) = \det(\mathbf{x}) \sum_{k=0}^r (D_{j_1} \dots D_{j_k} \xi^\sigma) \sum_{l=k}^r (\partial^{i_1,\dots,i_l}_\nu L) \circ \rho \quad (\Delta^{j_1,\dots,j_k}_\sigma y^\nu_{i_1,\dots,i_l}) + \\ \det(\mathbf{x}) \sum_{k=0}^r (D_{j_1} \dots D_{j_k} \xi^p) \sum_{l=k}^r (\partial^{i_1,\dots,i_l}_\nu L) \circ \rho \quad (\Delta^{j_1,\dots,j_k}_p y^\nu_{i_1,\dots,i_l}) + \\ \det(\mathbf{x}) \xi^p (\partial_p L) \circ \rho + (D_j \xi^p) [\Delta^j_p \det(\mathbf{x})] L \circ \rho \end{aligned} \tag{3.1.28}$$

We will now consider that the functions ξ^A depend only of the variables (x^i, y^σ) i.e. there exist the smooth real functions Ξ^A on W^r such that

$$\xi^A = \Xi^A \circ \rho. \tag{3.1.29}$$

(ii) To compute further the expression above one starts from (2.4.10) and firstly proves directly:

$$\Delta_\sigma^{j_1,\dots,j_k} y_I^\nu = \delta_\sigma^\nu \mathcal{S}^+_{j_1,\dots,j_k} \sum_{(I_1,\dots,I_k)} z_{I_1}^{j_1} \cdots z_{I_k}^{j_k}, \quad 1 \leq k \leq |I| \leq r. \tag{3.1.30}$$

Next one proves

Lemma 3.9 *If we have $1 \leq k \leq |I| \leq r$ then one has:*

$$\Delta_p^{j_1,\dots,j_k} y_I^\sigma = -\mathcal{S}^+_{j_1,\dots,j_k} \sum_{(I_0,\dots,I_k)} z_{I_1}^{j_1} \cdots z_{I_k}^{j_k} y_{pI_0}^\sigma. \tag{3.1.31}$$

Here we understand that any of the subsets $I_1, \dots, I_k$ cannot be the emptyset; on the contrary, it is allowed to have $I_0 = \emptyset$.

Proof By induction on $|I|$ starting from $|I| = k$. For this smallest possible value one uses (2.4.10). Then one supposes that the formula from the statement is valid for $k \leq |I| < r$, uses the defining recurrence relation (2.4.6) for the invariants and establishes the relation for iI. The cases $k = 1$ and $k > 1$ must be treated separately. □

Another auxiliary result is contained in the well known result:

Lemma 3.10 *The following formulas are true:*

$$\Delta_j^i \det(\mathbf{x}) = z_j^i \det(\mathbf{x}) \tag{3.1.32}$$

and

$$D_i[z_j^i \det(\mathbf{x})] = 0. \tag{3.1.33}$$

Proof The first result is general i.e. valid for any invertible matrix and it is proved directly from the definition of the determinant. The second formula is a corollary of the first. □

One must use these results together with (2.5.11); we have from (3.1.28) after permuting the two summation signs:

$$\begin{aligned} P_{\mathcal{L}}(\xi) = \det(\mathbf{x})[&\xi^\sigma (\partial_\sigma L) \circ \rho \\ &+ \sum_{|I|=1}^{r} (\partial_\sigma^I L) \circ \rho \quad \sum_{k=1}^{I} \sum_{(I_1,\dots,I_k)} z_{I_1}^{j_1} \cdots z_{I_k}^{j_k} (D_{j_1} \dots D_{j_k} \xi^\sigma) \\ &- \sum_{|I_0| \leq r} \sum_{|I|=1}^{r-|I_0|} \binom{|I| + |I_0|}{|I|} y_{pI_0}^\sigma (\partial_\sigma^{II_0} L) \circ \rho \\ &\qquad \sum_{k=1}^{I} \sum_{(I_1,\dots,I_k)} z_{I_1}^{j_1} \cdots z_{I_k}^{j_k} (D_{j_1} \dots D_{j_k} \xi^p) \\ &+ \xi^p (\partial_p L) \circ \rho - \xi^p (d_j L) \circ \rho] \\ &+ D_j[\xi^p z_p^j \det(\mathbf{x}) L \circ \rho]. \end{aligned} \tag{3.1.34}$$

(iii) One can proceed further one must generalise the formula (2.5.11). We have the following formula

$$D_I(f\circ\rho) = \sum_{k=1}^{I} \sum_{(I_1,\dots,I_k)} x^{j_1}_{I_1}\cdots x^{j_k}_{I_k}(d_{j_1}\dots d_{j_k}f)\circ\rho \tag{3.1.35}$$

which can be proved by induction on $|I|$. This formula can be "inverted" rather easily and we get:

$$(d_I f)\circ\rho = \sum_{k=1}^{I} \sum_{(I_1,\dots,I_k)} z^{j_1}_{I_1}\cdots z^{j_k}_{I_k} D_{j_1}\dots D_{j_k}(f\circ\rho). \tag{3.1.36}$$

The expression (3.1.34) can be considerably simplified to

$$P_{\mathcal{L}}(\xi) = \det(\mathbf{x})[\sum_{|I|\leq 1}(d_I\Xi^\sigma)P^I_\sigma - (d_I\Xi^p)P^I_p]\circ\rho + D_j[\xi^p z^j_p\det(\mathbf{x})L\circ\rho] \tag{3.1.37}$$

where we have defined

$$P^I_\sigma \equiv \partial^I_\sigma L; \quad P^I_p \equiv \sum_{|J|\leq r-|I|}\binom{|I|+|J|}{|J|} y^\sigma_{pJ}\partial^{IJ}_\sigma L.$$

(iv) Now it this the time to apply lemma 3.1 and to obtain in this way:

$$P_{\mathcal{L}}(\xi) = \det(\mathbf{x})\sum_{|I|\leq r}[d_I(\Xi^\sigma Q^I_\sigma - \Xi^p Q^I_p)]\circ\rho + D_j[\xi^p z^j_p\det(\mathbf{x})L\circ\rho] \tag{3.1.38}$$

where one gets by some computations the following formulas for the expressions Q^I_A:

$$Q^I_\sigma = \sum_{|J|\leq r-|I|}(-1)^{|J|}\binom{|I|+|J|}{|J|} d_J P^{IJ}_\sigma = E^I_\sigma(L) \tag{3.1.39}$$

and

$$Q^I_p = \sum_{|J|\leq r-|I|}(-1)^{|J|}\binom{|I|+|J|}{|J|} d_J P^{IJ}_p = y^\sigma_p E^I_\sigma(L). \tag{3.1.40}$$

(v) We want to compare the expression (3.1.38) with the right hand side of the formula from lemma 3.1. To do this we need one more combinatorial result valid for any smooth real function on the chart W^r:

$$\det(\mathbf{x})(d_I f)\circ\rho = \sum_{k=1}^{I}(-1)^{k-|I|}D_{j_1}\cdots D_{j_k}\left[\det(\mathbf{x})\sum_{(I_1,\dots,I_k)} z^{j_1}_{I_1}\cdots z^{j_k}_{I_k}(f\circ\rho)\right], \quad |I|\leq r. \tag{3.1.41}$$

One proves this formula by induction on $|I|$ and so the final expression for the total differential operator is

$$P_{\mathcal{L}}(\xi) = \det(\mathbf{x})(\Xi^\sigma Q_\sigma - \Xi^p Q_p^I) \circ \rho + D_j[\xi^p z_p^j \det(\mathbf{x}) L \circ \rho] + \sum_{k=1}^{r} (-1)^k D_{j_1} \cdots D_{j_k} \left[\det(\mathbf{x}) \sum_{|I| \geq k} (-1)^{|I|} \sum_{(I_1,\ldots,I_k)} z_{I_1}^{j_1} \cdots z_{I_k}^{j_k} (\Xi^\sigma Q_\sigma^I - \Xi^p Q_p^I) \circ \rho \right] \tag{3.1.42}$$

If one uses the unicity statement from lemma 3.1 one obtains the desired formulas. □

Immediate consequences of the preceding theorem are

Corollary 3.11 *If $\mathcal{L}$ is a homogeneous Lagrangian, then the following relations are true on the manifold $T_n^s X, \quad s \geq 2r$:*

$$\mathcal{E}_A(\mathcal{L})(x \cdot a) = \det(a) \mathcal{E}_A(\mathcal{L})(x), \quad \forall a \in L_n^r. \tag{3.1.43}$$

In the conditions of the above theorem we have:

$$\mathcal{E}_A(\mathcal{L}) \equiv 0 \Longleftrightarrow E_\sigma(L) \equiv 0. \tag{3.1.44}$$

We also have the analogue of proposition 3.6:

Proposition 3.12 *In the condition of the preceding theorem let us consider two overlapping charts $(V, \psi), \quad (\bar{V}, \bar{\psi})$. Then one has on the overlap of the associated charts: $\rho_n^s(W^s \cap \bar{W}^s)$ the following relation:*

$$E_\sigma(L) = \mathcal{J} Q_\sigma^\nu \bar{E}_\nu(\bar{L}) \tag{3.1.45}$$

where the matrix Q_σ^ν has been defined by the formula (2.5.20).

Proof We start from the transformation formula (3.1.14) for $A \to \sigma$:

$$\mathcal{E}_\sigma(\mathcal{L}) = (\partial_\sigma \bar{x}^i) \bar{\mathcal{E}}_i(\bar{\mathcal{L}}) + (\partial_\sigma \bar{y}^\nu) \bar{\mathcal{E}}_\nu(\bar{\mathcal{L}})$$

and substitute (3.1.23) and (3.1.25). Using (2.5.15) we obtain by elementary computations the relation from the statement. □

Remark 3.13 For a different proof of this result see [28].

Now we have the analogue of theorem 3.7:

Theorem 3.14 *If $\mathcal{L}$ is a homogeneous Lagrangian on T_n^r, then one can globally define the equivalence class*

$$[E(L)] \in \Omega_{n+1}^s(PX)/\Omega_{n+1(c)}^s(PX)$$

on $P_n^s, \quad s \geq 2r$ such that the local expression of $E(L)$ is

$$E(L) = E_\sigma(L) \omega^\sigma \wedge \theta_0. \tag{3.1.46}$$

Proof Follows the lines of theorem 3.7 and it is elementary. □

We also note the following property:

Proposition 3.15 *If $\mathcal{L}$ is a homogeneous Lagrangian on $T^r_n X$, then the corresponding Euler-Lagrange form verifies on $T^s_n X, \quad s \geq 2r$ the following identity:*

$$(j^s_0 \gamma)^* \mathcal{E}(\mathcal{L}) = 0, \quad \forall \gamma \in \mathrm{Imm} T^s_n. \tag{3.1.47}$$

Proof By direct computation we get

$$(j^s_0 \gamma)^* \mathcal{E}(\mathcal{L}) = \left(\mathcal{E}_A(\mathcal{L}) x^A_j \right) \circ j^s_0 \gamma \quad dt^j.$$

Now one uses (3.1.23) and (3.1.25) to prove that the expression in the bracket is identically zero. □

We close this subsection with some remarks.

Remark 3.16 If $\mathcal{L}$ is a homogeneous Lagrangian, one can expect some homogeneity property for the total differential operator associated to it. Indeed, one has for an arbitrary Lagrangian

$$P_{\mathcal{L}} \circ \phi_a(\xi) = P_{\mathcal{L}}(\xi) \circ \phi_a \tag{3.1.48}$$

where ϕ_a denotes the right action of the differential group L^r_n. As a consequence, one has for a homogeneous Lagrangian

$$P_{\mathcal{L}}(\xi) \circ \phi_a = \det(a) P_{\mathcal{L}}(\xi), \quad \forall a \in L^r_n. \tag{3.1.49}$$

3.2 Differential equations on Grassmann manifolds

An element $\mathcal{T} \in \Omega^s_{n+1}(X)$ is called a *differential equation* on $T^s_n X$. In the chart (V^s, ψ^s) the differential equation $\mathcal{T}$ has the following local expression:

$$\mathcal{T} = \mathcal{T}_A dx^A. \tag{3.2.1}$$

It is clear that the Euler-Lagrange form defined by (3.1.13) is a differential equation. A differential equation $\mathcal{T}$ is called *variational* if there exists a Lagrangian $\mathcal{L}$ on $T^r_n X, \quad s \geq 2r$ such that we have $\mathcal{T} = \mathcal{E}(\mathcal{L})$. If the function $\mathcal{L}$ is only locally defined, then such a differential equation is called *locally variational* (or, of the *Euler-Lagrange type*). If $\gamma : \mathbb{R}^n \to X$ is a immersion, then on says that the differential equation $\mathcal{T}$ *verifies the differential equation iff* we have

$$(j^s_0 \gamma)^* i_Z \mathcal{T} = 0 \tag{3.2.2}$$

for any vector field Z on $J^s_n X$. In local coordinates we have on V^s:

$$\mathcal{T}_A \circ j^s_0 \gamma = 0 \quad (A = 1, ..., N). \tag{3.2.3}$$

Guided by corollary 3.11 and prop 3.15 we also introduce the following definition. We say that $\mathcal{T}$ is a *homogeneous differential equation* on $\mathrm{Imm} T^s_n X$ if it verifies the following conditions:

$$(\phi_a)^* \mathcal{T} = \det(a) \mathcal{T}, \quad \forall a \in L^s_n \tag{3.2.4}$$

and

$$(j_0^s\gamma)^*\mathcal{T} = 0, \quad \forall\gamma \in \mathrm{Imm}T_n^s. \tag{3.2.5}$$

Then we have the following result which can be proved by elementary computations suggested by the similar results obtained for a differential equation of the Euler-Lagrange type.

Theorem 3.17 *Let $\mathcal{T}$ be a homogeneous differential equation on $\mathrm{Imm}T_n^sX$. Then there exist some local smooth real functions T_σ in every chart $\rho_n^s(W^s)$ such that one has:*

$$\mathcal{T}_\sigma = \det(\mathbf{x})T_\sigma \circ \rho, \qquad \mathcal{T}_i = -\det(\mathbf{x})y_i^\sigma T_\sigma \circ \rho. \tag{3.2.6}$$

As a consequence, if (V,ψ), $(\bar{V},\bar{\psi})$ are two overlapping charts on X, then one has on the intersection $\rho_n^s(W^s \cap \bar{W}^s)$ the following transformation formula:

$$T_\sigma = \mathcal{J}P_\sigma^\nu\bar{T}_\nu \tag{3.2.7}$$

and the class

$$[T] \in \Omega_{n+1}^s(PX)/\Omega_{n+1(c)}^s(PX)$$

can be properly defined such that we have locally

$$T = T_\sigma\omega^\sigma \wedge \theta_0. \tag{3.2.8}$$

One calls T the *associated (local) non-homogenous differential equation*. We now prove the existence of the (globally) defined Helmholtz-Sonin form associated to a differential equation. By analogy with [13] we have the following result

Theorem 3.18 *Let $\mathcal{T}$ be a differential equation on T_n^sX with the local form given by (3.2.1). We define the following expressions in any chart V^t, $t > 2s$:*

$$\mathcal{H}_{AB}^J \equiv \Delta_B^J\mathcal{T}_A - (-1)^{|J|}\mathcal{E}_A^J(\mathcal{T}_B), \quad |J| \leq s. \tag{3.2.9}$$

Then there exists a globally defined 2-form, denoted by $\mathcal{H}(\mathcal{T})$ such that in any chart V^t we have:

$$\mathcal{H}(\mathcal{T}) = \sum_{|J|\leq s}\mathcal{H}_{AB}^J dx_J^B \wedge dx^A. \tag{3.2.10}$$

Proof We sketch briefly the argument from [13]. Let ξ be a vector field on X; we define a (global) 1-form $\mathcal{H}_\xi(\mathcal{T})$ according to:

$$\mathcal{H}_\xi(\mathcal{T}) \equiv L_{pr(\xi)}\mathcal{T} - \mathcal{E}\left(i_{pr(\xi)}\mathcal{T}\right) \tag{3.2.11}$$

and the following local expression is obtained:

$$\mathcal{H}_\xi(\mathcal{T}) = \sum_{|I|\leq s}(D_I\xi^B)\mathcal{H}_{AB}^I dx^A. \tag{3.2.12}$$

The transformation formula for a change of charts for the expressions $d_I\xi^B$ is:

$$\bar{D}_I\bar{\xi}^A = \sum_{|J|\leq|I|} \left(\Delta_B^J\bar{x}_I^A\right)\left(D_J\xi^B\right), \quad |I| = 0, ..., s. \tag{3.2.13}$$

Using the transformation formula (3.2.13) one can obtain the transformation formula for the expressions $\mathcal{H}_{AB}^I$: one has in the overlap $V^t \cap \bar{V}^t, \quad t \geq 2s$:

$$\mathcal{H}_{DB}^J = \sum_{|I|\geq|J|} \left(\Delta_B^J\bar{x}_I^A\right)\left(\Delta_D\bar{x}^C\right)\bar{\mathcal{H}}_{CA}^I. \tag{3.2.14}$$

This transformation formula leads now to the fact that $\mathcal{H}(\mathcal{T})$ has an invariant meaning. □

$\mathcal{H}(\mathcal{T})$ is called the *Helmholtz-Sonin form* associated to $\mathcal{T}$ and $\mathcal{H}_{AB}^I$ are the *Helmholtz-Sonin expressions* associated to $\mathcal{T}$. A well-known corollary of the theorem above is:

Corollary 3.19 *The differential equation $\mathcal{T}$ is locally variational iff $\mathcal{H}(\mathcal{T}) = 0$ iff*

$$\mathcal{H}_{AB}^I = 0, \quad \forall A, B = 1, \ldots N, \quad \forall |I| \leq r. \tag{3.2.15}$$

The proof is identical with the one presented in [13]. The preceding equations are called the *Helmholtz-Sonin equations*. As in the preceding subsection, if $\mathcal{T}$ is a homogeneous differential equation, we have a very precise connection between the Helmholtz-Sonin expressions of $\mathcal{T}$ and of T from theorem 3.17.

Theorem 3.20 *Suppose $\mathcal{T}$ is a homogeneous equation defined on $\mathrm{Imm}T_n^sX$ and T_σ the components of the associated non-homogeneous equation. Then the following relations are valid on the chart $W^t, \quad t \geq 2s$:*

$$\mathcal{H}_{\sigma\nu}(\mathcal{T}) = \det(\mathbf{x})H_{\sigma\nu}(T) \circ \rho_n^t, \tag{3.2.16}$$

$$\mathcal{H}_{\sigma\nu}^{j_1,\ldots,j_k}(\mathcal{T}) = \det(\mathbf{x}) \sum_{|I|\geq k} \sum_{(I_1,\ldots,I_k)} \mathcal{S}_{j_1,\ldots,j_k}^+ z_{I_1}^{j_1} \ldots z_{I_k}^{j_k} H_{\sigma\nu}^I(T) \circ \rho_n^t, \quad k = 1, \ldots, s. \tag{3.2.17}$$

$$\mathcal{H}_{\sigma p}(\mathcal{T}) = -\det(\mathbf{x}) \sum_{I_0\leq s} y_{pI_0}^\nu H_{\sigma\nu}^{I_0}(T) \circ \rho_n^t, \tag{3.2.18}$$

$$\mathcal{H}_{\sigma p}^{j_1,\ldots,j_k}(\mathcal{T}) = -\det(\mathbf{x}) \sum_{I_0\leq s} \sum_{|I|=k}^{s-|I_0|} \binom{|I|+|I_0|}{|I|} \sum_{(I_1,\ldots,I_k)} \mathcal{S}_{j_1,\ldots,j_k}^+ z_{I_1}^{j_1} \ldots z_{I_k}^{j_k} y_{pI_0}^\nu H_{\sigma\nu}^{II_0}(T) \circ \rho_n^t, \quad k = 1, \ldots, s, \tag{3.2.19}$$

$$\mathcal{H}_{p\nu}(\mathcal{T}) = -\det(\mathbf{x}) y_p^\sigma H_{\sigma\nu}(T) \circ \rho_n^t, \tag{3.2.20}$$

$$\mathcal{H}_{p\nu}^{j_1,\ldots,j_k}(\mathcal{T}) = -\det(\mathbf{x}) y_p^\sigma \sum_{I_0\leq s} \sum_{|I|\geq k} \sum_{(I_1,\ldots,I_k)} \mathcal{S}_{j_1,\ldots,j_k}^+ z_{I_1}^{j_1} \ldots z_{I_k}^{j_k} H_{\sigma\nu}^I(T) \circ \rho_n^t, \quad k = 1, \ldots, s \tag{3.2.21}$$

and

$$\mathcal{H}_{pq}(\mathcal{T}) = \det(\mathbf{x}) y^\sigma_p \sum_{I_0 \leq s} y^\nu_{qI_0} H^{I_0}_{\sigma\nu}(T) \circ \rho^t_n, \tag{3.2.22}$$

$$\mathcal{H}^{j_1,\dots,j_k}_{pq}(\mathcal{T}) = \det(\mathbf{x}) y^\sigma_p \sum_{I_0 \leq s} \sum_{|I|=k}^{s-|I_0|} \binom{|I|+|I_0|}{|I|} \sum_{(I_1,\dots,I_k)} \mathcal{S}^+_{j_1,\dots,j_k} z^{j_1}_{I_1} \dots z^{j_k}_{I_k} y^\nu_{qI_0} H^{II_0}_{\sigma\nu}(T) \circ \rho^t_n, \quad k = 1,\dots,s. \tag{3.2.23}$$

The proof is tedious but elementary. One must use the formulas derived in theorem 3.8 combined with the derivation property from proposition 3.4 to prove case by case the formulas from the statement. Occasionally, one must study separately the cases $k = 0, 1$ and $k > 1$. In the conditions of the above theorem we have:

Corollary 3.21

$$\mathcal{H}^I_{AB}(\mathcal{T}) \equiv 0 \Longleftrightarrow H_{\sigma\nu}(T) \equiv 0. \tag{3.2.24}$$

As it can be expected we have the analogues of propositions 3.6 and 3.12:

Proposition 3.22 *In the condition of the preceding theorem let us consider two overlapping charts* $(V, \psi), \quad (\bar{V}, \bar{\psi})$*. Then one has on the overlap of the associated charts:* $\rho^t_n(W^t \cap \bar{W}^t)$ *the following relations:*

$$H_{\sigma\nu}(T) = \mathcal{J} P^\alpha_\sigma \sum_{|J| \leq s} \left[\left(\partial_\nu \bar{y}^\beta_J\right) - P^i_j \left(d_i \bar{y}^\beta_J\right) \left(\partial_\nu \bar{x}^j\right) \right] \bar{H}^J_{\alpha\beta}(\bar{T}) \tag{3.2.25}$$

$$H^I_{\sigma\nu}(T) = \mathcal{J} P^\alpha_\sigma \sum_{|J| \geq |I|} \left(\partial_\nu \bar{y}^\beta_J\right) \bar{H}^J_{\alpha\beta}(\bar{T}), \quad \forall I \neq \emptyset. \tag{3.2.26}$$

Proof (i) It is convenient to introduce the expression $g_{j_1,\dots,j_k}, \quad k = 0,\dots,r$ completely symmetric in all indices (with the convention $g_\emptyset = g$) and to use the (3.2.16) and (3.2.17) to obtain:

$$\det(\mathbf{x}) \sum_{|I| \leq s} (g \cdot \mathbf{z})_I H^I_{\sigma\nu}(T) \circ \rho = \sum_{k=0}^{s} g_{j_1,\dots,j_k} \mathcal{H}^{j_1,\dots,j_k}_{\sigma\nu}(\mathcal{T}) = \sum_{|I| \leq s} g_I \sum_{|J| \geq |I|} \left(\Delta^J_\sigma \bar{x}^A_I\right) \left(\Delta_\nu \bar{x}^C\right) \bar{\mathcal{H}}^I_{CA}(\bar{\mathcal{T}})$$

where use have been made of the transformation formula (3.2.14) for $B \to \sigma, \quad D \to \nu$.

If we make here $g \to g \cdot \mathbf{x}$ we obtain after elementary prelucrations:

$$\det(\mathbf{x}) H_{\sigma\nu}(T) \circ \rho = \sum_{|I| \leq s} \sum_{|J| \geq |I|} [(\Delta_\nu \bar{x}^\beta_I)(\Delta_\sigma \bar{x}^\alpha) \bar{\mathcal{H}}^I_{\alpha\beta}(\bar{\mathcal{T}}) + (\Delta_\nu \bar{x}^\beta_I)(\Delta_\sigma \bar{x}^p) \bar{\mathcal{H}}^I_{p\beta}(\bar{\mathcal{T}}) + (\Delta_\nu \bar{x}^q_I)(\Delta_\sigma \bar{x}^\alpha) \bar{\mathcal{H}}^I_{\alpha q}(\bar{\mathcal{T}}) + (\Delta_\nu \bar{x}^q_I)(\Delta_\sigma \bar{x}^p) \bar{\mathcal{H}}^I_{pq}(\bar{\mathcal{T}})]$$

(3.2.27)

and for $k \geq 1$:

$$\det(\mathbf{x}) H^{j_1,\dots,j_k}_{\sigma\nu}(T) \circ \rho = \sum_{|J| \geq k} \sum_{(J_1,\dots,J_k)} \mathcal{S}^{+}_{j_1,\dots,j_k} x^{j_1}_{J_1} \dots x^{j_k}_{J_k} \sum_{|J| \geq |I|} [(\Delta^J_\nu \bar{x}^\beta_I)(\Delta_\sigma \bar{x}^\alpha) \bar{\mathcal{H}}^I_{\alpha\beta}(\bar{\mathcal{T}}) + (\Delta^J_\nu \bar{x}^\beta_I)(\Delta_\sigma \bar{x}^p) \bar{\mathcal{H}}^I_{p\beta}(\bar{\mathcal{T}}) + (\Delta^J_\nu \bar{x}^q_I)(\Delta_\sigma \bar{x}^\alpha) \bar{\mathcal{H}}^I_{\alpha q}(\bar{\mathcal{T}}) + (\Delta^J_\nu \bar{x}^q_I)(\Delta_\sigma \bar{x}^p) \bar{\mathcal{H}}^I_{pq}(\bar{\mathcal{T}})] \tag{3.2.28}$$

The second relation can be considerably simplified if one uses (2.4.7), more precisely

$$\partial^{j_1,\dots,j_k}_\nu x^\mu_I = \delta^\mu_\nu \sum_{(I_1,\dots,I_k)} \mathcal{S}^{+}_{j_1,\dots,j_k} x^{j_1}_{I_1} \dots x^{j_k}_{I_k}; \tag{3.2.29}$$

the chain rule gives from here

$$\partial^{j_1,\dots,j_k}_\nu \bar{x}^A_I = \sum_{|J|=k}^{|I|} (\Delta^J_\nu \bar{x}^A_I) \sum_{(J_1,\dots,J_k)} \mathcal{S}^{+}_{j_1,\dots,j_k} x^{j_1}_{J_1} \dots x^{j_k}_{J_k}. \tag{3.2.30}$$

So, the relation (3.2.28) becomes:

$$\det(\mathbf{x}) H^{j_1,\dots,j_k}_{\sigma\nu}(T) \circ \rho = \sum_{|I| \geq k} \{(\partial^{j_1,\dots,j_k}_\nu \bar{x}^\beta_I)[(\Delta_\sigma \bar{x}^\alpha) \bar{\mathcal{H}}^I_{\alpha\beta}(\bar{\mathcal{T}}) + (\Delta_\sigma \bar{x}^p_I) \bar{\mathcal{H}}^I_{p\beta}(\bar{\mathcal{T}})] + (\partial^{j_1,\dots,j_k}_\nu \bar{x}^q_I)[(\Delta_\sigma \bar{x}^\alpha) \bar{\mathcal{H}}^I_{\alpha q}(\bar{\mathcal{T}}) + (\Delta_\sigma \bar{x}^p) \bar{\mathcal{H}}^I_{pq}(\bar{\mathcal{T}})]\} \tag{3.2.31}$$

If we compare with (3.2.27) we see that the preceding relation stays true for $k = 0$ also. Now we use again the theorem above in the right hand side of the relation just derived and obtains after some computations (using the relations (3.1.30) and (3.1.31) and the chain rule) the relations from the statement of the theorem. □

Now we have the analogue of theorems 3.7 and 3.14

Theorem 3.23 *If $\mathcal{T}$ is a homogeneous differential equation on* $\mathrm{Imm}T^s_n$*, then one can globally define the equivalence class*

$$[H(T)] \in \Omega^t_{n+2}(PX)/\Omega^t_{n+2(c)}(PX)$$

on $P^t_n, \quad t \geq 2s$ such that the local expression of $H(T)$ is

$$H(T) = \sum_{|I| \leq s} H^I_{\sigma\nu}(L) \omega^\nu_I \wedge \omega^\sigma \wedge \theta_0. \tag{3.2.32}$$

Proof Follows the lines of theorem 3.14 and it is elementary. □

As a consequence of the theorems 3.7, 3.14 and 3.23 we can apply the exactness of the variational sequence and obtain that the expressions of an arbitrary variationally trivial Lagrangian and of a locally variational differential equation are the same as those from [14]. For a related analysis see [29] and references quoted there.

4 Lagrangian formalism on second order Grassmann bundles

4.1 The second order Grassmann bundle

Here we particularize the results obtained in the preceding sections for the case $r = 2$. The coordinates on $T^2_n X$ are (x^A, x^A_j, x^A_{ij}) and with the help of the derivative operators (see (2.2.1))

$$\Delta_A \equiv \frac{\partial}{\partial x^A}, \quad \Delta^j_A \equiv \frac{\partial}{\partial x^A_j}, \quad \Delta^{ij}_A \equiv \begin{cases} \frac{\partial}{\partial x^A_{ij}}, & \text{for } i = j \\ \frac{1}{2}\frac{\partial}{\partial x^A_{ij}}, & \text{for } i \neq j \end{cases} \tag{4.1.1}$$

we have for any smooth function f (see (2.2.4)):

$$df = (\partial^A f)dx^A + (\Delta^A_i f)dx^A_i + (\Delta^A_{ij} f)dx^A_{ij} \tag{4.1.2}$$

We have the following formulas (see (2.2.2)):

$$\Delta_A x^B = \delta^B_A, \quad \Delta^i_A x^B_j = \delta^B_A \delta^i_j, \quad \Delta^{ij}_A x^B_{lm} = \frac{1}{2}\delta^B_A(\delta^i_l \delta^j_m + \delta^i_m \delta^j_l) \tag{4.1.3}$$

and the other derivatives are zero. The formal derivatives (see (2.2.5)) are in this case:

$$D^r_i \equiv x^A \Delta_A + x^A_{ij}\Delta^j_A \tag{4.1.4}$$

and from here we immediately have (see (2.2.6)):

$$D_i x^A = x^A_i, \quad D_i x^A_j = x^A_{ij}. \tag{4.1.5}$$

The formulas for the induces change of charts (see (2.2.13)) are in this case:

$$F^A_i = x^B_i \Delta_B F^A, \quad F^A_{i_1,i_2} = x^B_{i_1,i_2}\Delta_B F^A + x^{B_1}_{i_1} x^{B_2}_{i_2} \Delta_{B_1}\Delta_{B_2} F^A \tag{4.1.6}$$

The elements of the differential group are of the form

$$a = (a^j_i, a^j_{i_1,i_2}), \quad \det(a^j_i) \neq 0 \tag{4.1.7}$$

with the composition law (see (2.3.5)):

$$(a \cdot b)^k_i = b^k_j a^j_i, \quad (a \cdot b)^k_{i_1,i_2} = b^j_{i_1,i_2} a^k_j + b^{j_1}_{i_1} b^{j_2}_{i_2} a^k_{j_1,j_2} \tag{4.1.8}$$

and the inverse element given by

$$a^{-1} = ((a^{-1})^j_i, -(a^{-1})^j_k (a^{-1})^{j_1}_{i_1} (a^{-1})^{j_2}_{i_2} a^k_{j_1,j_2}). \tag{4.1.9}$$

The action of this group on $T^2_n X$ is (see (2.3.7)):

$$(a \cdot x)^A = x^A, \quad (a \cdot x)^A_i = a^j_i x^A_j, \quad (a \cdot x)^A_{i_1,i_2} = a^j_{i_1,i_2} x^A_j + a^{j_1}_{i_1} a^{j_2}_{i_2} x^A_{j_1,j_2}. \tag{4.1.10}$$

The expressions for the invariants of this action are (see (2.4.10)):

$$y^\sigma = x^\sigma, \quad y^\sigma_i = z^j_i x^\sigma_j, \quad y^\sigma_{i_1,i_2} = z^{j_1}_{i_1} z^{j_2}_{i_2}(x^\sigma_{j_1,j_2} - z^k_p y^\sigma_k x^p_{j_1,j_2}) \tag{4.1.11}$$

and they are, together with x^i, local coordinates on $P^2_n X$. The inverse of these formulas are (see (2.4.7)):

$$x^\sigma = y^\sigma, \quad x^\sigma_i = x^j_i y^\sigma_j, \quad x^\sigma_{i_1,i_2} = x^{j_1}_{i_1} x^{j_2}_{i_2} y^\sigma_{j_1,j_2} + y^\sigma_k x^k_{i_1,i_2} \tag{4.1.12}$$

On the factor manifold $P^2_n X$ one introduces the derivatives operators (see (2.5.8) and (2.5.9)):

$$\partial_\sigma \equiv \frac{\partial}{\partial y^\sigma}, \quad \partial^j_\sigma \equiv \frac{\partial}{\partial y^\sigma_j}, \quad \partial^{ij}_\sigma \equiv \begin{cases} \frac{\partial}{\partial y^\sigma_{ij}}, & \text{for } i = j \\ \frac{1}{2}\frac{\partial}{\partial y^\sigma_{ij}}, & \text{for } i \neq j \end{cases} \tag{4.1.13}$$

and

$$d_i \equiv \frac{\partial}{\partial x^i} + y^\sigma_i \partial_\sigma + y^\sigma_{ij} \partial^j_\sigma. \tag{4.1.14}$$

The formula for the change of charts on $P^2_n X$ is (see (2.5.13)):

$$\bar{y}^\sigma_i = P^j_i d_j \bar{y}^\sigma, \quad \bar{y}^\sigma_{i_1,i_2} = P^{j_1}_{i_1} P^{j_2}_{i_2} \left[-P^m_l (d_{j_1} d_{j_2} \bar{x}^l)(d_m \bar{y}^\sigma) + d_{j_1} d_{j_2} \bar{y}^\sigma \right]. \tag{4.1.15}$$

Finally, the expressions for the contact forms are (see (2.6.3)):

$$\omega^\sigma \equiv dy^\sigma - y^\sigma_i dx^i, \quad \omega^\sigma_j \equiv dy^\sigma_j - y^\sigma_{ij} dx^i \tag{4.1.16}$$

and their transformation for a change of charts is (see (2.6.8) and (2.6.9)):

$$\bar{\omega}^\sigma = P^\sigma_\nu \omega^\nu, \quad \bar{\omega}^\sigma_i = P^\sigma_\nu P^j_i \bar{\omega}^\nu_j + \left[\partial_\nu \bar{y}^\sigma_i - P^l_k (\partial_\nu \bar{x}^k)(d_l \bar{y}^\sigma_i) \right] \omega^\nu. \tag{4.1.17}$$

4.2 Lagrangian formalism

Here we give a different approach to the Lagrangian formalism based on a certain $(n+1)$-form defined on the Grassmann manifold P^2_n [18]. The description of the formalism will be slightly different and some new material will appear. As in [18], [11], we base our formalism on the operator K (see prop. 2.26). We define the space of *Lagrange-Souriau forms* according to:

$$\Omega^2_{LS} \equiv \{\alpha \in \Omega^2_{n+1,hor}(PX) | \quad d\alpha = 0, \quad K\alpha = 0, \quad i_{V_1} i_{V_2} \alpha = 0, \quad \forall V_i \in \mathrm{Vert}(X)\} \tag{4.2.1}$$

where $\mathrm{Vert}(X)$ is the space of vertical vector fields on P^2_n with respect to the projection $\rho^{2,1}_n$. By definition, a *Lagrangian system* on P^2_n is a couple (E, α) where E is a open sub-bundle of P^2_n and α is a Lagrange-Souriau form. If $\gamma \in \mathrm{Imm} T^2_n$ we say that it *verifies the Euler-Lagrange equations iff*

$$[j^2_0 \gamma]^* \alpha = 0. \tag{4.2.2}$$

It is easy to see that the local expression of a Lagrange-Souriau form is

$$\alpha = \sum_{k=0}^{n} \frac{1}{k!} F^{i_0,\dots,i_k}_{\sigma_0,\dots,\sigma_k} \omega^{\sigma_0}_{i_0} \wedge \omega^{\sigma_1} \wedge \cdots \omega^{\sigma_k} \wedge \theta_{i_1,\dots,i_k}$$

$$+\sum_{k=0}^{n}\frac{1}{(k+1)!}E^{i_1,\ldots,i_k}_{\sigma_0,\ldots,\sigma_k}\omega^{\sigma_0}\wedge\omega^{\sigma_1}\wedge\cdots\omega^{\sigma_k}\wedge\theta_{i_1,\ldots,i_k} \tag{4.2.3}$$

where we have defined:

$$\theta_{i_1,\ldots,i_k}\equiv\binom{n}{k}\varepsilon_{i_1,\ldots,i_n}dx^{i_{k+1}}\wedge\cdots\wedge dx^{i_n},\quad k=0,\ldots,n. \tag{4.2.4}$$

We can admit, without loosing generality, some (anti)-symmetry properties.

$$F^{i_0,i_{P(1)},\ldots,i_{P(k)}}_{\sigma_0,\sigma_{Q(1)},\ldots,\sigma_{Q(k)}}=(-1)^{|P|+|Q|}F^{i_0,i_1,\ldots,i_k}_{\sigma_0,\sigma_1,\ldots,\sigma_k},\quad \forall P,Q\in\mathcal{P}_k \tag{4.2.5}$$

and

$$E^{i_{P(1)},\ldots,i_{P(k)}}_{\sigma_{Q(0)},\ldots,\sigma_{Q(k)}}=(-1)^{|P|+|Q|}E^{i_1,\ldots,i_k}_{\sigma_0,\sigma_1,\ldots,\sigma_k},\quad \forall P\in\mathcal{P}_k,\forall Q\in\mathcal{P}_{k+1}. \tag{4.2.6}$$

The condition

$$K\alpha=0$$

appearing in the definition (4.2.1) of a Lagrange-Souriau form, has the following local form (see [11]):

$$\mathcal{S}^-_{i_0,\ldots,i_k}\mathcal{S}^-_{\sigma_0,\ldots,\sigma_k}F^{i_0,\ldots,i_k}_{\sigma_0,\ldots,\sigma_k}=0,\quad k=1,\ldots,n \tag{4.2.7}$$

where $\mathcal{S}^{\pm}_{\sigma_0,\ldots,\sigma_k}$ are defined similarly to (2.2.3). The local form of the Euler-Lagrange equation (4.2.2) is simply:

$$E_\sigma\circ[j^2_0\gamma]=0 \tag{4.2.8}$$

Remark 4.1 For the case $n=1$, the 2-form α above appears in [34] and the condition (4.2.7) is investigated in [20] and [19].

The justification of the terminology for (4.2.8) is contained in the following result:

Proposition 4.2 *The expressions E_σ verify the Helmholtz equations (3.2.24).*

Proof One writes in detail the closeness condition $d\alpha=0$ and find out, in particular, the following equations:

$$d_jE^j_{\sigma_0,\sigma_1}+\partial_{\sigma_0}E_{\sigma_1}-\partial_{\sigma_1}E_{\sigma_0}=0, \tag{4.2.9}$$

$$d_lF^{jl}_{\nu,\sigma}-\partial_\sigma F^j_\nu+\partial^j_\nu E_\sigma+E^j_{\nu,\sigma}=0, \tag{4.2.10}$$

$$\partial^{jk}_\nu E_\sigma+\frac{1}{2}\left(F^{jk}_{\nu,\sigma}+F^{kj}_{\nu,\sigma}\right)=0, \tag{4.2.11}$$

$$F^{jl}_{\sigma,\nu}-F^{lj}_{\nu,\sigma}=\partial^l_\nu F^j_\sigma-\partial^j_\sigma F^l_\nu. \tag{4.2.12}$$

From (4.2.7) we get, in particular:

$$F^j_\sigma=0 \tag{4.2.13}$$

and

$$F^{jl}_{\nu,\sigma} - F^{jl}_{\sigma,\nu} - F^{lj}_{\nu,\sigma} + F^{lj}_{\sigma,\nu} = 0. \tag{4.2.14}$$

We use (4.2.13) in (4.2.10) and get:

$$d_l F^{jl}_{\nu,\sigma} + \partial^j_\nu E_\sigma + E^j_{\nu,\sigma} = 0,$$

If we substitute the last term of this relation into (4.2.9) and use (4.2.11) we get:

$$\partial_{\sigma_0} E_{\sigma_1} - \partial_{\sigma_1} E_{\sigma_0} = d_j \partial^j_{\sigma_0} E_{\sigma_1} - d_j d_l \partial^{jl}_{\sigma_0} E_{\sigma_1}. \tag{4.2.15}$$

Next, we take the symmetric part in ν, σ of (4.2.10) and obtain:

$$\partial^j_\nu E_\sigma + \partial^j_\sigma E_\nu = -d_l \left(F^{jl}_{\nu,\sigma} + F^{jl}_{\sigma,\nu} \right).$$

One uses here (4.2.14) and next (4.2.12) + (4.2.13) to get:

$$\partial^j_\nu E_\sigma + \partial^j_\sigma E_\nu = 2 d_l \partial^{jl}_\sigma E_\nu. \tag{4.2.16}$$

Finally, the antisymmetric part of (4.2.11) in σ, ν is:

$$\partial^{jk}_\nu E_\sigma = \partial^{jk}_\sigma E_\nu. \tag{4.2.17}$$

The equations (4.2.15), (4.2.16) and (4.2.17) are the Helmholtz-Sonin equations for the expressions E_σ. □

We now mention a result derived in [18]:

Proposition 4.3 *There exists in every chart a local n-form β on $P^1_n X$ having the local expression*

$$\beta = \sum_{k=0}^{n} \frac{1}{k!} L^{i_1,\dots,i_k}_{\sigma_1,\dots,\sigma_k} \omega^{\sigma_1} \wedge \cdots \omega^{\sigma_k} \wedge \theta_{i_1,\dots,i_k} \tag{4.2.18}$$

where

$$L^{i_1,\dots,i_k}_{\sigma_1,\dots,\sigma_k} = \mathcal{S}^-_{i_1,\dots,i_k} \mathcal{S}^-_{\sigma_1,\dots,\sigma_k} \partial^{i_1}_{\sigma_1} \dots \partial^{i_k}_{\sigma_k} L, \quad k = 0, \dots, n \tag{4.2.19}$$

such that:

$$\alpha = d(\rho^{2,1}_n)^* \beta. \tag{4.2.20}$$

Remark 4.4 The form β is a generalization of the Poincaré-Cartan form [6], [30]. It had appeared in the literature in [22], [4], [5], [32]. For other generalizations of the Poincaré-Cartan form see [31], [21], [9], [8], [10], [33] and [23] where the notion of Lepage is introduced for such generalizations.

As a consequence we can express the coefficients of the form α given by (4.2.3) in terms of the smooth function L [11]:

$$F^{i_0,\dots,i_k}_{\sigma_0,\dots,\sigma_k} = \partial^{i_0}_{\sigma_0} L^{i_1,\dots,i_k}_{\sigma_1,\dots,\sigma_k} - L^{i_0,\dots,i_k}_{\sigma_0,\dots,\sigma_k}, \quad k = 0, \dots, n \tag{4.2.21}$$

and

$$E^{i_1,\ldots,i_k}_{\sigma_0,\ldots,\sigma_k} = \mathcal{S}^-_{i_0,\ldots,i_k} \partial_{\sigma_0} L^{i_1,\ldots,i_k}_{\sigma_1,\ldots,\sigma_k} - d_{i_0} L^{i_0,\ldots,i_k}_{\sigma_0,\ldots,\sigma_k}, \quad k = 0,\ldots,n. \tag{4.2.22}$$

In particular, for $k = 0$ the preceding formula is

$$E_\sigma = \partial_\sigma L - d_i \partial^i_\sigma L \tag{4.2.23}$$

i.e. the Euler-Lagrange operator. This is another justification of the terminology for the equations (4.2.2) (and (4.2.8).) This shows that the expressions $E^{i_1,\ldots,i_k}_{\sigma_0,\ldots,\sigma_k}, \quad k = 0,\ldots,n$ are also some generalizations of the Euler-Lagrange expressions, however, different from the Lie-Euler expressions introduced in [1] and given by (3.1.10). The coefficients of α verify some recurrence relations:

$$E^{i_1,\ldots,i_k}_{\sigma_0,\ldots,\sigma_k} = -\mathcal{S}^-_{i_1,\ldots,i_k} \mathcal{S}^-_{\sigma_0,\ldots,\sigma_k} \left(\partial^{i_1}_{\sigma_0} + d_l \partial^{li_1}_{\sigma_0} \right) E^{i_2,\ldots,i_k}_{\sigma_1,\ldots,\sigma_k}, \quad k = 1,\ldots,n \tag{4.2.24}$$

and

$$F^{i_0,\ldots,i_k}_{\sigma_0,\ldots,\sigma_k} = \text{const} \times \mathcal{S}^-_{\sigma_0,\ldots,\sigma_k} \mathcal{S}^-_{i_0,\ldots,i_k} \partial^{i_0}_{\sigma_0} F^{i_1,\ldots,i_k}_{\sigma_1,\ldots,\sigma_k}, \quad k = 2,\ldots,n. \tag{4.2.25}$$

Proof By exploiting the conditions $d\alpha = 0$ and $K\alpha = 0$ written in local coordinates. Alternatively, one computes the right hand sides of these formulas using (4.2.21) and (4.2.22) and obtains the left hand sides. □

These recurrence relations have an important consequence:

Corollary 4.5 *In the conditions above we have*

$$\alpha \equiv 0 \Longleftrightarrow E_\sigma \equiv 0. \tag{4.2.26}$$

We close this subsection giving the connection between the form α introduced here and the Lagrange-Souriau form σ introduced in [11], [18]. We have

Proposition 4.6 *Let α be a Lagrange-Souriau form. Then there exists a closed $(n+1)$-form σ on P^1_n such that*

$$\alpha = (\rho^{2,1}_n)^* \sigma. \tag{4.2.27}$$

Proof One makes explicit the condition

$$d\alpha = 0$$

appearing in the definition of a Lagrange-Souriau form (4.2.1) and finds out in particular:

$$\partial^{jk}_\nu F^{i_0,\ldots,i_k}_{\sigma_0,\ldots,\sigma_k} = 0, \quad k = 0,\ldots,n \tag{4.2.28}$$

$$\partial^{jk}_\nu E^{i_1,\ldots,i_k}_{\sigma_0,\ldots,\sigma_k} + \frac{1}{2}\left[F^{j,k,i_0,\ldots,i_k}_{\nu,\sigma_0,\ldots,\sigma_k} + (j \leftrightarrow k) \right] = 0, \quad k = 0,\ldots,n. \tag{4.2.29}$$

It follows that the generic expression of the coefficients $E^{i_1,\ldots,i_k}_{\sigma_0,\ldots,\sigma_k}$ is:

$$E^{i_1,\ldots,i_k}_{\sigma_0,\ldots,\sigma_k} = G^{i_1,\ldots,i_k}_{\sigma_0,\ldots,\sigma_k} - y^\nu_{jk} F^{j,k,i_1,\ldots,i_k}_{\nu,\sigma_0,\ldots,\sigma_k} \tag{4.2.30}$$

where $G^{i_1,\ldots,i_k}_{\sigma_0,\ldots,\sigma_k}$ have the antisymmetry property (4.2.6) and verify

$$\partial^{jk}_{\nu} G i_0, \ldots, i_{k\,\sigma_0,\ldots,\sigma_k} = 0, \quad k = 0, \ldots, n. \tag{4.2.31}$$

Substituting (4.2.30) into the expression (4.2.3) one obtains (4.2.27) with

$$\sigma = \sum_{k=0}^{n} \frac{1}{k!} F^{i_0,\ldots,i_k}_{\sigma_0,\ldots,\sigma_k} dy^{\sigma_0}_{i_0} \wedge \omega^{\sigma_1} \wedge \cdots \omega^{\sigma_k} \wedge \theta_{i_1,\ldots,i_k}$$
$$+ \sum_{k=0}^{n} \frac{1}{(k+1)!} G^{i_1,\ldots,i_k}_{\sigma_0,\ldots,\sigma_k} \omega^{\sigma_0} \wedge \omega^{\sigma_1} \wedge \cdots \omega^{\sigma_k} \wedge \theta_{i_1,\ldots,i_k}. \tag{4.2.32}$$

This finishes the proof. □

The formalism presented in this section can be extended to Grassmann manifolds of arbitrary order $r > 2$. Some steps in this direction are contained in [25] and a detailed analysis is contained in [17].

5 Some applications

The most interesting applications are related to infinite dimensional groups, in particular the so-called gauge groups. We have in mind especially non-Abelian gauge theories, gravitation theory and string theory. The general setting is the following: we have a group of transformations ϕ_ξ of the basic manifold X which are parametrised by some set of smooth functions ξ and act naturally on the set of solutions of some Lagrangian theory; we want to classify the action functionals which are invariant with respect to these transformations. We will consider that we are in the conditions of the preceding Section. It can be easily proved that this condition is equivalent to a more convenient form: if $\dot{\phi} \equiv j^1\phi_\xi$ is the extension of the transformation ϕ_ξ to P^1X then $\dot{\phi}$ are Noetherian symmetries i.e.

$$(\dot{\phi})^* \sigma = \sigma.$$

This condition will impose severe restrictions on the coefficients F and G from (4.2.32); for the three cases we consider these restrictions can be completely solved i.e. the most general form can be obtained. It remains to imppose the closeness condition $d\sigma = 0$ and determine the most general form of the (local) Lagrangian following as in proposition 4.2.20. The details can be found in [11] and references quoted there.

5.1 Non-Abelian gauge theories

We consider here only the case of pure gauge fields i.e. no matter fields. In the framework of the preceding Section we take $X = M \times V$ where $V = M \times Lie(G)$; here G is a non-Abelian Lie group and $Lie(G)$ is the associated Lie algebra. If $\{e_a\}^r_{a=1}$ is a basis in $Lie(G)$ then the coordinates on X are: $(x^\mu, v^{a\nu})$ where $v^{a\nu}$ are the Yang-Mills potentials. We take as global coordinates on $E = P^1_n X$ to be $(x^\mu, v^{a\nu}, v^{a\nu}_\mu)$.

First we impose Poincaré invariance; the action of $\mathcal{P}^{\uparrow}_{+}$ on X is:

$$\phi_{\Lambda,a}(x, v) = (\Lambda \cdot x + a, \Lambda \cdot v). \tag{5.1.1}$$

and we require that $\phi_{\Lambda,a}$ are Noetherian symmetries for any $(\Lambda, a) \in \mathcal{P}^{\uparrow}_{+}$. This conditions translates very easily on conditions on the functions $F^{\cdots}_{\cdots}$ and $G^{\cdots}_{\cdots}$ from (4.2.32). For $\Lambda = 1$ this is equivalent to the x-independence of the functions and for $a = 0$ to the Lorentz covariance of these functions.

Let us impose now the condition of gauge invariance in a similar way. The gauge group $Gau(G)$ consists of smooth maps $g : M \to G$ with pointwise multiplication as composition law. We will consider only gauge transformations in the connected component of the identity $Gau(G)^0$ of the gauge group. So, we can use only infinitesimal gauge transformations $\xi : M \to Lie(G)$ which act on the set of evolutions as follows:

$$\Phi_\xi(x^\mu, v^{a\nu}(x)) = (x^\mu, v^{a\nu}(x) + \xi^b(x) f^a_{bc} v^{c\nu}(x) + (\partial^\nu \xi^a)(x)) \tag{5.1.2}$$

where f^a_{bc} are the structure constants of $Lie(G)$ defined through

$$[e_a, e_b] = f^c_{ab} e_c \qquad \forall\, a, b, c = 1, ..., r. \tag{5.1.3}$$

The transformation (5.1.2) is induced by the following transformation of X:

$$\phi_\xi(x^\mu, v^{a\nu}) = (x^\mu, v^{a\nu} + \xi^b(x) f^a_{bc} v^{c\nu} + (\partial^\nu \xi^a)(x)) \tag{5.1.4}$$

and we postulate that ϕ_ξ is a Noetherian symmetry i.e.:

$$(\dot{\phi}_\xi)^* \sigma = \sigma. \tag{5.1.5}$$

It is not very difficult to prove that this gauge invariance implies that the functions $F^{\cdots}_{\cdots}$ from (4.2.32) depend on the variables $(x^\mu, v^{a\nu}, v^{a\nu}_\mu)$. only of the *Yang-Mills field strength* which is defined by:

$$F^a_{\mu\nu} \equiv v^a_{\mu\nu} - v^a_{\nu\mu} + f^a_{bc} v^b_\mu v^c_\nu. \tag{5.1.6}$$

One can easily discover the “Lagrangian” nature of the functions $F^{\cdots}_{\cdots}$ i.e. one shows that there exists a F-dependent function L_{YM} such that:

$$F^{\mu_0,\ldots\ldots,\mu_k}_{a_0\nu_0,\ldots,a_k\nu_k} = \frac{\partial (L_{YM})^{\mu_1,\ldots\ldots,\mu_k}_{a_1\nu_1,\ldots,a_k\nu_k}}{\partial F^{a_0\nu_0}{}_{\mu_0}} - (L_{YM})^{\mu_0,\ldots\ldots,\mu_k}_{a_0\nu_0,\ldots,a_k\nu_k} \tag{5.1.7}$$

where:

$$(L_{YM})^{\mu_1,\ldots\ldots,\mu_k}_{a_1\nu_1,\ldots,a_k\nu_k} = \frac{1}{k!} \sum_{\sigma \in P_k} (-1)^{|\sigma|} \frac{\partial^k L_{YM}}{\partial F^{a_1\nu_1}{}_{\mu_{\sigma(1)}} \ldots \partial F^{a_k\nu_k}{}_{\mu_{\sigma(k)}}}. \tag{5.1.8}$$

Here L_{YM} must verify

$$X_b L_{YM} = P_b \tag{5.1.9}$$

where X_b is the differential operator:

$$X_b \equiv \frac{1}{2} f^d_{bc} F^c_{\rho\zeta} \frac{\partial}{\partial F^d_{\rho\zeta}}. \tag{5.1.10}$$

and P_b is a polynomial in F (an "anomaly") which verifies the consistency condition:

$$X_b P_c - X_c P_b = -f^a_{bc} P_a. \tag{5.1.11}$$

One can show that in fact, $P_b = 0$ and this expresses the invariance of L_{YM} with respect to the adjoint action of G^0; also L_{YM} can be chosen strictly Lorentz invariant. All these facts are of cohomological nature and use of standard results as Whitehead lemmas is required.

The analysis of the functions $G^{\cdots}_{\cdots}$ from (4.2.32) is greatly simplified if we separate in σ the contribution of the Yang-Mills Lagrangian L_{YM} that is, we define:

$$\sigma_{CS} \equiv \sigma - \sigma_{L_{YM}}. \tag{5.1.12}$$

In this way, σ_{CS} will contain only the functions of the type $G^{\cdots}_{\cdots}$ and the structure equations following from the closeness of the form σ are giving the generic form:

$$G^{\mu_1,\ldots,\mu_k,\nu_0,\ldots,\nu_k}_{a_0,\ldots\ldots\ldots\ldots,a_k}(F) = \sum_{p=k}^{m} \frac{1}{(p-k)!2^{p-k}} C^{\mu_1,\ldots,\mu_p,\nu_0,\ldots,\nu_p}_{a_0,\ldots\ldots\ldots\ldots\ldots,a_p} \prod_{i=k+1}^{p} F^{a_i}_{\mu_i\nu_i} \tag{5.1.13}$$

where the constants $C^{\mu_1,\ldots\ldots,\mu_p,\nu_0,\ldots,\nu_p}_{a_0,\ldots\ldots\ldots\ldots\ldots,a_p}$ are completely antisymmetric in the upper indices and completely symmetric in the lower indices and verify:

$$\sum_{i=0}^{p} C^{\mu_1,\ldots\ldots,\mu_p,\nu_0,\ldots\ldots,\nu_p}_{a_0,\ldots,a_{i-1},c,a_{i+1},\ldots,a_p} f^c_{ba_i} = 0 \tag{5.1.14}$$

for all $p = 0, \ldots, n$. We have two distinct cases:

(a) $n = 2m$

In this case we have for any k:

$$C^{\cdots}_{\cdots} = 0 \qquad \Longleftrightarrow \qquad G^{\cdots}_{\cdots} = 0. \tag{5.1.15}$$

It follows that in this case:

$$\sigma = \sigma_{L_{YM}} \tag{5.1.16}$$

where L_{YM} is constrained only by invariance with respect to the Poincaré group and with respect to the coadjoint action of the group G^0.

(b) $n = 2m + 1$

In this case we have for $k < m$:

$$C^{\mu_1,\ldots,\mu_k,\nu_0,\ldots,\nu_k}_{a_0,\ldots\ldots\ldots\ldots\ldots,a_k} = 0 \tag{5.1.17}$$

and:

$$C^{\mu_1,\ldots,\mu_m,\nu_0,\ldots,\nu_m}_{a_0,\ldots\ldots\ldots\ldots\ldots,a_m} = C_{a_0,\ldots,a_m} \varepsilon^{\mu_1,\ldots,\mu_m,\nu_0,\ldots,\nu_m} \tag{5.1.18}$$

where the tensor $C_{\ldots}$ is completely symmetric in the indices $a_0, \ldots, a_k$ and invariant with respect to the adjoint action of G^0.

The corresponding expression for the Lagrangian is of Chern-Simons type:

$$L_{CS}(x,A,\chi) = \sum_{k=0}^{n} \frac{1}{k!} l^{\mu_1,\ldots\ldots\ldots,\mu_k}_{a_1\nu_1,\ldots,a_k\nu_k}(x,v) \prod_{i=1}^{k} v^{a_i\nu_i}_{\mu_i} \tag{5.1.19}$$

where one can choose:

$$l^{\mu_1,\ldots\ldots\ldots\ldots,\mu_k}_{a_1\nu_1,\ldots,a_k\nu_k}(v) = \frac{1}{n-k}\frac{1}{(m-k)!2^{m-p}} v^{a_0\nu_0} C_{a_0,\ldots,a_m} \varepsilon^{\mu_1,\ldots,\mu_m}{}_{\nu_0,\ldots,\nu_m} \prod_{i=k+1}^{m} f^{a_i}_{b_ic_i} v^{b_i\nu_i} v^{c_i}_{\mu_i}. \tag{5.1.20}$$

So in this case, like in the Abelian case we have:

$$\sigma = \sigma_{L_{YM}} + \sigma_{CS}. \tag{5.1.21}$$

One can find explicit solutions for the Chern-Simons part as follows. Let $\{t^a\}_{a=1}^{r}$ be any finite dimensional representation of $Lie(G)$ i.e.

$$[t_a, t_b] = f^c_{ab} t_c. \tag{5.1.22}$$

Then we can take:

$$C_{a_0,\ldots,a_m} = Sym_{a_0,\ldots,a_m} Tr(t_{a_0} \ldots t_{a_m}). \tag{5.1.23}$$

If we impose in addition dilation invariance, we again have two distinct cases:
(a) $n = 2m$
In this case L_{YM} is a polynomial in F, homogeneous of degree m:

$$L_{YM}(\lambda^2 F) = \lambda^n L_{YM}(F). \tag{5.1.24}$$

and we have only the Yang-Mills term.
(b) $n = 2m + 1$
In this case:

$$L_{YM} = 0. \tag{5.1.25}$$

and we have only the Chern-Simons term.

5.2 Gravitation theory

Gravitation theory can be also analyzed in the general framework used in this paper. We will not obtain the Hilbert Lagrangian from which Einstein equations are usually derived because this is a second order Lagrangian. Nevertheless, we will obtain the most general first-order Lagrangian for the gravitation theory, imposing the invariance with respect to the general coordinate transformations; this first order Lagrangian is equivalent to the Hilbert Lagrangian, up to a divergence.

Let M be a differentiable manifold, interpreted as the space-time (or events) manifold. To construct the basic manifold X for the gravitation theory, we need the tensor bundle

$T^0_2(M)$ of covariant tensors of rank $(0,2)$ on M. If (x^μ) $\mu = 1,...,n$ is a local coordinates system on M, then a local coordinates system on $T^0_2(M)$ is of the form: $(x^\mu, g_{\alpha\beta})$ where $g_{\alpha\beta}$ is a symmetric matrix (the components of the metric tensor $g \equiv g_{\alpha\beta} dx^\alpha \otimes dx^\beta \in T^0_2(M)$).

We take as the basic manifold X, the open subset $S \subseteq T^0_2(M)$ corresponding to the Lorentzian metrics i.e. to tensors $g_{\alpha\beta}$ which are equivalent to the Minkowski metric $G_{\alpha\beta}$ up to a change of basis; explicitly $\exists A \in GL(n, R)$ such that:

$$A^\alpha_\gamma A^\beta_\delta g_{\alpha\beta} = G_{\gamma\delta}. \tag{5.2.1}$$

The dimension of X is $N = \frac{n(n+3)}{2}$. The local coordinates on $E \equiv P^1_n X$ will be denoted by $(x^\mu, g_{\alpha\beta}, g_{\alpha\beta,\mu})$ where $g_{\alpha\beta,\mu}$ will be supposed symmetric in the indices α and β.

We impose the basic condition of invariance with respect to general coordinate transformations as follows. If $\xi \in \mathit{Diff}(M)$ let us define $A^\rho_\mu(x)$ to be the inverse of the matrix $\frac{\partial \xi^\mu}{\partial x^\nu}$ i.e.

$$\frac{\partial \xi^\mu}{\partial x^\nu}(x) A^\rho_\mu(x) = \delta^\rho_\nu. \tag{5.2.2}$$

The group $\mathit{Diff}(M)$ acts on the set of evolutions as follows:

$$\Phi_\xi(x, g_{\alpha\beta})(x)) = (\xi(x), A^\gamma_\alpha(x) A^\delta_\beta(x) g_{\gamma\delta} \circ \xi^{-1}(x)). \tag{5.2.3}$$

Let us define the following action of $\mathit{Diff}(M)$ on X:

$$\phi_\xi(x, g_{\alpha\beta}) = (\xi(x), A^\gamma_\alpha(x) A^\delta_\beta(x) g_{\gamma\delta}). \tag{5.2.4}$$

Then, we postulate that $\phi_\xi \in \mathit{Diff}(X)$ is a Noetherian symmetry i.e.:

$$(\dot{\phi}_\xi)^* \sigma = \sigma. \tag{5.2.5}$$

We consider for the moment only the group $\mathit{Diff}(M)_+$ of reparametrizations of positive Jacobian. As before, we translate this condition into equations satisfied by the coefficient functions $F^{\cdot\cdot}$ and $G^{\cdots}$. from (4.2.32). It is convenient to consider infinitesimal diffeomorphisms of the form

$$\xi^\mu(x) = x^\mu + \lambda^\mu(x) \tag{5.2.6}$$

with λ^μ arbitrary functions.

In a familiar way, we start with the analysis of the functions $F^{\cdots}$. Using standard facts in invariant theory [33] and taking into account the various symmetry properties one arrives at an essential simplification, namely

$$\sigma^{\mu_0,\ldots,\mu_k;\alpha_0\beta_0,\ldots,\alpha_k\beta_k} = 0 \qquad \forall k > 1 \tag{5.2.7}$$

and

$$\sigma^{\mu_1,\mu_2;\alpha_1\beta_1,\alpha_2\beta_2}(g_{\alpha\beta}) = \frac{1}{2}\kappa(|\det(g)|)^{1/2}\times$$

$$[g^{\alpha_1\alpha_2}(g^{\mu_1\beta_1}g^{\mu_2\beta_2}+g^{\mu_1\beta_2}g^{\mu_2\beta_1})+g^{\beta_1\beta_2}(g^{\mu_1\alpha_1}g^{\mu_2\alpha_2}+g^{\mu_1\alpha_2}g^{\mu_2\alpha_1})$$
$$+g^{\alpha_1\beta_2}(g^{\mu_1\alpha_2}g^{\mu_2\beta_1}+g^{\mu_1\beta_1}g^{\mu_2\alpha_2})+g^{\alpha_2\beta_1}(g^{\mu_1\alpha_1}g^{\mu_2\beta_2}+g^{\mu_1\beta_2}g^{\mu_2\alpha_1})$$
$$+4g^{\mu_1\mu_2}g^{\alpha_1\beta_1}g^{\alpha_2\beta_2}-2g^{\mu_1\mu_2}(g^{\alpha_1\alpha_2}g^{\beta_1\beta_2}+g^{\alpha_1\beta_2}g^{\alpha_2\beta_1})$$
$$-2g^{\alpha_1\beta_1}(g^{\mu_1\alpha_2}g^{\mu_2\beta_2}+g^{\mu_1\beta_2}g^{\mu_2\alpha_2})-2g^{\alpha_2\beta_2}(g^{\mu_1\alpha_1}g^{\mu_2\beta_1}+g^{\mu_1\beta_1}g^{\mu_2\alpha_1})]. \tag{5.2.8}$$

The hardest part is the determination of the functions $G^{\cdots}$. We get another essential simplification, namely:

$$G^{\mu_1,\dots,\mu_k;\alpha_0\beta_0,\dots,\alpha_k\beta_k}=0 \qquad \forall k>2 \tag{5.2.9}$$

so, we must determine only the expressions $G^{\alpha_0\beta_0}, G^{\mu_1;\alpha_0\beta_0,\alpha_1\beta_1}, G^{\mu_1\mu_2;\alpha_0\beta_0,\alpha_1\beta_1,\alpha_2\beta_2}$. We obtain

$$G^{\mu_1;\alpha_0\beta_0,\alpha_1\beta_1}(g_{\alpha\beta},g_{\alpha\beta,\mu})=\frac{1}{2}\left(G^{\mu_1\mu_2;\alpha_0\beta_0,\alpha_1\beta_1,\alpha_2\beta_2}\right.$$
$$\left.+\frac{\partial F^{\mu_1\mu_2;\alpha_1\beta_1,\alpha_2\beta_2}}{\partial g_{\alpha_0\beta_0}}-\frac{\partial F^{\mu_1\mu_2;\alpha_0\beta_0,\alpha_2\beta_2}}{\partial g_{\alpha_1\beta_1}}\right)g_{\alpha_2\beta_2,\mu_2} \tag{5.2.10}$$

$$G^{\alpha_0\beta_0}(g_{\alpha\beta},g_{\alpha\beta,\mu})=\frac{1}{8}\Bigg(G^{\mu_1\mu_2;\alpha_0\beta_0,\alpha_1\beta_1,\alpha_2\beta_2}$$
$$+\frac{\partial F^{\mu_1\mu_2;\alpha_1\beta_1,\alpha_2\beta_2}}{\partial g_{\alpha_0\beta_0}}-\frac{\partial F^{\mu_1\mu_2;\alpha_0\beta_0,\alpha_2\beta_2}}{\partial g_{\alpha_1\beta_1}}$$
$$-\frac{\partial F^{\mu_1\mu_2;\alpha_0\beta_0,\alpha_1\beta_1}}{\partial g_{\alpha_2\beta_2}}\,g_{\alpha_1\beta_1,\mu_1}\Bigg)\,g_{\alpha_2\beta_2,\mu_2}$$
$$+\lambda(|\det(g)|)^{1/2}g^{\alpha_0\beta_0} \tag{5.2.11}$$

for some $\lambda\in\mathbb{R}$

$$G^{\mu_1\mu_2;\alpha_0\beta_0,\alpha_1\beta_1,\alpha_2\beta_2}=$$
$$\frac{\partial l^{\mu_1\mu_2;\alpha_1\beta_1,\alpha_2\beta_2}}{\partial g_{\alpha_0\beta_0}}-\frac{\partial l^{\mu_1\mu_2;\alpha_0\beta_0,\alpha_2\beta_2}}{\partial g_{\alpha_1\beta_1}}+\frac{\partial l^{\mu_1\mu_2;\alpha_0\beta_0,\alpha_1\beta_1}}{\partial g_{\alpha_2\beta_2}} \tag{5.2.12}$$

where we have defined:

$$l^{\mu_1\mu_2;\alpha_1\beta_1,\alpha_2\beta_2}(g_{\alpha\beta})\equiv\frac{1}{2}\kappa(|\det(g)|)^{1/2}[g^{\alpha_1\alpha_2}\left(g^{\mu_1\beta_2}g^{\mu_2\beta_1}-g^{\mu_1\beta_1}g^{\mu_2\beta_2}\right)$$
$$+g^{\beta_1\beta_2}\left(g^{\mu_1\alpha_2}g^{\mu_2\alpha_1}-g^{\mu_1\alpha_1}g^{\mu_2\alpha_2}\right)$$
$$+g^{\alpha_1\beta_2}\left(g^{\mu_1\alpha_2}g^{\mu_2\beta_1}-g^{\mu_1\beta_1}g^{\mu_2\alpha_2}\right)$$
$$+g^{\alpha_2\beta_1}\left(g^{\mu_1\beta_2}g^{\mu_2\alpha_1}-g^{\mu_1\alpha_1}g^{\mu_2\beta_2}\right)]. \tag{5.2.13}$$

To make connection with the usual formalism one tries to determine a Lagrangian. It is not very hard to show that the following expression

$$L(x,g,\chi) = \kappa(|\det(g)|)^{1/2} g^{\mu_1\alpha_2}\left(g^{\alpha_1\beta_2}g^{\mu_2\beta_1} - g^{\alpha_1\beta_1}g^{\beta_2\mu_2}\right)\Gamma_{\alpha_1\beta_1,\mu_1}\Gamma_{\alpha_2\beta_2,\mu_2} + \lambda(|\det(g)|)^{1/2} \quad (5.2.14)$$

where $\Gamma_{...}$ are the Christoffel symbols:

$$\Gamma_{\alpha\beta,\mu} \equiv \frac{1}{2}\left(g_{\mu\alpha,\beta} + g_{\mu\beta,\alpha} - g_{\alpha\beta,\mu}\right) \quad (5.2.15)$$

produces $\sigma = \sigma_L$.

We have found out an well-known expression: the first term from this expression gives the usual Einstein equations (see e.g. [27], $ 93) and the second term is the cosmological contribution.

One can easily prove that σ determined above admits dilation invariance:

$$\phi_a(x,g) = (ax, a^{-2}g) \qquad \forall a \in \mathbb{R}_+ \quad (5.2.16)$$

are Noetherian symmetries. Next we note that if $\xi \in \mathit{Diff}(M)$ is arbitrary (not necessarily of positive Jacobian) then we have:

$$(\dot{\phi}_\xi)^*\sigma = sign\left(\frac{\partial \xi^\mu}{\partial x^\nu}\right)\sigma \quad (5.2.17)$$

and ϕ_ξ continues to be a symmetry.

5.3 Extended objects

A very interesting case is provided by the so-called extended objects as strings, membranes and, generally, p-branes and we will show that they can be completely studied using the framework developed above. We take in the general scheme from preceding Section $X \equiv P \times M$ where P is a p-dimensional parameter manifold and M is, as usual, the n-dimensional Minkowski space. It is natural to suppose that $n \geq p$. The coordinates on X are (τ^a, X^μ) $a = 1,...,p$ $\mu = 0,...,n-1$. The global coordinates on $E \equiv P^1_n X$ are: $(\tau^a, X^\mu, U^\mu{}_a)$.

We impose now the reparametrization invariance. If $\xi \in \mathit{Diff}(P)$ then the action of this reparametrization on X is simply:

$$\phi_\xi(\tau, X) = (\xi(\tau), X). \quad (5.3.1)$$

We will impose the condition that ϕ_ξ is a Noetherian symmetry for any $\xi \in \mathit{Diff}(P)_+$ i.e. for diffeomorphisms of positive Jacobian:

$$(\dot{\phi}_\xi)^*\sigma = \sigma. \quad (5.3.2)$$

It is not hard to translate this condition into equations on $F^{\cdots}_{\cdots}$ and $G^{\cdots}_{\cdots}$. Let $A_a{}^b(\tau)$ be the inverse matrix of $\frac{\partial \xi^a}{\partial \tau^a}$ i.e.

$$\frac{\partial \xi^b}{\partial \tau^a}(\tau) A_b{}^c(\tau) = \delta^c_a. \quad (5.3.3)$$

Then we have:

$$\prod_{i=0}^{k} A_{a_i}{}^{b_i} F^{a_0,\dots,a_k}_{\mu_0,\dots,\mu_k} \circ \dot{\phi}_\xi = (\det(A)) F^{b_0,\dots,b_k}_{\mu_0,\dots,\mu_k} \tag{5.3.4}$$

and:

$$\prod_{i=1}^{k} A_{a_i}{}^{b_i} \left(G^{a_1,\dots,a_k}_{\mu_0,\dots,\mu_k} \circ \dot{\phi}_\xi - \frac{\partial A_b{}^d}{\partial \tau^c} U^\nu{}_b F^{b,c,a_1,\dots,a_k}_{\nu,\mu_0,\dots,\mu_k} \circ \dot{\phi}_\xi \right) = (\det(A)) G^{b_1,\dots,b_k}_{\mu_0,\dots,\mu_k} \tag{5.3.5}$$

for $k = 0, \dots, p$

Because $\xi \in \mathit{Diff}(P)_+$ we can consider ξ an infinitesimal diffeomorphism i.e.

$$\xi^a(\tau) = \tau^a + \theta^a(\tau) \tag{5.3.6}$$

with θ^a arbitrary functions.

In addition we impose Poincaré invariance in the "target" space M. Namely:

$$\phi_{\Lambda,a}(\tau, X) = (\tau, \Lambda X + a) \tag{5.3.7}$$

is a Noetherian symmetry for any $(\Lambda, a) \in \mathcal{P}^\uparrow_+$. This condition translates into the X-independence of those functions and into their Lorentz covariance as functions of U.

According to the general strategy, we first determine the functions $F^{\cdots}_{\cdots}$. One can show as before that there is a U-dependent smooth function L_0 such that:

$$\sigma^{a_0,\dots,a_k}_{\mu_0,\dots,\mu_k} = \frac{\partial (L_0)^{a_1,\dots,a_k}_{\mu_1,\dots,\mu_k}}{\partial U^{\mu_0}{}_{a_0}} - (L_0)^{a_0,\dots,a_k}_{\mu_0,\dots,\mu_k} \tag{5.3.8}$$

where:

$$(L_0)^{a_0,\dots,a_k}_{\mu_0,\dots,\mu_k} = \frac{1}{k!} \sum_{\sigma \in P_k} (-1)^{|\sigma|} \frac{\partial^k L_0}{\partial U^{\mu_1}{}_{a_{\sigma(1)}} \dots \partial U^{\mu_k}{}_{a_{\sigma(k)}}}. \tag{5.3.9}$$

We exploit now the Lorentz covariance of the functions $\sigma^{\cdots}_{\cdots}$. Using a cohomology argument as in previous Subsections one shows that, without modifying σ one can redefine L_0 such that:

$$L_0(\Lambda^\mu{}_\nu U^\nu{}_a) = L_0(U^\mu{}_a) \qquad \forall \Lambda \in \mathcal{L}^\uparrow_+ \tag{5.3.10}$$

and

$$L_0(A_a{}^b U^\mu{}_b) = L_0(U^\mu{}_a) \qquad \forall A \in GL(p, R)_+. \tag{5.3.11}$$

From these relations one can completely determine L_0. We use well-known results in invariant theory for (pseudo)-orthogonal groups. Because we have admitted from the beginning that $n \geq p$ one applies the Gramm trick and shows that the first relation implies that L_0 is a function only of the Lorentz scalars (U_a, U_b) i.e.:

$$L_0(U) = l((U_a, U_b)) \tag{5.3.12}$$

where l is a smooth symmetric function of the variables ζ_{ab} $a, b = 1, ..., p$. Next, second relation translates for l into:

$$l(A_a{}^c A_b{}^d \zeta_{cd}) = l(\zeta_{ab}) \qquad \forall A \in GL(p, R)_+. \tag{5.3.13}$$

Now it is elementary to find out that l is a smooth function of the variable $\det(\zeta)$ i.e:

$$L(\zeta_{ij}) = l_0(\det(\zeta)). \tag{5.3.14}$$

It remains to consider the case $A = \lambda I \qquad \forall \lambda \in \mathbb{R}_+$. This translates into the following homogeneity property for l_0 namely:

$$l_0(\lambda^2 x) = \lambda l_0(x) \qquad \forall \lambda \in \mathbb{R}_+. \tag{5.3.15}$$

We can conclude easily that if l_0 is defined on the whole real axis then the smoothness condition forces $l_0 = 0$. On the contrary, if l_0 is defined only on one of the the domains: $D^+ \equiv \{x > 0\}$ or $D^- \equiv \{x < 0\}$ then l_0 is of the form:

$$l_0(x) = \kappa \sqrt{|x|} \tag{5.3.16}$$

for some $\kappa \in \mathbb{R}$.

So, we have two cases:

a) $E = J_p^1 X$

In this case:

$$L_0 = 0. \tag{5.3.17}$$

b) $E = \{(\tau, X, U) | \det(U_a, U_b) > 0\}$ or $E = \{(\tau, X, U) | \det(U_a, U_b) < 0\}$.

In this case:

$$L_0(\tau, X, U) = \kappa \sqrt{|\det(U_a, U_b)|}. \tag{5.3.18}$$

This completely elucidates the structure of the functions $F^{\cdots}_{\cdots}$.

The determination of the functions $G^{\cdots}_{\cdots}$ is very simple. They are of the form:

$$G^{a_1,\dots,a_k}_{\mu_0,\dots,\mu_k} = \sum_{l=k}^{p} \frac{1}{(l-k)!} C^{a_1,\dots,a_l}_{\mu_0,\dots,\mu_l} \prod_{i=k+1}^{l} U^{\mu_i}{}_{a_i} \tag{5.3.19}$$

where $C^{\cdots}_{\cdots}$ are some constants, completely antisymmetric in the upper and also in the lower indices.

The Lorentz covariance of $G^{\cdots}_{\cdots}$ is the equivalent to the Lorentz invariance of these tensors and reparametrization invariance is equivalent to the $GL(p, R)_+$-invariance of the same tensors (in the upper indices).

We have two possibilities:

A) $n \neq p + 1$.

In this case:

$$C_{\mu_0,\dots,\mu_p} = 0 \qquad \Longleftrightarrow \qquad G^{a_1,\dots,a_k}_{\mu_0,\dots,\mu_k} = 0 \qquad \forall k = 0, ..., p. \tag{5.3.20}$$

So, in this case we get:

$$\sigma = \sigma_{L_0}. \tag{5.3.21}$$

B) $n = p + 1$.
In this case:

$$C_{\mu_0,\dots,\mu_p} = \lambda \varepsilon_{\mu_0,\dots,\mu_p} \tag{5.3.22}$$

for some $\lambda \in \mathbb{R}$. So, we get for

$$G^{a_1,\dots,a_k}_{\mu_0,\dots,\mu_k}(U) = \lambda \frac{1}{(p-k)!} \varepsilon^{a_1,\dots,a_p} \varepsilon_{\mu_0,\dots,\mu_p} \prod_{i=k+1}^{p} U^{\mu_i}{}_{a_i}. \tag{5.3.23}$$

These expressions are following from a Chern-Simons Lagrangian:

$$L_{CS}(X,U) \equiv \lambda \frac{1}{(p+1)!} \varepsilon^{a_1,\dots,a_p} \varepsilon_{\mu_0,\dots,\mu_p} X^{\mu_0} \prod_{i=1}^{p} U^{\mu_i}{}_{a_i}. \tag{5.3.24}$$

So, in this case:

$$\sigma = \sigma_{L_0 + L_{CS}}. \tag{5.3.25}$$

We note that for $p = 1$ (particles) we get:

$$L_0(U) = \kappa \sqrt{\|U\|^2} \tag{5.3.26}$$

i.e. the usual form of the (homogeneous) relativistic Lagrangian. If $n = 2$ then we pick a Chern-Simons term:

$$L_{CS}(X,U) = \frac{\lambda}{2} \varepsilon_{\mu\nu} X^{\mu} U^{\nu}. \tag{5.3.27}$$

Finally for $p = 2$ (strings) we get for L_0 the well-known Nambu-Goto action:

$$L_0(U) = \kappa \sqrt{|\|U_1\|^2 \|U_2\|^2 - (U_1, U_2)^2|}. \tag{5.3.28}$$

If $n = 3$ the Chern-Simons term is:

$$L_{CS}(X,U) = \frac{\lambda}{6} \varepsilon^{a_1,a_2} \varepsilon_{\mu_0,\mu_1,\mu_2} X^{\mu_0} U^{\mu_1}{}_{a_1} U^{\mu_2}{}_{a_2}. \tag{5.3.29}$$

Acknowledgements

The conjecture that it is possible to define the Lagrange, Euler-Lagrange and Helmholtz-Sonin classes in this more general setting of non-fibred manifolds was suggested to the author by professor Demeter Krupka.

Bibliography

[1] S. J. Aldersley: Higher Euler Operators and some of their Applications *Journ. Math. Phys.* **20** (1979) 522–531

[2] I. M. Anderson: The Variational Bicomplex *Preprint* Utah State Univ. (1989)

[3] I. M. Anderson and T. Duchamp: On the Existence of Global Variational Principles *American Journ. Math.* **102** (1980) 781–868

[4] D. E. Betounes: Extensions of the Classical Cartan Form *Phys. Rev.* **D 29** (1984) 599–606

[5] D. E. Betounes: Differential Geometric Aspects of the Cartan Form: Symmetry Theory *J. Math. Phys.* **28** (1987) 2347–2353

[6] E. Cartan: *Leçons sur les Invariants Integraux* (Hermann, 1922)

[7] J. Dieudonné: *Éléments d'Analyse 3* (Gauthier-Villars, Paris, 1970)

[8] P. L. Garcia: The Poincaré-Cartan Invariant in the Calculus of Variations *Symp. Math.* **XIV** (1974) 219–246

[9] H. Goldschmits, S. Sternberg: The Hamilton-Cartan Formalism in the Calculus of Variations *Ann. Inst. Fourier* **23** (1973) 203–267

[10] M. J. Gotay: A Multisymplectic Framework for Classical Field Theory and the Calculus of Variations. I. Covariant Hamiltonian Formalism *In: Mechanics, Analysis and Geometry: 200 Years after Lagrange, M. Francaviglia and D. D. Holms, eds.* (North-Holland, Amsterdam, 1990) 203–235

[11] D. R. Grigore: A Generalized Lagrangian Formalism in Particle Mechanics and Classical Field Theory *Fortschr. der Phys.* **41** (1993) 569–617

[12] D. R. Grigore: Higher-Order Lagrangian Theories and Noetherian Symmetries *Romanian Journ. Phys.* **39** (1995) 11–35; `arXiv:hep-th/9305009`

[13] D. R. Grigore: Variational Sequence on Finite Jet Bundle Extensions and the Lagrangian Formalism *Diff. Geometry and its Applications* **10** (1999) 43–77; `arXiv:dg-ga/9702016`

[14] D. R. Grigore: Variationally Trivial Lagrangians and Locally Variational Differential Equations of Arbitrary Order *Diff. Geometry and its Applications* **10** (1999) 79–105

[15] D. R. Grigore and D. Krupka: Invariants of Velocities and Higher Order Grassmann Bundles *Journ. Geom. Phys.* **24** (1997) 244–266; `arXiv:dg-ga/9708013`

[16] D. R. Grigore: Higher-Order Lagrangian Formalism on Grassmann Manifolds *Preprint* `arXiv:dg-ga/9709005` (1995)

[17] D. R. Grigore: On a Generalisation of the Poincaré-Cartan Form to Classical Field Theory *Preprint* `arXiv:math.DG/9801073` (1998)

[18] D. R. Grigore and O. T. Popp: On the Lagrange-Souriau Form in Classical Field Theory *Mathematica Bohemica* **123** (1998) 73–86

[19] P. Horvàthy: Variational Formalism for Spinning Particles *Journ. Math. Phys.* **20** (1979) 49–52

[20] J. Klein: Espaces Variationels et Mécanique *Ann. Inst. Fourier* **12** (1962) 1–124

[21] I. Kijowski: A Finite-Dimensional Canonical Formalism in Classical Field Theory *Comm. Math. Phys.* **30** (1973) 99-128

[22] D. Krupka: A Map Associated to the Lepagean Forms of the Calculus of Variations in Fibered Manifolds *Czech. Math. Journ.* **27** (1977) 114–118

[23] D. Krupka: Lepagean Forms in Higher Order Variational Theory *In: Proceedings of the IUTAM-ISIMM Symposium on Modern Developments in Analytical Mechanics, Turin, 1982* (Atti della Academia delle Scienze di Torino, Suppl. al Vol. 117, 1983) 198–238

[24] D. Krupka: Variational Sequence on Finite Order Jet Spaces *In: Proceedings of the Conference "Differential Geometry and its Applications", August, 1989* (World Scientific, Singapore, 1990) 236–254

[25] D. Krupka: Topics in the Calculus of Variation: Finite Order Variational Sequences *In: "Differential Geometry and its Applications", proceedings Conf. Opava, 1992* (Open Univ. Press, 437–495)

[26] D. Krupka: The Contact Ideal *Diff. Geometry and its Applications* **5** (1995) 257–276

[27] L. Landau and E. L. Lifschitz: *Théorie des Champs* (Mir, Moscou, 1970)

[28] P. J. Olver: *Applications of Lie Groups to Differential Equations* (Springer, 1986)

[29] G. Manno and R. Vitolo: Variational Sequence on Finite Order Jets of Submanifolds *In: "Differential Geometry and its Applications", proceedings Conf. Opava, 2001* (Silesian Univ. at Opava, 435-446)

[30] H. Poincaré: *Leçons sur les Méthodes Nouvelles de la Mécanique Céleste* (Gauthier-Villars, Paris, 1892)

[31] H. Rund: A Cartan Form for the Field Theory of Charathéodory in the Calculus of Variations of Multiple Integrals *Lect. Notes in Pure and Appl. Math.* **100** (1985) 455–469

[32] H. Rund: Integral FormulæAssociated with the Euler-Lagrange Operator of Multiple Integral Problems in the Calculus of Variation *Æquationes Math.* **11** (1974) 212–229

[33] D. J. Saunders: An Alternative Approach to the Cartan Form in Lagrangian Field Theories *J. Phys.* **A 20** (1987) 339–349

[34] J. M. Souriau: *Structure des Systemes Dynamique* (Dunod, Paris, 1970)

D. R. Grigore

Department of Theoretical Physics,
"Horia Hulubei" National Institute of Physics and Nuclear Engineering,
Institute of Atomic Physics,
Bucharest-Măgurele, P.O. Box MG 6, România

E-mail: grigore@theory.nipne.ro
grigore@ifin.nipne.ro

This page intentionally left blank

Handbook of Global Analysis
Demeter Krupka and David Saunders

Sobolev spaces on manifolds

Emmanuel Hebey and Frédéric Robert

Contents

1 Motivations

Sobolev spaces are natural and powerful tools in nonlinear analysis and differential geometry. They are of great help in solving partial differential equations. For instance, given Ω a domain of $\mathbb{R}^n$ $(n \geq 1)$ and $f \in C^1(\mathbb{R})$, it is classical to consider the problem of finding a function $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$ such that

$$\Delta u = f'(u)$$

in Ω and $u = 0$ on $\partial\Omega$, where $\Delta = -\sum_i \partial_{ii}$. Such a function u can be seen as a critical point of the functional

$$J(u) := \frac{1}{2}\int_\Omega |\nabla u|^2\,dx - \int_\Omega f(u)\,dx.$$

When f has a reasonable growth, for instance of quadratic type, the definition of $J(u)$ makes sense as soon as $u, |\nabla u| \in L^2(\Omega)$, which is, roughly speaking, the definition of one of the most used Sobolev spaces. In several situations Sobolev spaces are naturally associated to variational problems. Their definition in the Euclidean context can be extended to the context of Riemannian manifolds. Sobolev spaces on $\mathbb{R}^n$ are well understood. Surprises and subtilities occur in the context of Riemannian manifolds. In the sequel, C denotes a positive constant, the value of which might change from line to line, and even in the same line.

2 Sobolev spaces on manifolds: definition and first properties

Let us start with few lines on the theory of Sobolev spaces in $\mathbb{R}^n$. Classical references on the subject are the books by Adams [1] and by Mazj'a [57]. Given a function $u \in L^1(\Omega)$ and given a multi-index α, we define the distributional derivative $D_\alpha u$ of u by

$$\langle D_\alpha u, \varphi\rangle := (-1)^{|\alpha|} \int_\Omega u D_\alpha \varphi \, dx$$

for all $\varphi \in C_c^\infty(\Omega)$. Here, and in what follows, $C_c^\infty(X)$ stands for the set of smooth functions with compact support in X. A distribution T is said to be in L^p if there exists $f \in L^p(\Omega)$ such that $\langle T, \varphi\rangle = \int_\Omega f\varphi \, dx$. Given $k \in \mathbb{N}$, $k \geq 1$, and $p \geq 1$, we define

$$W_k^p(\Omega) := \{u \in L^p(\Omega) / \text{ for any } |\alpha| \leq k,\ D_\alpha u \in L^p(\Omega)\}.$$

By definition, $W_k^p(\Omega)$ is the Sobolev space of order p in integrability and order k in differentiability. The space $W_k^p(\Omega)$ is a Banach space when endowed with the norm

$$\|u\|_{W_k^p(\Omega)} := \sum_{i=0}^{k} \sum_{|\alpha|=i} \|D_\alpha u\|_{L^p(\Omega)}.$$

Another possibility is to define a Sobolev space $H_k^p(\Omega)$ as the completion with respect to the above norm $\|\cdot\|_{W_k^p(\Omega)}$ of the set consisting of the functions $u \in C^\infty(\Omega)$ for which $\|u\|_{W_k^p(\Omega)} < +\infty$. By a theorem of Meyers and Serrin [58], for any integer k, and any $p \geq 1$, $W_k^p(\Omega) = H_k^p(\Omega)$. Now we turn our attention to the case of Riemannian manifolds. Possible references for the material we use in differential and Riemannian geometry are the books by Chavel [22], Do Carmo [30], Gallot, Hulin and Lafontaine [43], Hebey [46], Sakai [65], and Spivak [69]. In what follows we let (M, g) be a (smooth) Riemannian manifold of dimension $n \geq 1$. Also, for $k \in \mathbb{N}$ and $p \geq 1$, we define $\mathcal{C}_k^p(M)$ to be the set of the $u \in C^\infty(M)$ for which $\int_M |\nabla^i u|_g^p \, dv_g < +\infty$ for all $i \in \{0, ..., k\}$. Then, mimicking the above definition of $H_k^p(\Omega)$, we define the Sobolev space $H_k^p(M, g)$ as follows.

Definition 2.1 The Sobolev space $H_k^p(M, g)$ is the completion in $L^p(M)$ of $\mathcal{C}_k^p(M)$ for the norm

$$u \mapsto \|u\|_{H_k^p} := \sum_{i=0}^{k} \|\nabla^i u\|_p.$$

Here, $\|\nabla^i u\|_p$ is the L^p-norm of the function $|\nabla^i u|$ which, by definition, is the pointwise norm of the tensor $\nabla^i u$ with respect to g. The Sobolev space $\dot{H}_k^p(M, g)$ is the completion of $C_c^\infty(M)$ for the norm $\|\cdot\|_{H_k^p}$ in $H_k^p(M, g)$ or, similarly, the closure of $C_c^\infty(M)$ in $H_k^p(M, g)$.

In order to make this definition consistent, we need to prove that a Cauchy sequence for $\|\cdot\|_{H_k^p}$ that converges to 0 in $L^p(M)$ is actually converging to 0 for $\|\cdot\|_{H_k^p}$. We prove this claim for $k = 1$. The case of arbitrary $k \geq 1$ can be treated in a similar way. Let $(u_i)_{i\in\mathbb{N}} \in \mathcal{C}_1^p(M)$ a Cauchy sequence for $\|\cdot\|_{H_1^p}$. Let $x_0 \in M$ and let $\varphi : U \subset M \to \Omega \subset \mathbb{R}^n$ be a local chart at $x_0 \in U$, $\overline{U}$ compact. Then $u_i \circ \varphi^{-1} \in C^1_{loc}(\Omega)$ for all $i \in \mathbb{N}$ and we have that

$$\begin{aligned}
&\int_U |\nabla u_i - \nabla u_j|^p \, dv_g \\
&= \int_\Omega \left(g^{kl}\partial_k(u_i \circ \varphi^{-1} - u_j \circ \varphi^{-1})\partial_l(u_i \circ \varphi^{-1} - u_j \circ \varphi^{-1})\right)^{p/2} \sqrt{|g|}\, dx \\
&\geq C \int_\Omega |\nabla(u_i \circ \varphi^{-1}) - \nabla(u_j \circ \varphi^{-1})|^p \, dx
\end{aligned}$$

and we check that $(u_i \circ \varphi^{-1})_{i\in\mathbb{N}}$ is a Cauchy sequence in $W_1^p(\Omega)$. In particular, it converges in $W_1^p(\Omega)$ when $i \to +\infty$. Clearly, $(u_i \circ \varphi^{-1})_{i\in\mathbb{N}}$ converges to 0 in $L^p(\Omega)$ when $i \to +\infty$ if $u_i \to 0$ in $L^p(M)$ when $i \to +\infty$. Therefore, $u_i \circ \varphi^{-1} \to 0$ in $W_1^p(\Omega)$. Coming back to the manifold via the charts, we get that $\int_U |\nabla u_i|^p \, dv_g \to 0$ and it follows that for any compact $K \subset M$, $\int_K |\nabla u_i|^p \, dv_g \to 0$ when $i \to +\infty$. On the other hand, it follows from the triangle inequality that the sequence of functions $(|\nabla u_i|)_{i\in\mathbb{N}}$ is a Cauchy sequence in $L^p(M)$, and then, it converges to a function $f \in L^p(M)$. According to what we just proved, $\int_K |f|^p \, dv_g = 0$ for all K compact. Since M is Riemannian, it is paracompact, and thus M can be written as a countable union of compact sets. It follows that $\int_M |f|^p \, dv_g = 0$, and then that $f \equiv 0$. In particular, $u_i \to 0$ in $H_1^p(M, g)$ when $i \to +\infty$. The claim is proved.

An immediate consequence of the definition, and of the fact that L^p-spaces are reflexive spaces when $p > 1$, is the following.

Proposition 2.1 *When endowed with the norm $\|\cdot\|_{H_k^p}$, $H_k^p(M, g)$ and $\dot{H}_k^p(M, g)$ are Banach spaces. They are reflexive spaces when $p > 1$.*

Given a function $u \in H_1^p(M, g)$, it is natural to define its gradient. Let $x_0 \in M$ and let $\varphi : U \subset M \to \Omega \subset \mathbb{R}^n$ be a local chart at $x_0 \in U$. If $u \in H_1^p(M, g)$, then $u \circ \varphi^{-1} \in W_1^p(\Omega)$ and its derivatives are defined almost everywhere. We can then define $\nabla u(x)$ by the equation

$$\nabla u(x).X = \partial_k(u \circ \varphi^{-1})_{\varphi(x)} X^k$$

for almost every $x \in U$ and all $X \in T_x M$, where the X^k's are the components of X in the chart. This definition is independent of the choice of the chart and we defined $\nabla u(x)$ for almost every $x \in M$. We have that

$$\int_M |\nabla u(x)|^p \, dv_g < +\infty.$$

The construction extends to the spaces $H_k^p(M, g)$ for $k \geq 2$ since one has explicit expressions for the higher order derivatives $\nabla^i v$ for smooth v.

A natural question on Sobolev spaces is to wonder in which measure the definition depends on the metric. When M is compact, the dependancy on the metric turns out to be inexistant.

Proposition 2.2 *Let M be a compact manifold. Let g, g' be two metrics on M. Then $H_k^p(M,g) = H_k^p(M,g')$, and $\dot{H}_k^p(M,g) = \dot{H}_k^p(M,g')$.*

Proof. Given any point $x \in M$, there exists an open neighborhood U_x of x such that $c^{-1} \cdot g \le g' \le c \cdot g$ in U_x in the sense of bilinear forms. Since M is compact, there is a finite covering of such neighborhoods, and the H_k^p-norms corresponding to g and g' are equivalent. In particular, the two Sobolev spaces coincide. □

A simple but important remark is that this proposition does not extend to noncompact manifolds. For instance we may consider the two Riemannian spaces $(\mathbb{R}^n, \xi)$, where ξ is the standard Euclidean metric, and $(\mathbb{R}^n, 4(1+|x|^2)^{-2}\xi)$, which corresponds to the standard sphere after the stereographic projection. It is not difficult to see that

$$1 \in H_k^p\left(\mathbb{R}^n, \frac{4}{(1+|x|^2)^2}\xi\right) \text{ and } 1 \notin H_k^p\left(\mathbb{R}^n, \xi\right).$$

Despite the dependancy for noncompact manifolds, we often write $H_k^p(M)$ for $H_k^p(M,g)$ and $\dot{H}_k^p(M)$ for $H_k^p(M,g)$ when there is no ambiguity. In what follows we mention a few useful properties of the function space $H_1^p(M)$.

Proposition 2.3 *Let $u : M \to \mathbb{R}$ be a Lipschitz function with compact support. Then $u \in \dot{H}_1^p(M)$ for all $p \ge 1$.*

Proof. Let K be the compact support of u. The notation $B_\delta(x)$ stands for the ball centered at x of radius δ. The ball, depending on weither $x \in M$ or $x \in \mathbb{R}^n$, is either a ball in M or in the Euclidean space. Let $N \in \mathbb{N}$, let $(x_i)_{i \in \{1,\ldots,N\}} \in M$ and $\delta > 0$ be such that

$$K \subset \bigcup_{i=1}^{N} B_{\delta_i}(x_i)$$

and such that for any $i \in \{1, \ldots, N\}$, there exists a local chart

$$\varphi_i : B_{3\delta}(x_i) \subset M \to B_{3\delta}(0).$$

We let $\tilde{\eta} \in C^\infty(\mathbb{R}^n)$ be such that $\tilde{\eta} \equiv 1$ in $B_\delta(0)$ and $\tilde{\eta} \equiv 0$ in $\mathbb{R}^n \setminus B_{2\delta}(0)$. For any $i \in \{1, ..., N\}$, we let $\tilde{\eta}_i = \tilde{\eta} \circ \varphi_i$ and

$$\eta_i = \frac{\tilde{\eta}_i}{\sum_j \tilde{\eta}_j} \text{ in } \bigcup_{i=1}^{N} B_{\delta_i}(x_i).$$

Clearly $\sum_i \eta_i = 1$ in a neighborhood of K, and $u = \sum \eta_i u$ makes sense in M. For any $i \in \{1, \ldots, N\}$, we let $v_i = (\eta_i u) \circ \varphi_i^{-1}$. Clearly, v_i has compact support in $B_{2\delta}(0)$ and is Lipschitz continuous. It follows from standard theory of Sobolev spaces in Euclidean space that $v_i \in W_1^p(B_{2\delta}(0))$, and therefore, since $H_k^p(\Omega) = W_k^p(\Omega)$ for open subsets of $\mathbb{R}^n$, that $v_i \in H_1^p(B_{2\delta}(0))$. Since v_i has compact support in $B_{2\delta}(0)$, we get that $v_i \in \dot{H}_1^p(B_{2\delta}(0))$. Then, coming back to the initial manifold, we get that $\eta_i u = v_i \circ \varphi_i \in \dot{H}_1^p(B_{3\delta}(x_i)) \subset \dot{H}_1^p(M)$. Since $u = \sum \eta_i u$, we it follows that $u \in \dot{H}_1^p(M)$. □

In the same spirit, namely by using local charts to come back to the corresponding result in $\mathbb{R}^n$, we get that the following proposition holds true.

Proposition 2.4 *Let (M,g) be a complete manifold. Let $u \in H_1^p(M)$ and let $h : \mathbb{R} \to \mathbb{R}$ be a Lipschitz function. Assume that $h \circ u \in L^p(M)$, $p \geq 1$. Then $h \circ u \in H_1^p(M)$ and*

$$|(\nabla(h \circ u))|(x) = |h'(u(x))| \cdot |(\nabla u)|(x)$$

for almost every $x \in M$. In particular, $|u| \in H_1^p(M)$ and

$$|(\nabla |u|)|(x) = |(\nabla u)|(x)$$

for almost every $x \in M$.

Note that this latest assertion is not true for Sobolev spaces of order greater than one. We easily find examples of functions $u \in H_2^p(\mathbb{R}^n)$ such that $|u| \notin H_2^p(\mathbb{R}^n)$.

Proof. Let $u \in H_1^p(M)$ and $v = h \circ u$. As a preliminary remark, note that $v \in L^p(M)$ is given for free when $Vol_g(M) < +\infty$, and is equivalent to $h(0) = 0$ when $Vol_g(M) = +\infty$. Given $x_0 \in M$ and $\varphi : U \to \Omega$ a local chart at x_0 as in the proof of Proposition 2.3 (in particular, $\overline{\Omega}$ is compact), we get that $u \circ \varphi^{-1} \in W_1^p(\Omega)$. It follows from standard theory of Sobolev spaces in Euclidean space that $v \circ \varphi^{-1} = h \circ (u \circ \varphi^{-1}) \in W_1^p(\Omega) = H_1^p(\Omega)$ and that

$$\nabla(v \circ \varphi^{-1})(x) = h'(u \circ \varphi^{-1}(x)) \cdot \nabla\left(u \circ \varphi^{-1}\right)(x)$$

for almost every $x \in \Omega$ (note that it follows from Rademacher's Theorem that a Lipschitz function is differentiable almost everywhere, so that $h'(y)$ makes sense for almost every $y \in \mathbb{R}$). Coming back to the initial manifold, we get that $v \in H_1^p(U)$ and that

$$|\nabla v|(x) = |h'(v(x))| \cdot |\nabla u|(x) \tag{1}$$

for almost every $x \in U$. By a covering argument, we get that for any $V \subset M$ relatively compact, then $v \in H_1^p(V)$ and (1) holds for almost every $x \in V$. Proposition 2.4 is then proved when v has compact support. Now we consider the general case. We let $\eta \in C^\infty(\mathbb{R})$ such that $\eta \equiv 1$ on $(-\infty, 0]$ and $\eta \equiv 0$ on $[1, +\infty)$. We let $x_0 \in M$ and we let $\eta_i(x) = \eta(d_g(x, x_0) - i)$ for all $i \in \mathbb{N}$ and all $x \in M$. Since M is complete, we get that η_i is Lipschitz continuous and has compact support. Then $\eta_i \in \dot{H}_1^q(M)$ for all $q \geq 1$. Since $v \in H_1^p(V)$ for all $V \subset\subset M$, we get that $\eta_i v \in H_1^p(M)$ for all $i \in \mathbb{N}$. Since $v \in L^p(M)$, it follows from Lebesgue's convergence theorem that

$$\lim_{i \to +\infty} \eta_i v = v \text{ in } L^p(M).$$

Given $i < j$, we get that

$$\begin{aligned} \int_M |\nabla(\eta_i v) - \nabla(\eta_j v)|^p \, dv_g &= \int_M |(\eta_i - \eta_j)\nabla v + v \nabla(\eta_i - \eta_j)|^p \, dv_g \\ &\leq c \int_{\{d_g(x,x_0) \geq i\}} (|\nabla u|^p + |v|^p) \, dv_g. \end{aligned}$$

It then follows from Lebesgue's convergence theorem that $(\eta_i v)_{i \in \mathbb{N}}$ is a Cauchy sequence for the H_1^p-norm. Since it converges to v in $L^p(M)$, we actually proved that $v \in H_1^p(M)$. $\square$

3 Equality and density issues

Let $k \in \mathbb{N}$ and $p \geq 1$. When the manifold M is compact, it is clear that $\mathcal{C}_k^p(M) = C^\infty(M) = C_c^\infty(M)$. Therefore, $H_k^p(M) = \dot{H}_k^p(M)$ when M is compact. This equality does not hold for arbitrary domains. Let Ω be an open bounded subset of $\mathbb{R}^n$: we have that

$$\dot{H}_k^p(\Omega) \subset H_k^p(\Omega) \text{ and } 1 \in H_k^p(\Omega) \setminus \dot{H}_k^p(\Omega)$$

when $k \geq 1$. Concerning the whole space $\mathbb{R}^n$, one can write that $H_k^p(\mathbb{R}^n) = \dot{H}_k^p(\mathbb{R}^n)$ for all $p \geq 1$ and all $k \in \mathbb{N}$. When $k = 1$, and we are dealing with complete manifolds, Aubin [5] proved the following result.

Theorem 3.1 *Let (M, g) be a complete Riemannian manifold and let $p \geq 1$. Then $H_1^p(M) = \dot{H}_1^p(M)$.*

Proof. Let $u \in C_1^p(M)$. We claim that there exists a sequence $(u_i)_{i\in\mathbb{N}} \in \dot{H}_1^p(M)$ such that $\lim_{i\to+\infty} u_i = u$ in $H_1^p(M)$. Clearly the theorem is equivalent to this claim. Let $\eta \in C^\infty(\mathbb{R})$ such that $\eta(x) = 1$ for $x \leq 0$ and $\eta(x) = 0$ for $x \geq 1$. For any $i \in \mathbb{N}$, we let $u_i(x) = \eta(d_g(x, x_0) - i)u(x)$ for all $x \in M$ and all $i \in \mathbb{N}$. As easily checked, u_i is Lipschitz continuous and u_i has compact support in M (this latest assertion holds since M is complete). It follows from Proposition 2.3 that $u_i \in \dot{H}_1^p(M)$. We get that

$$\|u - u_i\|_p = \left(\int_M |\eta(d_g(x, x_0) - i) - 1|^p |u|^p \, dv_g\right)^{1/p} = o(1) \tag{2}$$

when $i \to +\infty$. Concerning the first order derivative, one gets that

$$\begin{aligned}\|\nabla(u - u_i)\|_p \quad &\leq \quad C\left(\int_M |\eta(d_g(x, x_0) - i) - 1|^p |\nabla u|^p \, dv_g\right)^{1/p} \\ &\quad + C\left(\int_M |\nabla(\eta(d_g(x, x_0) - i) - 1)|^p |u|^p \, dv_g\right)^{1/p}.\end{aligned}$$

With Proposition 2.4, we get that

$$|\nabla(\eta(d_g(x, x_0) - i) - 1)| = |\eta'(d_g(x, x_0) - i)| \leq C\mathbf{1}_{d_g(x,x_0)\geq i}$$

for all $x \in M$ and all $i \in \mathbb{N}$. Since $|u|, |\nabla u| \in L^p(M)$, we get that $\|\nabla(u - u_i)\|_p = o(1)$ when $i \to +\infty$. This latest result and (2) yield

$$\lim_{i\to+\infty} u_i = u \text{ in } H_1^p(M),$$

and the claim is proved. As already mentioned, this proves the theorem. □

Concerning the spaces $H_k^p(M)$ and $\dot{H}_k^p(M)$ with $k \geq 2$, the above proof does not work, and the situation is more intricate. Equality results require assumptions on the Ricci tensor of (M, g). For instance, the following proposition of Hebey [47] holds true.

Proposition 3.1 *Let (M, g) be a complete Riemannian manifold with positive injectivity radius and Ricci curvature bounded from below. Then $\dot{H}_2^2(M) = H_2^2(M)$.*

The proof of this proposition uses the existence of harmonic coordinates. We refer to [47] for more results in this direction. If one wants to avoid geometric assumptions on the manifold, and get general results as in the Euclidean case, it follows from the results in [47] that $\dot{H}_k^p(M) = H_k^p(M)$ for any k integer, and any $p \geq 1$, when (M, g) is a Riemannian covering of a compact Riemannian manifolds.

4 Embedding theorems, Part I

In his paper [68], Sobolev proved that a function in L^p with derivatives in L^p is actually in L^q, for some $q > p$. The result is now refered to as Sobolev's theorem. The original statement concerned Euclidean spaces. However, such an embedding makes sense on a manifold. The question of the validity of the embedding leads to surprising results when the manifold is noncompact.

Definition 4.1 Let (M, g) be a manifold. We we say that the the Sobolev embedding theorem in its first part holds true on (M, g) if for any real numbers $1 \leq p < q$, and any integers $0 \leq m < k$, we have that $H_k^p(M) \subset H_m^q(M)$ and that the embedding is continuous as soon as $\frac{1}{q} = \frac{1}{p} - \frac{k-m}{n}$. Assuming that $kp < n$, we say that the Sobolev embedding $(\mathcal{S}_{H_k^p(M)})$ holds true if $H_k^p(M) \subset L^q(M)$ and the embedding is continuous as soon as $\frac{1}{q} = \frac{1}{p} - \frac{k}{n}$.

By definition, if $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ are two Banach spaces, and $E \subset F$, we say that the embedding $E \subset F$ is continuous if there exists $C > 0$ such that

$$\|u\|_F \leq C\|u\|_E$$

for all $u \in E$. In the sequel, the notation $E \subset F$ refers to continuous embeddings. A general very useful result is that there is an ordering in the embeddings. In particular, all the Sobolev inequalities can be reduced to the proof of one.

Theorem 4.1 *Let (M, g) be a complete manifold of dimension $n \geq 2$. Assume that $(\mathcal{S}_{H_1^1(M)})$ holds true. Then for any real numbers $1 \leq p < q$, and any integers $0 \leq m < k$ such that $\frac{1}{q} = \frac{1}{p} - \frac{k-m}{n}$, we have that $H_k^p(M) \subset H_m^q(M)$. In particular, the Sobolev embedding theorem in its first part holds true on (M, g).*

Proof. We prove the results for $k = 1$ and the embeddings $(\mathcal{S}_{H_1^p(M)})$. We refer to Aubin [9] for the other cases. We assume that there exists $C > 0$ such that

$$\left(\int_M |u|^{\frac{n}{n-1}}\, dv_g\right)^{\frac{n-1}{n}} \leq C \int_M (|\nabla u| + |u|)\, dv_g \tag{3}$$

for all $u \in H_1^1(M)$. We let $u \in C_c^\infty(M)$, and we consider $v = |u|^{\frac{p(n-1)}{n-p}}$. It follows from Proposition 2.3 that $v \in H_1^p(M)$. By Hölder's inequality we can write that

$$\begin{aligned} \|v\|_1 &= \int_M |u|^{\frac{1}{p}\cdot p} \cdot |u|^{(1-\frac{1}{p})\cdot\frac{pn}{n-p}} \\ &\leq \|u\|_p \cdot \|u\|_{\frac{pn}{n-p}}^{\frac{n(p-1)}{n-p}} \end{aligned}$$

and

$$\begin{aligned}\|\nabla v\|_1 &= \int_M |\nabla v|\, dv_g = \frac{p(n-1)}{n-p}\int_M |u|^{\frac{n(p-1)}{n-p}} |\nabla u|\, dv_g \\ &\leq \frac{p(n-1)}{n-p}\|u\|_{\frac{pn}{n-p}}^{\frac{n(p-1)}{n-p}} \cdot \|\nabla u\|_p.\end{aligned}$$

We apply (3) to v, and we get that

$$\|u\|_{\frac{pn}{n-p}}^{\frac{p(n-1)}{n-p}} \leq C\|u\|_{\frac{pn}{n-p}}^{\frac{n(p-1)}{n-p}} \cdot \|\nabla u\|_p + C\|u\|_{\frac{pn}{n-p}}^{\frac{n(p-1)}{n-p}} \cdot \|u\|_p$$

and then

$$\|u\|_{\frac{pn}{n-p}} \leq C\|\nabla u\|_p + C\|u\|_p$$

for all $u \in C_c^\infty(M)$. By density, this inequality holds for $u \in \dot{H}_1^p(M)$, and then on $H_1^p(M)$ by Theorem 3.1. □

4.1 Discussion on the exponent

A natural question is: what about the "magical" exponent $q > p$? A justification is easy to give in terms of rescaling arguments. Consider the Euclidean space $\mathbb{R}^n$ and assume that there exists $C > 0$ such that

$$\|u\|_q \leq C\|\nabla^k u\|_p \tag{4}$$

for all $u \in C_c^\infty(\mathbb{R}^n)$. Let $u \in C_c^\infty(\mathbb{R}^n)$ be nonzero, $\lambda \in \mathbb{R}$ be positive, and define $v(x) = u(\lambda x)$ for all $x \in \mathbb{R}^n$. Plugging v into (4), we get that

$$\|u\|_q \leq C\lambda^{k-\frac{n}{p}+\frac{n}{q}}\|\nabla^k u\|_p$$

Letting λ go to 0 on the one hand, and to $+\infty$ on the other hand, it follows that we need to have that $k - \frac{n}{p} + \frac{n}{q} = 0$. In other words $\frac{1}{q} = \frac{1}{p} - \frac{k}{n}$, and this is the Sobolev exponent. The following result holds true.

Proposition 4.1 *Let (M, g) be a Riemannian manifold of dimension n. Let $p \geq 1$ and $k \in \mathbb{N}$ such that $p < n$. Let $r > 1$ such that $\frac{1}{r} < \frac{1}{p} - \frac{1}{n}$. Then $H_1^p(M)$ is not continuously embedded in $L^r(M)$.*

Proof. We argue by contradiction and assume that there is a continuous embedding of $H_1^p(M)$ in $L^r(M)$. Then there exists $C > 0$ such that

$$\|u\|_r \leq C\|u\|_{H_1^p} \tag{5}$$

for all $u \in H_1^p(M)$. We let $u \in C_c^\infty(\mathbb{R}^n) \setminus \{0\}$ and $R > 0$ be such that supp $u \subset B_R(0)$. Let $x_0 \in M$. Given $\epsilon > 0$, we define

$$\varphi_\epsilon(x) = \epsilon^{1-n/p} u(\epsilon^{-1} \exp_{x_0}^{-1}(x))$$

for $d_g(x, x_0) < i_g(x_0)$ and 0 elsewhere, where $i_g(x_0)$ denotes the injectivity radius at x_0. Clearly $\varphi_\epsilon \in C_c^\infty(M)$ for all $\epsilon > 0$. Working in the exponential chart, one gets that

$$\int_M |\nabla \varphi_\epsilon|^p \, dv_g = \int_{\mathbb{R}^n} \left(g^{ij}(\epsilon x) \partial_i u \partial_j u \right)^{p/2} \sqrt{|g|(\epsilon x)} \, dx$$

and

$$\int_M |\varphi_\epsilon|^s \, dv_g = \epsilon^{n - s \frac{n-p}{p}} \int_{\mathbb{R}^n} |u|^s \sqrt{|g|(\epsilon x)} \, dx$$

for all $s \geq 1$. Letting $\epsilon \to 0$ and plugging these estimates into (5), we get that $\|u\|_q = O\left(\epsilon^{nr(1/p - 1/n - 1/r)}\right)$ when $\epsilon \to 0$, and then $u \equiv 0$. A contradiction. This proves the proposition. □

4.2 The Euclidean setting

In the Euclidean setting, things go for the best. The Sobolev embeddings are all valid and the following fundamental theorem, due to Sobolev [68], holds true.

Theorem 4.2 *For any real numbers $1 \leq p < q$, and any integers $0 \leq m < k$ such that $\frac{1}{q} = \frac{1}{p} - \frac{k-m}{n}$, we have that $H_k^p(\mathbb{R}^n) \subset H_m^q(\mathbb{R}^n)$. In particular, given $n \geq 2$, $k \in \mathbb{N}$, and $p \geq 1$ such that $kp < n$, the embedding $(\mathcal{S}_{H_k^p(\mathbb{R}^n)})$ holds true and there exists $C_{n,k,p} > 0$ such that*

$$\left(\int_{\mathbb{R}^n} |u|^{\frac{pn}{n-kp}} \right)^{\frac{n-kp}{n}} \leq C_{n,k,p} \int_{\mathbb{R}^n} |\nabla^k u|^p \, dx$$

for all $u \in H_k^p(\mathbb{R}^n)$.

Proof. We present here the very elegant proof by Gagliardo [42] and Nirenberg [62]. Let $u \in C_c^\infty(\mathbb{R}^n)$. Any point $x \in \mathbb{R}^n$ is written $x = (x_1, ..., x_n)$. We have that

$$u(x) = \int_{-\infty}^{x_i} \partial_i u(x_1, ..., x_{i-1}, t, x_{i+1}, ..., x_n) \, dt$$

for all $x \in \mathbb{R}^n$ and all $i \in \{1, ..., n\}$, and then

$$|u(x)| \leq \int_{\mathbb{R}} |\nabla u|(..., x_{i-1}, t, x_{i+1}, ...) \, dt$$

for all $x \in \mathbb{R}^n$ and all $i \in \{1, ..., n\}$. Therefore, we get that

$$|u(x)|^{\frac{n}{n-1}} \leq \Pi_{i=1}^n \left(\int_{\mathbb{R}} |\nabla u|(..., x_{i-1}, t, x_{i+1}, ...) \, dt \right)^{\frac{1}{n-1}}$$

for all $x \in \mathbb{R}^n$. Integrating with respect to $x_n \in \mathbb{R}$ and using Hölder's inequality, we get that

$$\int_{x_n \in \mathbb{R}} |u(x)|^{\frac{n}{n-1}} \, dx_n$$

$$\leq \int_{x_n\in\mathbb{R}} \Pi_{i=1}^n \left(\int_{\mathbb{R}} |\nabla u|(..., x_{i-1}, t, x_{i+1}, ...)\, dt\right)^{\frac{1}{n-1}}$$

$$\leq \left(\int_{\mathbb{R}} |\nabla u|(x_1, ..., x_{n-1}, t)\, dt\right)^{\frac{1}{n-1}} \Pi_{i=1}^{n-1} \left(\int_{(x_i,x_n)\in\mathbb{R}^2} |\nabla u|(x)\, dx_i dx_n\right)^{\frac{1}{n-1}} .$$

With the same method, integrating with respect to $x_{n-1}, ..., x_1$, we get that

$$\int_{\mathbb{R}^n} |u(x)|^{\frac{n}{n-1}}\, dx \leq c \left(\int_{\mathbb{R}^n} |\nabla u|\, dx\right)^{\frac{n}{n-1}}$$

for all $u \in C_c^\infty(\mathbb{R}^n)$. Since $\mathbb{R}^n$ with the Euclidean metric is obviously a complete manifold, the theorem follows from Theorems 3.1 and 4.1. □

4.3 The compact setting

In some sense a compact manifold is just a finite union of small pieces of $\mathbb{R}^n$. From this, and since Sobolev embeddings are true in Euclidean space, we expect that Sobolev embeddings are also true on compact manifolds. This is indeed the case and the following theorem holds true.

Theorem 4.3 *Let (M, g) be a compact manifold. For any real numbers $1 \leq p < q$, and any integers $0 \leq m < k$ such that $\frac{1}{q} = \frac{1}{p} - \frac{k-m}{n}$, we have that $H_k^p(\mathbb{R}^n) \subset H_m^q(\mathbb{R}^n)$. In particular, the Sobolev embedding theorem in its first part holds true on (M, g).*

Proof. It follows from Theorem 4.1 above that it is sufficient to prove that the embedding $(\mathcal{S}_{H_k^p(M)})$ holds true for $k = p = 1$. Since M is compact, it can be covered by a finite number of charts $(\Omega_m, \varphi_m)_{m=1,...,N}$ such that for any m, the components g_{ij}^m of g in (Ω_m, φ_m) satisfy

$$\frac{1}{2}\xi_{ij} \leq g_{ij}^m \leq 2\xi_{ij}$$

in the sense of bilinear forms. Let (η_m) be a smooth partition of unity subordinate to the covering (Ω_m). For any $u \in C^\infty(M)$ and any m, one has that

$$\int_M |\eta_m u|^{\frac{n}{n-1}}\, dv_g \leq 2^{n/2} \int_{\mathbb{R}^n} |(\eta_m u) \circ \varphi_m^{-1}|^{\frac{n}{n-1}}\, dx$$

and

$$\int_M |\nabla(\eta_m u)|\, dv_g \geq 2^{-(n+1)/2} \int_{\mathbb{R}^n} |\nabla((\eta_m u) \circ \varphi_m^{-1})(x)|\, dx.$$

Independently, by Theorem 4.2, we have that

$$\left(\int_{\mathbb{R}^n} |(\eta_m u) \circ \varphi_m^{-1}|^{\frac{n}{n-1}}\, dx\right)^{\frac{n-1}{n}} \leq C \int_{\mathbb{R}^n} |\nabla((\eta_m u) \circ \varphi_m^{-1})(x)|\, dx$$

for all m. Therefore, we get that

$$\left(\int_M |u|^{\frac{n}{n-1}}\, dv_g\right)^{\frac{n-1}{n}} \quad = \quad \left(\int_M |\sum_m \eta_m u|^{\frac{n}{n-1}}\, dv_g\right)^{\frac{n-1}{n}}$$

$$\leq \sum_{m=1}^{N} \left(\int_M |\eta_m u|^{\frac{n}{n-1}} \, dv_g \right)^{\frac{n-1}{n}}$$
$$\leq C \sum_{m=1}^{N} \int_M |\nabla(\eta_m u)| \, dv_g$$
$$\leq C \int_M |\nabla u| \, dv_g + C \left(\max_M \sum_{i=1}^{N} |\nabla \eta_m|_g \right) \int_M |u| \, dv_g .$$

Hence, there exists $A > 0$ such that

$$\left(\int_M |u|^{\frac{n}{n-1}} \, dv_g \right)^{\frac{n-1}{n}} \leq A \left(\int_M |\nabla u| \, dv_g + \int_M |u| \, dv_g \right)$$

for all $u \in C^\infty(M)$. By density, this inequality holds true for all $u \in H_1^1(M)$, and we get that $(\mathcal{S}_{H_1^1(M)})$ holds true. By Theorem 4.1 the Sobolev embedding theorem in its first part holds true on (M, g). □

In particular, it follows from Theorem 4.3 that the embeddings $(\mathcal{S}_{H_k^p(M)})$ hold true. On a compact manifolds, $L^{q'}(M) \subset L^q(M)$ when $q' \leq q$. It follows that on compact manifolds, $H_k^p(M) \subset L^{q'}(M)$ as soon as $q' \leq np/(n - kp)$.

4.4 Results in the noncompact setting

A possible model for complete noncompact manifolds is the Euclidean space $(\mathbb{R}^n, \xi)$, a manifold for which all the Sobolev embeddings are valid. The model is misleading. The picture in terms of Sobolev embeddings turns out be more tricky for arbitrary complete noncompact manifolds with nontrivial geometries at infinity. In particular, the following result due to Carron [19] holds true. We refer also to, Chavel [22] for the case $k = p = 1$.

Proposition 4.2 *Let (M, g) be a complete Riemannian manifold of dimension n. We assume that $(\mathcal{S}_{H_1^p(M)})$ holds true for some $p \geq 1$. Then for any $r > 0$, there exists $v(r) > 0$ such that*

$$Vol_g(B_r(x)) \geq v(r)$$

for all $x \in M$.

The result is actually more precise and we get that there exists a constant $c > 0$ such that

$$\mathrm{Vol}_g(B_r(x)) \geq c \inf\{1, r^n\}$$

for all $x \in M$ and all $r > 0$.

Proof. Since $(\mathcal{S}_{H_1^p(M)})$ holds true, there exists $A > 0$ such that

$$\|u\|_q \leq A \|\nabla u\|_p + A \|u\|_p \tag{6}$$

for all $u \in H_1^p(M)$, where $q > 1$ is such that $\frac{1}{q} = \frac{1}{p} - \frac{1}{n}$. We fix $x \in M$ and $r > 0$ and we define the function u as follows: $u(y) = r - d_g(x, y)$ if $d_g(x, y) \leq r$ and $u(y) = 0$

elsewhere. Since M is complete, it follows from Propositions 2.3 and 2.4 that $u \in H_1^p(M)$ and we have that $|\nabla u|(x) = 1$ for a.e. $x \in M$. Plugging u into (6) and using Hölder's inequality, we get that

$$\|u\|_q \leq A|B_r(x)|^{\frac{1}{p}} + A|B_r(x)|^{\frac{1}{n}}\|u\|_q, \tag{7}$$

where we use the notation $|B_r(x)| = \mathrm{Vol}_g(B_r(x))$. We distinguish two cases:

Case 1: $A|B_r(x)|^{\frac{1}{n}} \geq \frac{1}{2}$. In this case, we get that

$$|B_r(x)| \geq \left(\frac{1}{2A}\right)^n. \tag{8}$$

Case 2: $A|B_r(x)|^{\frac{1}{n}} \leq \frac{1}{2}$. In this case, (7) yields

$$\|u\|_p \leq |B_r(x)|^{\frac{1}{n}}\|u\|_q \leq 2A|B_r(x)|^{\frac{1}{p}+\frac{1}{n}}. \tag{9}$$

Moreover, we have that

$$\|u\|_p = \left(\int_M |u|^p \, dv_g\right)^{1/p} \geq \left(\int_{B_{r/2}(x)} |u|^p \, dv_g\right)^{1/p} = \frac{r}{2}|B_{r/2}(x)|^{\frac{1}{p}}, \tag{10}$$

and plugging (10) into (9), we get that

$$r|B_{r/2}(x)|^{\frac{1}{p}} \leq 4A|B_r(x)|^{\frac{1}{p}+\frac{1}{n}}, \tag{11}$$

and therefore

$$|B_r(x)| \geq \left(\frac{r}{4A}\right)^{\frac{np}{n+p}} |B_{r/2}(x)|^{\frac{n}{n+p}}.$$

With an induction argument, we get that

$$|B_r(x)| \geq \left(\frac{r}{A}\right)^{n(1-\theta^N)} \frac{1}{2^{p\sum_{i=1}^{N}(i+1)\theta^i}} |B_{2^{-N}r}(x)|^{\theta^N} \tag{12}$$

for all $N \in \mathbb{N}$, where $\theta := n/(n+p)$. Since

$$\lim_{r \to 0} \frac{|B_r(x)|}{r^n} = \omega_{n-1},$$

letting $N \to +\infty$ in (12) yields that there exists $\lambda = \lambda(n,p) > 0$ depending only on n, p such that

$$|B_r(x)| \geq \lambda \left(\frac{r}{A}\right)^n.$$

By cases 1 and 2 we get that the proposition holds true. □

With the above result it is quite easy to construct complete manifolds for which the Sobolev embeddings are not valid.

Proposition 4.3 *Let $M = \mathbb{R} \times \mathbb{S}^{n-1}$ endowed with the warped product metric $g(x, \theta) = dx^2 + u(x)h_\theta$, where h is the standard metric on $\mathbb{S}^{n-1}$ and $u(x) = e^{-x^2}$ for $x \in \mathbb{R}$. Then (M, g) is a complete Riemannian manifold on which none of the Sobolev inequality $(\mathcal{S}_{H_1^p(M)})$ is valid. In particular, there exist complete manifolds of arbitrary dimension such that none of the Sobolev inequalities $(\mathcal{S}_{H_1^p(M)})$ holds true.*

Proof. Given $y = (x_1, \theta_1)$, $z = (x_2, \theta_2) \in M$, we have that $|x_1 - x_2| \le d_g(y, z)$. This implies that M is complete. We let $y = (x, \theta) \in M$, where $x \ge 1$. Then $B_1(y) \subset (x-1, x+1) \times \mathbb{S}^{n-1}$. Therefore, we get that

$$\mathrm{Vol}_g(B_1(x,\theta)) \le \mathrm{Vol}_g((x-1, x+1) \times \mathbb{S}^{n-1}) \le \omega_{n-1} e^{-\frac{n-1}{2}(x-1)^2}, \tag{13}$$

where ω_{n-1} denotes the volume of $\mathbb{S}^{n-1}$. It follows that for any $\theta \in \mathbb{S}^{n-1}$, we have that

$$\lim_{x \to +\infty} \mathrm{Vol}_g(B_1(x,\theta)) = 0,$$

and it follows from Proposition 4.2 that none of the Sobolev inequalities $(\mathcal{S}_{H_1^p(M)})$ holds true. ☐

A natural question is to ask about minimal hypothesis to recover the Sobolev inequalities on complete noncompact manifolds. Fairly general results on the question can be obtained. We mention only the following result due to Varopoulos [73].

Theorem 4.4 *Let (M, g) be a complete Riemannian manifold of dimension n. We assume that the Ricci curvature is bounded from below and that there exists $v > 0$ such that*

$$\mathit{Vol}_g(B_1(x)) \ge v$$

for all $x \in M$. Then for any real numbers $1 \le p < q$, and any integers $0 \le m < k$ such that $\frac{1}{q} = \frac{1}{p} - \frac{k-m}{n}$, we have that $H_k^p(M) \subset H_m^q(M)$. In particular, the Sobolev embedding theorem in its first part holds true on (M, g).

We refer to Coulhon and Saloff-Coste [27] for complementary results. For an exposition in book form of such results we refer to Hebey [47] and Saloff-Coste [66].

4.5 The Nash inequality

Many inequalities can be derived from the generic Sobolev inequalities. This is the case for the so-called Gagliardo-Nirenberg inequalities, which we can write as

$$\|u\|_r \le C \|\nabla u\|_q^\alpha \|u\|_s^{1-\alpha}$$

for particular values of r, q, s, and α. We restrict ourselves here to one of these inequalities, refered to as the Nash inequality. The inequality first appeared in the celebrated Nash [61] when discussing the Hölder regularity of solutions of divergence form of uniformly elliptic equations. The Nash inequality in Euclidean space asserts that there exists a positive constant A such that for any function $u \in C_c^\infty(\mathbb{R}^n)$,

$$\left(\int_{\mathbb{R}^n} u^2 dx\right)^{1+\frac{2}{n}} \le A \int_{\mathbb{R}^n} |\nabla u|^2 dx \left(\int_{\mathbb{R}^n} |u| dx\right)^{\frac{4}{n}}.$$

The inequality has a transcription to Riemannian manifolds. In the case of compact manifolds (M,g) the inequality holds with an additional term and we get that

$$\Big(\int_M u^2 dv_g\Big)^{1+\frac{2}{n}} \le A\int_M |\nabla u|^2 dv_g \Big(\int_M |u| dv_g\Big)^{\frac{4}{n}} + B\Big(\int_M |u| dv_g\Big)^{2+\frac{4}{n}}$$

for all $u \in H_1^2(M)$, where A and B are positive constant independent of u. We refer the reader to the original papers by Gagliardo [42] and Nirenberg [62], and to the exhaustive paper [13] by Bakry, Coulhon, Ledoux, and Saloff-Coste for more material on the subject.

5 Euclidean type inequalities

Let (M,g) be a complete Riemannian n-manifold of infinite volume, and let $p \in [1,n)$ real. We say that the Euclidean-type Sobolev inequality of order p is valid if there exists $C_p > 0$ real such that for any $u \in H_1^p(M)$,

$$\left(\int_M |u|^q \, dv_g\right)^{p/q} \le C_p \int_M |\nabla u|^p \, dv_g \ , \tag{14}$$

where $1/q = 1/p - 1/n$. We know that such an inequality holds true for the Euclidean space. In the case of an arbitrary complete Riemannian n-manifold of infinite volume, similar arguments to the ones used in the proof of Theorem 4.1 give that if (14) holds true for $p = 1$, and thus $q = n/(n-1)$, then it also holds true for all $p \in [1,n)$. By Hoffman and Spruck [51], (14) with $p = 1$ holds true on any complete simply connected Riemannian manifold of nonpositive sectional curvature. It follows that for $p \in [1,n)$ arbitrary, (14) holds true on such manifolds. In general, the question of the validity of (14) is closely related to the nonparabolicity of the manifold. Let (M,g) be a complete, noncompact Riemannian manifold and let x be some point of M. One can prove that, uniformly with respect to x, either there exist positive Green functions of pole x, and in particular there exists a positive minimal Green's function of pole x, or there does not exist any positive Green function of pole x. More precisely, let $\Omega \subset M$ be such that $x \in \Omega$ and let G be the solution of $\Delta_g G = \delta_x$ in Ω, and $G = 0$ on $\partial\Omega$. Set $G_x^\Omega(y) = G(y)$ when $y \in \Omega$, $G_x^\Omega(y) = 0$ otherwise. Obviously, $G_x^\Omega \le G_x^{\Omega'}$ if $\Omega \subset \Omega'$. Set $G_x(y) = \sup_{\{\Omega \text{ s.t. } x\in\Omega\}} G_x^\Omega(y)$, $y \in M$. One then has that either $G_x(y) = +\infty$ for all $y \in M$, or $G_x(y) < +\infty$ for all $y \in M\backslash\{x\}$. Moreover, the alternative does not depend on x and in case G_x exists, it is the positive minimal Green function of pole x. When G_x does not exist (is not finite), the manifold is said to be parabolic; otherwise the manifold is said to be nonparabolic. By Cheng-Yau [23], one has that if for some $x \in M$,

$$\liminf_{r\to+\infty} \frac{Vol_g\left(B_r(x)\right)}{r^2} < +\infty$$

then (M,g) is parabolic. This explains, for instance, why $\mathbb{R}^2$ is parabolic while $\mathbb{R}^3$ is not. More results are in Grigor'yan [45] and Varopoulos [74]. Returning to (14), the following result can be proved.

Theorem 5.1 *Let (M,g) be a complete Riemannian n-manifold of infinite volume, $n \ge 3$. If (M,g) has nonnegative Ricci curvature, then (14) is true if and only if (M,g) is* *nonparabolic and there exists $K > 0$ such that for any $x \in M$ and any $t > 0$,*

$$Vol_g\big(\{y \in M \text{ s.t. } G_x(y) > t\}\big) \le Kt^{-n/(n-2)}$$

where G_x is the positive minimal Green function of pole x.

We refer to Carron [20], Coulhon-Ledoux [26], and Varopoulos [75] for more details and complements in this direction.

6 Embedding theorems, Part II

We focus on the case $kp > n$ and refer to the Sobolev embedding theorem in its second part. In this case, the order of differentiability or the order of integrability is so large that the Sobolev space can be embedded in Hölder spaces. These results have their origins in the contributions of Sobolev. Following standard notations, we let $C^k(M)$ be the space of k times continuously differentiable functions on M. We define $C^k_B(M)$ the space of functions $u \in C^k(M)$ for which

$$\|u\|_{C^k(M)} = \sum_{i=1}^{k} \|\nabla^i u\|_\infty$$

is finite. The norm $\|\cdot\|_{C^k(M)}$ induces a structure of Banach space on the space $C^k_B(M)$. Note that in case M is compact, we have that $C^k_B(M) = C^k(M)$. Given $\alpha \in [0,1]$, we let $C^{0,\alpha}(M)$ be the space of functions $u \in C^0(M)$ such that there exists $C > 0$ such that $|u(x) - u(y)| \le C d_g(x,y)^\alpha$ for all $x, y \in M$. In the same spirit, we define $C^{0,\alpha}_B(M)$ to be the space of functions $u \in C^{0,\alpha}(M)$ for which

$$\|u\|_{C^{0,\alpha}(M)} = \|u\|_{C^0(M)} + \sup_{x \ne y} \frac{|u(x) - u(y)|}{d_g(x,y)^\alpha}$$

is finite. This norm induces a structure of Banach space on $C^{0,\alpha}_B(M)$. Concerning the spaces $C^{k,\alpha}(M)$ and $C^{k,\alpha}_B(M)$, where $k \ge 1$ and $\alpha \in [0,1]$, a possible definition is the following: a function $u : M \to \mathbb{R}$ is in $C^{k,\alpha}(M)$ if it is in $C^k(M)$ and, given a system of charts on M, the coordinates of the tensor $\nabla^k u$ are in $C^{0,\alpha}$ when read via a chart. This definition is naturally independent of the choice of a C^∞ system of charts. However, the choice of a natural and good norm is a nontrivial question. We refer to [56] for a possible and natural norm on $C^{k,\alpha}(M)$.

6.1 The Euclidean case

The results in the Euclidean case can be seen as guidelines for the Riemannian case. The following result is due to Sobolev [68] and Morrey [59].

Theorem 6.1 *Let $p > n$. Then $H^p_1(\mathbb{R}^n)$ is continuously embedded in $C^{0,\alpha}_B(\mathbb{R}^n)$, with $\alpha = 1 - \frac{n}{p}$. More precisely, there exists $A > 0$ such that for any $u \in H^p_1(M)$, we have that*

$$\|u\|_{L^\infty(\mathbb{R}^n)} \le A\|u\|_{H^p_1(\mathbb{R}^n)}$$

and

$$|u(x) - u(y)| \leq A\|\nabla u\|_{L^p(\mathbb{R}^n)}|x - y|^\alpha \tag{15}$$

for a.e. $x, y \in \mathbb{R}^n$.

A remark is necessary here. To be precise, the Sobolev theorem asserts that when $u \in H_1^p(\mathbb{R}^n)$, $p > n$, then u has a representative in $C^{0,\alpha}(\mathbb{R}^n)$. Clearly, this representative is unique, and then, the embedding makes sense. Here, as in Section 4, the Hölder exponent $\alpha = 1 - \frac{n}{p}$ turns out to be natural. Indeed, let $p > n$ and assume that there exists $\alpha \in (0,1)$ and $A > 0$ such that (15) holds true for a.e. $x \in \mathbb{R}^n$. We choose $u \in C_c^\infty(\mathbb{R}^n) \setminus \{0\}$, let $\lambda > 0$, and consider the function $\tilde{u}(x) := u(\lambda x)$ for all $x \in \mathbb{R}^n$. Plugging $\tilde{u}$ into (15) and performing a change of variables, we get that

$$\sup_{x \neq y} \frac{|u(x) - u(y)|}{|x - y|^\alpha} \leq C\|\nabla u\|_p \lambda^{1 - \frac{n}{p} - \alpha}$$

for all $\lambda > 0$. Letting $\lambda \to 0$ and $\lambda \to +\infty$, it follows that we need to have that $\alpha = 1 - \frac{n}{p}$.

Proof. We let $u \in C_c^\infty(\mathbb{R}^n)$. We let also $r > 0$, $x_0, z_0 \in \mathbb{R}^n$ and $Q = z_0 + (-r,r)^n$ such that $x_0 \in Q$. Given $x \in Q$, we have that

$$u(x) - u(x_0) = \int_0^1 \frac{d}{dt}(u(tx + (1-t)x_0))\, dt = \int_0^1 (x - x_0)^i \partial_i u(tx + (1-t)x_0)\, dt$$

which yields

$$|u(x) - u(x_0)| \leq c \cdot r \int_0^1 |\nabla u(tx + (1-t)x_0)|\, dt.$$

We let $\bar{u} = \text{Vol}(Q)^{-1} \int_Q u\, dx$. Integrating the above inequality over Q yields

$$\begin{aligned} |\bar{u} - u(x_0)| &\leq c \cdot r^{1-n} \int_0^1 \left(\int_Q |\nabla u(tx + (1-t)x_0)|\, dx \right) dt \\ &\leq c \cdot r^{1-n} \int_0^1 t^{-n} \left(\int_{t(Q - x_0) + x_0} |\nabla u(x)|\, dx \right) dt \end{aligned}$$

With Hölder's inequality, and using that $t(Q - x_0) + x_0 \subset Q$, we get that

$$\begin{aligned} |\bar{u} - u(x_0)| &\leq c \cdot r^{1-n} \int_0^1 t^{-n} \|\nabla u\|_{L^p(Q)} \text{Vol}(t(Q - x_0) + x_0)^{1 - \frac{1}{p}}\, dt \\ &\leq c \cdot r^{1 - \frac{n}{p}} \int_0^1 t^{-\frac{n}{p}} \|\nabla u\|_{L^p(Q)}\, dt \leq \frac{c}{1 - \frac{n}{p}} r^{1 - \frac{n}{p}} \|\nabla u\|_{L^p(Q)}. \end{aligned} \tag{16}$$

We let $x, y \in \mathbb{R}^n$. We let $r = 2|x - y|$ and $Q = x + (-r,r)^n$. The above inequality applied to $x_0 = x$ and then to $x_0 = y$ yields that

$$|u(x) - u(y)| \leq \frac{c}{1 - \frac{n}{p}} r^{1 - \frac{n}{p}} \|\nabla u\|_{L^p(Q)} \leq \frac{c}{1 - \frac{n}{p}} |x - y|^{1 - \frac{n}{p}} \|\nabla u\|_{L^p(\mathbb{R}^n)}.$$

Taking $r = 1$, $Q = x + (-1,1)^n$ and $x_0 = x$, we get with (16) and Hölder's inequality that

$$\begin{aligned} |u(x)| &\leq 2^{-n} \int_Q |u|\, dx + C\|\nabla u\|_{L^p(Q)} \leq 2^{-n/p}\|u\|_{L^p(Q)} + C\|\nabla u\|_{L^p(Q)} \\ &\leq C\|u\|_{H_1^p(\mathbb{R}^n)}, \end{aligned}$$

which yields that

$$\|u\|_{L^\infty(\mathbb{R}^n)} \leq C\|u\|_{H_1^p(\mathbb{R}^n)}.$$

Coming back to the definition of $H_1^p(\mathbb{R}^n)$, the theorem follows. □

In the same spirit, we get that if $kp > n$, then $H_k^p(\mathbb{R}^n)$ is continuously embedded in $L^\infty(\mathbb{R}^n)$, with embeddings in Hölder spaces $C_B^{l,\theta}(\mathbb{R}^n)$ for particular values of l and θ. A very good reference on the subject is Adams [1]

6.2 The compact setting

Given a compact Riemannian manifold (M, g), the Sobolev embeddings in their second part follow from the above Euclidean case. In particular, the following result holds true.

Theorem 6.2 *Let (M, g) be a compact Riemannian manifold of dimension $n \geq 1$. Let $p \geq 1$ and assume that $p > n$. Then $H_1^p(M) \subset C^{0,\alpha}(M)$, where $\alpha = 1 - \frac{n}{p}$. Moreover, the embedding is continuous.*

Proof. Here again, since M is compact, we let $(\Omega_m, \varphi_m)_{m=1,\dots,N}$ as in the proof of Theorem 4.3, and we consider (η_m) a partition of unity subordinate to the covering (Ω_m). We have that there exists $C_1, C_2 > 0$ such that for any $m = 1, \dots, N$ and $u \in C^\infty(M)$, we have that

$$\|\eta_m u\|_{C^{0,\alpha}(M)} \leq C_1\|(\eta_m u) \circ \varphi_m^{-1}\|_{C^{0,\alpha}(\mathbb{R}^n)}$$

and

$$\|(\eta_m u) \circ \varphi_m^{-1}\|_{H_1^p(\mathbb{R}^n)} \leq C_2\|\eta_m u\|_{H_1^p(M)}$$

where the norms in the right-hand-side of the first inequality and in the left-hand-side in the second inequality are with respect to the Euclidean space. By Theorem 6.1, we get that there exists $C_3 > 0$ such that for any $m = 1, \dots, N$ and any $u \in C^\infty(M)$,

$$\|\eta_m u\|_{C^{0,\alpha}(M)} \leq C_3\|\eta_m u\|_{H_1^p(M)}.$$

Independently, one clearly has that there exists $B > 0$ such that for any function $u \in C^\infty(M)$,

$$\sum_{m=1}^N \|\eta_m u\|_{H_1^p(M)} \leq B\|u\|_{H_1^p(M)}.$$

Hence, for any $u \in C^\infty(M)$, we get that

$$\|u\|_{C^{0,\alpha}(M)} \leq \sum_{m=1}^{N} \|\eta_m u\|_{C^{0,\alpha}(M)} \leq BC_3 \|u\|_{H_1^p(M)}.$$

This ends the proof of the theorem. □

Concerning Sobolev spaces of higher order in differentiability, the following result is easy to get. For Hölder's analogues of Theorem 6.3 a possible reference is Lee-Parker [56].

Theorem 6.3 *Let (M, g) be a compact Riemannian manifold of dimension $n \geq 1$. We let $k \in \mathbb{N}$ and $p \geq 1$. We assume that $kp > n$. Then, for any $m \in \mathbb{N}$ such that $m < k - \frac{n}{p}$, we have $H_k^p(M) \subset C^m(M)$. Moreover, this embedding is continuous.*

6.3 The noncompact setting

As for Sobolev embeddings in their first part, the picture turns out to be intricate for complete noncompact manifolds. However, see Coulhon [25], essentially no new difficulties arise with respect to the embeddings in their first part. For historical references, we refer to Aubin [6] and Cantor [17]. The following result can be proved following the standard scheme of going back to Euclidean space by using harmonic coordinates.

Theorem 6.4 *Let (M, g) be a complete Riemannian manifold of dimension n with Ricci curvature bounded from below and with positive injectivity radius. Then for $p \geq 1$ and for $\lambda \leq 1 - \frac{n}{p}$, we have that $H_1^p(M) \subset C_B^{0,\lambda}(M)$, and this embedding is continuous.*

Concerning higher order spaces, the following result holds true.

Theorem 6.5 *Let (M, g) be a complete Riemannian manifold of dimension n with Ricci curvature bounded from below and with positive injectivity radius. Given $p \geq 1$ and $m < k - \frac{n}{p}$, we have that $H_k^p(M) \subset C_B^m(M)$, and the embedding is continuous.*

Improvement of these results are in Coulhon [25]. In particular, we can replace the lower bound on the injectivity radius by the assumption that there exists $v > 0$ such that $Vol_g(B_1(x)) \geq v$ for all $x \in M$.

7 Embedding theorems, Part III

We briefly discuss the limit case $kp = n$. A rough extension of the embeddings of part I yields that $H_k^p(M)$ should be embedded in $L^q(M)$ for all $q \geq p$. Indeed, at least for $\mathbb{R}^n$, this is true.

Theorem 7.1 *Let $n \geq 1$, $p \geq 1$ and $k \in \mathbb{N}$ such that $kp = n$. Then $H_k^p(\mathbb{R}^n) \hookrightarrow L^q(\mathbb{R}^n)$ for all $q \geq p$. Moreover, all these embeddings are continuous.*

The result can be improved on bounded domains of $\mathbb{R}^n$. A first step in this direction goes back to Trudinger [72] and Moser [60]. In what follows we define $\tilde{\nabla}^k$ by

$$\tilde{\nabla}^k u = \begin{cases} \Delta^{k/2} u & \text{if } k \text{ is odd,} \\ \nabla \Delta^{(k-1)/2} u & \text{if } k \text{ is even.} \end{cases}$$

The following result was obtained by Adams [2].

Theorem 7.2 *Let $n \geq 1$, $k \in \mathbb{N}$ and $p > 1$ such that $kp = n$. Then there exists a constant $c_0 = c_0(k, n)$ such that for any Ω bounded domain of $\mathbb{R}^n$, we have that*

$$\int_\Omega e^{\beta|u|^{\frac{p}{p-1}}}\, dx \leq c_0(k, n) Vol_\xi(\Omega) \tag{17}$$

for all $u \in C_c^\infty(\Omega)$ such that $\|\tilde{\nabla}^k u\|_p = 1$ and all $\beta \leq \beta_0(k, n)$, where

$$\beta_0(k,n) = \begin{cases} \frac{n}{\omega_{n-1}} \left[\frac{\pi^{n/2} 2^k \Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{n-k+1}{2}\right)}\right]^{\frac{p}{p-1}} & \text{if } k \text{ is odd,} \\ \frac{n}{\omega_{n-1}} \left[\frac{\pi^{n/2} 2^k \Gamma\left(\frac{k}{2}\right)}{\Gamma\left(\frac{n-k}{2}\right)}\right]^{\frac{p}{p-1}} & \text{if } k \text{ is even.} \end{cases}$$

Furthermore, if $\beta > \beta_0(k, n)$, then there exists a smooth $u \in C_c^\infty(\Omega)$ with $\|\tilde{\nabla}^k u\|_p \leq 1$ for which the integral in (17) *can be made as large as desired. In other words, the constant $\beta_0(k, n)$ is optimal.*

In the theorem, ω_{n-1} is the volume of $\mathbb{S}^{n-1}$, the unit $(n-1)-$sphere in $\mathbb{R}^n$ and $\Gamma(x) := \int_0^\infty t^{x-1} e^{-t}\, dt$ is the Gamma function. Concerning the history of this problem, Trudinger proved the existence of some $\beta > 0$ such that (17) holds true when $k = 1$, and Moser computed $\beta_0(1, n)$ for all $n \geq 2$. Using local charts, it is possible to get similar inequalities on compact manifolds. In the two following statements, $\bar{u} = \mathrm{Vol}_g(M)^{-1} \int_M u\, dv_g$. Moser [60] obtained the $2-$dimensional result below.

Theorem 7.3 *Let (M, g) be a compact Riemannian manifold of dimension 2. Then there exists $C(M) > 0$ such that*

$$\int_M e^{\alpha(u-\bar{u})^2}\, dv_g \leq C(M) \tag{18}$$

for all $\alpha \leq 4\pi$ and all $u \in H_1^2(M)$ such that $\|\nabla u\|_2 = 1$. Moreover, the constant 4π is optimal in the following sense: if there exists $\alpha \in \mathbb{R}$ such that (18) *holds for all $u \in H_1^2(M)$ such that $\|\nabla u\|_2 = 1$, then $\alpha \leq 4\pi$.*

This result of Moser was extended to arbitrary dimensions and arbitrary order of integrability by Fontana [41]. In the spirit of the above result of Adams, Fontana proved the following (one should keep in mind that on a compact Riemannian manifold, $\tilde{\nabla}^k u = 0$ is equivalent to $u = C^{st}$).

Theorem 7.4 *Let (M, g) be a compact Riemannian manifold of dimension $n \geq 2$, $k \in \mathbb{N}$ and $p > 1$ such that $kp = n$. Then there exists a constant $c_0 = c_0(k, M)$ such that we have that*

$$\int_M e^{\alpha|u|^{\frac{p}{p-1}}}\, dx \leq c_0(k, M) \tag{19}$$

for all $u \in H_k^p(\Omega)$ such that $\|\tilde{\nabla}^k u\|_p = 1$ and all $\alpha \leq \beta_0(k, n)$, where $\beta_0(k, n)$ is as in Theorem 7.2. Moreover, the constant $\beta_0(k, n)$ is optimal in the following sense: if there exists $\alpha \in \mathbb{R}$ such that (19) *holds for all $u \in H_k^p(M)$ such that $\|\tilde{\nabla}^k u\|_p = 1$, then $\alpha \leq \beta_0(k, n)$.*

The question of the extremals of such inequalities has also been considered by various authors. We refer to Chang [21], Adimurthi-Druet [3], Struwe [70] and the references therein. In the case $n = kp$, there are other useful and interesting functional inequalities, like Onofri's inequality. We refer to Chang [21] and Aubin [9] for related results.

8 Compact embeddings

We briefly discuss improvements on the structure of the Sobolev embeddings. More precisely, we address the question of the compactness of these embeddings. Concerning terminology, we say that an embedding $f : X \to Y$ between two metric spaces is compact if the image of any bounded set of X is relatively compact in Y. In other words if $\overline{f(B)}$ is compact for all bounded subsets B in X. This question is of importance, in particular for the use of variational methods. The first result we state concerns Sobolev embeddings in their first part.

Theorem 8.1 *Let (M, g) be a compact Riemannian manifold of dimension n. We let $p \geq 1$ and $k \in \mathbb{N}^\star$ be such that $kp < n$. We let $q \geq 1$ and $m \in \mathbb{N}$ be such that $m < k$ and $1 \leq q < \frac{pn}{n-(k-m)p}$. Then the embedding $H_k^p(M) \subset H_m^q(M)$ is compact. In particular, for $p \in [1, n)$ and $q \in [1, \frac{np}{n-p})$, the embedding $H_1^p(M) \hookrightarrow L^q(M)$ is compact.*

In other words, we have compactness as soon as the exponent is subcritical. A key point in proving such a result is the characterisation of compact subsets of Lebesgue's spaces. Let Ω be an open subset of $\mathbb{R}^n$, $p \geq 1$, and X be a bounded subset of $L^p(\Omega)$. Then we characterise the relative compactness of X by the sufficient and necessary condition that for any $\varepsilon > 0$, there exists a compact subset $K \subset \Omega$ and there exists $0 < \delta < d(K, \partial\Omega)$ such that

$$\int_{\Omega \backslash K} |u(x)|^p dx < \varepsilon \text{ and } \int_K |u(x+y) - u(x)|^p dx < \varepsilon$$

for all $u \in X$ and all y such that $|y| < \delta$. An independent important remark on the above result is that we loose compactness as soon as one reaches the critical exponent. For instance, the family $(\varphi_\epsilon)_{\epsilon>0} \in C^\infty(M)$ of the proof of Proposition 4.1 satisfies that $\|\varphi_\epsilon\|_{H_1^p} = O(1)$ when $\epsilon \to 0$, but no subsequence of (φ_ϵ) converges in $L^{\frac{np}{n-p}}(M)$. Concerning Sobolev embeddings in their second part, a similar theorem holds true.

Theorem 8.2 *Let (M, g) be a compact Riemannian manifold of dimension n. We let $p \geq 1$ such that $p > n$. Then the embedding $H_1^p(M) \subset C^{0,\alpha}(M)$ is compact for all $\alpha \in (0, 1)$ such that $\alpha < 1 - \frac{n}{p}$. In particular, the embedding $H_1^p(M) \subset C^0(M)$ is compact.*

9 Best constants

We restrict ourselves for the sake of simplicity to the case of compact manifolds. Results in the case of complete noncompact manifolds can be found in [47]. Let (M, g) be a compact Riemannian manifold. Then the Sobolev embeddings and inequalities are all valid for M. In particular, given k, p such that $kp < n$, the continuity of the embedding of $H_k^p(M)$ into $L^{\frac{pn}{n-pk}}(M)$ implies that there exists $C > 0$ such that

$$\|u\|_{\frac{pn}{n-pk}} \leq C\|u\|_{H_k^p}$$

for all $u \in H_k^p(M)$. In the special case $k = 1$, we get that for any $p < n$, there exists $A, B > 0$ such that

$$\|u\|_{\frac{pn}{n-p}} \leq A\|\nabla u\|_p + B\|u\|_p \tag{I_p^1}$$

for all $u \in H_1^p(M)$. From the pde point of view (see subsection 9.1), such inequalities are of real interest. In particular, raising (I_p^1) to the power p, we get that there exists $A, B > 0$ such that

$$\|u\|^p_{\frac{pn}{n-p}} \leq A\|\nabla u\|_p^p + B\|\nabla u\|_p^p \tag{I_p^p}$$

More generally, raising (I_p^1) to the power $\theta \in [1, p]$, we get that there exists $A, B > 0$ such that

$$\|u\|^\theta_{\frac{pn}{n-p}} \leq A\|\nabla u\|_p^\theta + B\|\nabla u\|_p^\theta \tag{I_p^θ}$$

for all $u \in H_1^p(M)$. Natural questions are to ask for the best possible value of A in (I_p^θ), the best possible value of B, the validity of sharp inequalities, and the existence of extremal functions. At this stage we let

$$A_{\theta,p}(M) = \inf\left\{A > 0 \text{ s.t. there exists } B > 0 \text{ s.t. } (I_p^\theta) \text{ holds for all } u \in H_1^p(M)\right\}$$

and ask the following questions: What is the value of $A_{\theta,p}(M)$? Is the best constant $A_{\theta,p}(M)$ achieved? In other words, is there a constant $B > 0$ such that

$$\|u\|^\theta_{\frac{pn}{n-p}} \leq A_{\theta,p}(M)\|\nabla u\|_p^\theta + B\|u\|_p^\theta \tag{$I_{p,opt}^\theta$}$$

for all $u \in H_1^p(M)$? Clearly, we have that $(I_{p,opt}^\theta)$ holds true if $(I_{p,opt}^{\theta'})$ holds true for some $\theta' > \theta$. In particular, $(I_{p,opt}^\theta)$ holds true for all $\theta \in [1, p]$ if and only if $(I_{p,opt}^p)$ holds true.

9.1 Applications to pde's

We let $a \in C^\infty(M)$ and $q \in (2, \frac{2n}{n-2}]$. We assume that $\Delta_g + a$ is coercive. By coercive we mean that there exists $\lambda > 0$ such that

$$\int_M \left(|\nabla u|^2 + au^2\right) dv_g \geq \lambda \int_M u^2\, dv_g$$

for all $u \in H_1^2(M)$. We address the question of the existence of nontrivial functions $u \in C^\infty(M)$, $u > 0$, solutions of an equation like

$$\Delta_g u + au = u^{q-1} \tag{20}$$

in M. Without going into the details, this question is related to the Yamabe problem of finding conformal metrics with constant scalar curvature on compact manifolds. An interesting survey paper on the subject is by Lee and Parker [56]. One possibility for solving the above equation is to consider the associated functional

$$I_q(u) := \frac{\int_M (|\nabla u|^2 + au^2)\, dv_g}{\left(\int_M |u|^q\, dv_g\right)^{2/q}}$$

for $u \in H_1^2(M)\backslash\{0\}$, and try to minimize the functional. Positive smooth critical points of I_q are, up to a scale factor, solutions of (20). When $q < \frac{2n}{n-2}$ (we say that q is subcritical), then the embedding $H_1^2(M) \hookrightarrow L^q(M)$ is compact, and one easily gets the existence of a minimizer for I_q. The comparison principle and standard elliptic theory yields that the minimizer is smooth and positive. When $q = \frac{2n}{n-2}$, the situation is more intricate and the existence of minimizers can be obtained for small energies. More precisely, the following result of Aubin [8] holds true.

Theorem 9.1 *Let (M, g) be a compact Riemannian manifold of dimension $n > 2$. Let $a \in C^\infty(M)$ such that $\Delta_g + a$ is coercive. We assume that*

$$\inf_{u\in H_1^2(M)\backslash\{0\}} I_{\frac{2n}{n-2}}(u) < \frac{1}{A_{2,2}(M)}.$$

Then there exists $u \in C^\infty(M)$ such that $u > 0$ and $\Delta_g u + au = u^{\frac{n+2}{n-2}}$ in M.

As a remark, the large inequality always holds true. As another remark, there are situations in which there is no solution to equations like (20) when $q = \frac{2n}{n-2}$. The above Theorem 9.1 is by now classical. We refer to the original reference Aubin [8] for a detailed proof. If one is interested in p-Laplace critical equations, one also get existence results similar to the Theorem 9.1. The p-Laplace operator is defined by

$$\Delta_{g,p} u = -\mathrm{div}_g(|\nabla u|^{p-2}\nabla u).$$

As shown by Druet [31], if $a \in C^\infty(M)$ is such that

$$\int_M (|\nabla u|^p + a|u|^p)\, dv_g \ge \lambda \int_M |u|^p\, dv_g$$

for all $u \in H_1^p(M)$ and some $\lambda > 0$ independent of u, and if

$$\inf_{u\in H_1^p(M)\backslash\{0\}} \frac{\int_M (|\nabla u|^p + a|u|^p)\, dv_g}{\left(\int_M |u|^{\frac{np}{n-p}}\, dv_g\right)^{\frac{n-p}{n}}} < \frac{1}{A_{p,p}(M)},$$

then there exists $u \in \cap_{\alpha\in(0,1)} C^{1,\alpha}(M)$ such that $u > 0$ and

$$\Delta_{g,p} u + au^{p-1} = u^{\frac{np}{n-p}-1}$$

in the distributional sense. Moreover, the regularity $C^{1,\alpha}$ is sharp. Summarizing, the sharp constant $A_{p,p}(M)$, and then the inequality (I_p^p), are of particular interest for solving critical pde's.

9.2 The value of $A_{\theta,p}(M)$

We discuss the question of the exact value of the sharp constant $A_{\theta,p}(M)$. To answer this question, a few definitions are requested. We let

$$\frac{1}{K(n,p)} = \inf_{u\in C_c^\infty(\mathbb{R}^n)\backslash\{0\}} \frac{\left(\int_{\mathbb{R}^n} |\nabla u|^p\, dx\right)^{\frac{1}{p}}}{\left(\int_{\mathbb{R}^n} |u|^{\frac{np}{n-p}}\right)^{\frac{n-p}{np}}}. \tag{21}$$

This constant was computed by Aubin [7], Talenti [71], and Rodemich [64]. One finds that

$$K(n,1) = \frac{1}{n}\left(\frac{n}{\omega_{n-1}}\right)^{1/n}$$

and

$$K(n,p) = \frac{p-1}{n-p}\left(\frac{n-p}{n(p-1)}\right)^{\frac{1}{p}}\left(\frac{\Gamma(n+1)}{\Gamma(\frac{n}{p})\Gamma(n+1-\frac{n}{p})\omega_{n-1}}\right)^{\frac{1}{n}}$$

for $p > 1$, where ω_{n-1} is the volume of the unit $(n-1)$-sphere in $\mathbb{R}^n$ and

$$\Gamma(x) := \int_0^\infty t^{x-1}e^{-t}\,dt$$

is the Gamma function. In particular, when $p = 2$, we get the nice expression

$$K(n,2) = \sqrt{\frac{4}{n(n-2)\omega_n^{2/n}}}$$

for all $n \geq 3$. The computation of $K(n,p)$ relies on previous work by Bliss [16] where the value of the sharp constant was computed for radially symmetric functions. The argument goes as follows. By standard Morse theory, it suffices to prove the sharp inequality for continuous nonnegative functions u with compact support $\overline{\Omega}$, $\overline{\Omega}$ being itself smooth, u being smooth in $\overline{\Omega}$ and such that it has only nondegenerate critical points in $\overline{\Omega}$. For such an u, let $u^\star : \mathbb{R}^n \to \mathbb{R}$, radially symmetric, nonnegative, and decreasing in $|x|$ be defined by

$$Vol_\xi\left(\{x \in \mathbb{R}^n, u^\star(x) \geq t\}\right) = Vol_\xi\left(\{x \in \mathbb{R}^n, u(x) \geq t\}\right)$$

where ξ stands for the Euclidean metric, and $Vol_\xi(X)$ stands for the Euclidean volume of X. It is easily seen that $u^\star$ has compact support and is Lipschitz. Moreover, the co-area formula gives that for any $m \geq 1$,

$$\int_{\mathbb{R}^n} |\nabla u|^m dx \geq \int_{\mathbb{R}^n} |\nabla u^\star|^m dx \quad \text{and} \quad \int_{\mathbb{R}^n} |u|^m dx = \int_{\mathbb{R}^n} |u^\star|^m dx$$

It follows that it suffices to prove the sharp inequality for decreasing absolutely continuous radially symmetric functions which equal zero at infinity, and we are back to the Bliss argument. In particular, the argument provides extremals for (21). When passing to manifolds, the following result of Aubin [7] can be proved.

Theorem 9.2 *Let (M,g) be a compact Riemannian manifold of dimension n, $p \in [1,n)$, and $\theta \in [1,p]$. Then we have that $A_{\theta,p}(M) = K(n,p)^\theta$.*

In particular, it follows from this result that for any $\epsilon > 0$, there exists $B_\epsilon > 0$ such that

$$\|u\|^\theta_{\frac{pn}{n-p}} \leq (K(n,p)^\theta + \epsilon)\|\nabla u\|_p^\theta + B_\epsilon\|\nabla u\|_p^\theta$$

for all $u \in H_1^p(M)$.

Proof. The proof proceeds in two steps. First we claim that $A_{\theta,p}(M) \geq K(n,p)^\theta$. We prove the claim by contradiction and assume that $A_{\theta,p}(M) < K(n,p)^\theta$. In this case, there exists $\alpha < K(n,p)^\theta$ and $B > 0$ such that

$$\|u\|^\theta_{\frac{pn}{n-p}} \leq \alpha\|\nabla u\|^\theta_p + B\|\nabla u\|^\theta_p \tag{22}$$

for all $u \in H^p_1(M)$. Let $u \in C^\infty_c(\mathbb{R}^n)$. Let $x_0 \in M$. For any $\epsilon > 0$, we define

$$u_\epsilon(x) = \epsilon^{-\frac{n-p}{p}} u(\epsilon^{-1}\exp^{-1}_{x_0}(x))$$

for $d_g(x, x_0) < i_g(x_0)$, and $u_\epsilon(x) = 0$ elsewhere. Clearly $u_\epsilon \in C^\infty(M)$ and has compact support in $B_{R\epsilon}(x_0)$, where supp $u \subset B_R(0)$. Working in the local chart $\exp^{-1}_{x_0}$, we get that

$$\begin{aligned} \int_M |u_\epsilon|^{\frac{np}{n-p}}\, dv_g &= \int_{B_{R\epsilon}(0)} \epsilon^{-n}|u|^{\frac{np}{n-p}}(\epsilon^{-1}x)\sqrt{|g|}(x)\, dx \\ &= \int_{B_R(0)} |u|^{\frac{np}{n-p}}(x)\sqrt{|g|}(\epsilon x)\, dx \end{aligned}$$

and that

$$\begin{aligned} \int_M |u_\epsilon|^p\, dv_g &= \int_{B_{R\epsilon}(0)} \epsilon^{p-n}|u|^p(\epsilon^{-1}x)\sqrt{|g|}(x)\, dx \\ &= \epsilon^p \int_{B_R(0)} |u|^p(x)\sqrt{|g|}(\epsilon x)\, dx \end{aligned}$$

and

$$\begin{aligned} \int_M |\nabla u_\epsilon|^p\, dv_g &= \int_{B_{R\epsilon}(0)} \epsilon^{p-n}\left(g^{ij}(x)\partial_i(u(\epsilon^{-1}x))\partial_j(u(\epsilon^{-1}x))\right)^{p/2}\sqrt{|g|}(x)\, dx \\ &= \int_{B_R(0)} \left(g^{ij}(\epsilon x)\partial_i u\partial_j u\right)^{p/2}\sqrt{|g|}(\epsilon x)\, dx. \end{aligned}$$

Letting $\epsilon \to 0$ and using that $\exp^{-1}_{x_0}$ is a normal chart at x_0, we get that

$$\lim_{\epsilon\to 0}\|u_\epsilon\|_{\frac{pn}{n-p}} = \|u\|_{L^{\frac{np}{n-p}}(\mathbb{R}^n)}, \quad \lim_{\epsilon\to 0}\|\nabla u_\epsilon\|_p = \|\nabla u\|_{L^p(\mathbb{R}^n)}$$

and $\lim_{\epsilon\to 0}\|u_\epsilon\|_p = 0$. Plugging these three limits in (22), it follows that

$$\|u\|_{L^{\frac{np}{n-p}}(\mathbb{R}^n)} \leq \alpha^{1/\theta}\|\nabla u\|_{L^p(\mathbb{R}^n)}$$

for all $u \in C^\infty_c(\mathbb{R}^n)$. A contradiction with the definition of $K(n,p)$ since $\alpha < K(n,p)^\theta$. This proves the claim. Now we claim that $A_{\theta,p}(M) \leq K(n,p)^\theta$, and more precisely that for any $\epsilon > 0$, there exists $B_\epsilon > 0$ such that

$$\left(\int_M |u|^{p^\star}\, dv_g\right)^{p/p^\star} \leq (K(n,p)^p + \epsilon)\int_M |\nabla u|^p\, dv_g + B_\epsilon\int_M |u|^p\, dv_g \tag{23}$$

for all $u \in H^p_1(M)$. We just sketch the proof of this result. We let $x_0 \in M$ and let $\delta \in (0, i_g(x_0))$. We consider the chart given by the inverse of the exponential map at x_0. We let $\eta > 0$. Up to choosing δ small enough, we get that

$$\frac{1}{1+\eta}\delta_{ij} \leq g_{ij} \leq (1+\eta)\delta_{ij}$$

in $B_\delta(0)$ in the sense of bilinear forms in the exponential chart. We let $u \in C_c^\infty(B_\delta(x_0))$ and we let $\tilde{u} = u \circ \exp_{x_0} \in C_c^\infty(B_\delta(0))$. We get that

$$\int_M |u|^{p^\star}\, dv_g = \int_{B_\delta(0)} |\tilde{u}|^q \sqrt{|g|}\, dx \le (1+\eta)^{n/2} \int_{\mathbb{R}^n} |\tilde{u}|^{p^\star}\, dx$$

and

$$\int_M |\nabla u|^p\, dv_g = \int_{B_\delta(0)} \left(g^{ij}\partial_i\tilde{u}\partial_j\tilde{u}\right)^{p/2} \sqrt{|g|}\, dx \ge (1+\eta)^{-(p+n)/2} \int_{\mathbb{R}^n} |\nabla\tilde{u}|^p\, dx.$$

With the optimal inequality (21), we get that

$$\left(\int_M |u|^{p^\star}\, dv_g\right)^{p/p^\star} \le (1+\eta)^{\frac{np}{2p^\star}+\frac{p+n}{2}} K(n,p)^p \int_M |\nabla u|^p\, dv_g. \tag{24}$$

Choosing η as small as needed, we then get that for any $\epsilon > 0$, there exists $\delta_{x_0,\epsilon} > 0$ such that

$$\left(\int_M |u|^{p^\star}\, dv_g\right)^{p/p^\star} \le \left(K(n,p)^p + \frac{\epsilon}{2}\right) \int_M |\nabla u|^p\, dv_g$$

for all $u \in C_c^\infty(B_{\delta_{x_0,\epsilon}}(x_0))$. The general case now goes as follows: since M is compact, there exists $x_1, ..., x_N \in M$ such that $M \subset \cup_{s=1}^N B_{\delta_{x_s,\epsilon}}(x_s)$. We consider a partition of unity subordinate to this covering, and we use inequality (24) for the points $x_1, ..., x_N$. Inequality (23) follows. We refer to the original article by Aubin [7] or to Hebey [47] for an exposition in book form. With (23), and the inequality $A_{\theta,p}(M) \ge K(n,p)^\theta$, the theorem is proved. □

9.3 Attainability of the first best constant

It has been a long standing conjecture to know weither the optimal constant in (I_p^θ) is attained or not. When $p = 2$, the conjecture was solved by Hebey and Vaugon [49, 50]. Their result states as follows.

Theorem 9.3 *Let (M,g) be a compact Riemannian manifold of dimension $n \ge 3$. Then there exists $B > 0$ such that*

$$\left(\int_M |u|^{\frac{2n}{n-2}}\, dv_g\right)^{\frac{n-2}{n}} \le K(n,2)^2 \int_M |\nabla u|^2\, dv_g + B\int_M u^2\, dv_g$$

for all $u \in H_1^2(M)$. In other words, $(I_{2,opt}^2)$ holds true on M.

The geometry of the manifold plays no role in this result.

Proof. We give here a general idea of the proof. For the detailed proof, we refer to the original papers [49, 50]. The proof relies on intricate blow-up arguments. Let $\alpha > 0$ be some positive real number, and for $u \in H_1^2(M)$, let

$$I_\alpha(u) = \frac{\int_M \left(|\nabla u|^2 + \alpha u^2\right)\, dv_g}{\left(\int_M |u|^{\frac{2n}{n-2}}\, dv_g\right)^{\frac{n-2}{n}}}.$$

Clearly, Theorem 9.3 is equivalent to the existence of $\alpha_0 > 0$ such that

$$\inf_{u \in H_1^2(M)\backslash\{0\}} I_{\alpha_0}(u) \geq \frac{1}{K(n,2)^2}.$$

The proof goes by contradiction: we assume that for any $\alpha > 0$, we have that

$$\inf_{u \in H_1^2(M)\backslash\{0\}} I_\alpha(u) < \frac{1}{K(n,2)^2}.$$

Then it follows from Theorem 9.1 that for any $\alpha > 0$, there exists $u_\alpha \in C^\infty(M)$ and there exists $\lambda_\alpha \in (0, K(n,2)^{-2})$ such that

$$\Delta_g u_\alpha + \alpha u_\alpha = \lambda_\alpha u_\alpha^{\frac{n+2}{n-2}}, \; u_\alpha > 0$$

in M and $\int_M |u_\alpha|^{\frac{2n}{n-2}} \, dv_g = 1$. Now the idea is to prove that u_α does not exist for α large enough. Here again, the difficulty is due to the critical exponent. It is easy to prove that $u_\alpha \rightharpoonup 0$ when $\alpha \to 0$ in the weak sense, but

$$\lim_{\alpha \to 0} \|u_\alpha\|_{H_1^2(M)} = K(n,2)^{-2}.$$

In the Euclidean setting, a powerful tool for the proof of nonexistence results for such equations is the Pohozaev identity [63]. Hebey and Vaugon proved that there exist $(x_\alpha)_{\alpha>0} \in M$ and $(\mu_\alpha)_{\alpha>0} \in \mathbb{R}_{>0}$ such that $\lim_{\alpha\to 0} \mu_\alpha = 0$ and such that there exists $C > 0$ such that

$$u_\alpha(x) \leq C \left(\frac{\mu_\alpha}{\mu_\alpha^2 + d_g(x, x_\alpha)^2} \right)^{\frac{n-2}{2}} \tag{25}$$

for all $x \in M$ and all $\alpha > 0$. Such type of inequalities have been improved by Druet-Hebey-Robert [38]. The Euclidean Pohozaev identity asserts that for any $u \in C^2(\overline{\Omega})$, where Ω is a smooth bounded oriented open subset of $\mathbb{R}^n$, we have that

$$\begin{aligned} &\int_\Omega x^i \partial_i u \Delta_\xi u \, dx + \frac{n-2}{2} \int_\Omega u \Delta_\xi u \, dx \\ &= \int_{\partial\Omega} \left(-\frac{n-2}{2} u \partial_\nu u + \frac{(x,\nu)}{2} |\nabla u|^2 - x^i \partial_i u \partial_\nu u \right) d\sigma, \end{aligned}$$

where ν denotes the outer normal vector at $\partial\Omega$ and $d\sigma$ is the measure associated to $\partial\Omega$. Via the exponential chart at x_α, one can write $\tilde{u}_\alpha = u_\alpha \circ \exp_{x_\alpha}$ in an open ball of $\mathbb{R}^n$. It solves

$$\Delta_{g_\alpha} \tilde{u}_\alpha + \alpha \tilde{u}_\alpha = \lambda_\alpha \tilde{u}_\alpha^{\frac{n+2}{n-2}},$$

where $g_\alpha = \exp_{x_\alpha}^\star g$ is the pull-back of the metric g. Plugging $\tilde{u}_\alpha$ into the above Pohozaev identity and estimating carefully the difference $g_\alpha - \xi$ of the two metrics, thanks to (25), one gets a contradiction for α large enough. This proves that the optimal Sobolev inequality holds true. □

A major difficulty in the above proof is to manage the perturbation g_α of the Euclidean metric. It is interesting to see that managing this small perturbation requires a lot of pde material. In the general case, where p is arbitrary, the critical value for the exponent happens to be precisely 2. The following result was proved independently by Druet [32] and Aubin-Li [11].

Theorem 9.4 *Let (M, g) be a compact Riemannian manifold of dimension $n \geq 2$. Let $p \in (1, n)$. Assume that $\theta = p$ if $p \leq 2$ and that $\theta = 2$ if $p \geq 2$. Then $(I^\theta_{p,opt})$ holds true.*

As a remark, it follows from this result that $(I^1_{p,opt})$ holds true for all $p > 1$. When $p > 2$ and $\theta > 2$, the geometry of the manifold starts playing a role as shown by Druet [33]. In particular, Druet [33] proves that $(I^p_{p,opt})$ is false if $p > 2$, $n > 3p - 2$, and the scalar curvature of g is positive somewhere. On the other hand, by Aubin-Druet-Hebey [10] and Druet [33, 35], we can prove that $(I^p_{p,opt})$ is true if the sectional curvature of (M, g) is nonpositive and the Cartan-Hadamard conjecture is true, or if the scalar curvature of (M, g) is negative. When the Ricci curvature of the manifold is nonnegative, geometric rigidity is attached to $(I^p_{p,opt})$, as shown by the following result of Druet [33].

Theorem 9.5 *Let (M, g) be a compact Riemannian n-manifold of nonnegative Ricci curvature. Assume that $(I^p_{p,opt})$ is valid on (M, g) for some p with $p > 4$ and $5p - 4 < n$. Then g is flat, and M is covered by a torus.*

Proof. In order to prove the theorem, we claim first that if (M, g) is Ricci flat, but not flat, and if $p > 4$ and $5p - 4 < n$, then inequality $(I^p_{p,opt})$ is false on (M, g). Let W_g be the Weyl curvature of g, and let $x_0 \in M$ be such that $W_g(x_0) \not\equiv 0$. For $\varepsilon > 0$, we set

$$u_\varepsilon = \left(\varepsilon + r^{\frac{p}{p-1}}\right)^{1-\frac{n}{p}} \varphi(r)$$

where r denotes the distance to x_0, φ is smooth such that $0 \leq \varphi \leq 1$, $\varphi = 1$ on $\left(-\frac{\delta}{2}, \frac{\delta}{2}\right)$, and $\varphi = 0$ if $r \geq \delta$, and $\delta > 0$, δ small, is real. In order to prove the claim, it suffices to prove that for any $\alpha > 0$, and for ε small enough, $J(u_\varepsilon) < K(n, p)^{-p}$, where

$$J(u) = \frac{\int_M |\nabla u|^p dv_g + \alpha \int_M u^p dv_g}{\left(\int_M |u|^{p^\star} dv_g\right)^{p/p^\star}}$$

Computing the expansion of $J(u_\varepsilon)$ in terms of ε, we get that

$$J(u_\varepsilon) \leq \frac{1}{K(n,p)^p} + \frac{A|W_g(x_0)|^2}{120K(n,p)^p n(n+2)}\varepsilon^{4\frac{p-1}{p}} + o\left(\varepsilon^{4\frac{p-1}{p}}\right)$$

where

$$A = \frac{n-p}{n}\frac{\int_0^\infty \left(1 + s^{\frac{p}{p-1}}\right)^{-n} s^{n+3} ds}{\int_0^\infty \left(1 + s^{\frac{p}{p-1}}\right)^{-n} s^{n-1} ds} - \frac{\int_0^\infty \left(1 + s^{\frac{p}{p-1}}\right)^{-n} s^{\frac{p}{p-1}+n+3} ds}{\int_0^\infty \left(1 + s^{\frac{p}{p-1}}\right)^{-n} s^{\frac{p}{p-1}+n-1} ds}$$

As one can easily check, $A < 0$. It follows that there exists $\varepsilon > 0$ small enough such that

$$J(u_\varepsilon) < \frac{1}{K(n,p)^p}$$

and the above claim is proved. Let now (M, g) be a compact n-dimensional Riemannian manifold of nonnegative Ricci curvature, and let p be such that $p > 4$ and $5p - 4 < n$. Assume that $(I^p_{p,opt})$ is valid on (M, g). By the above mentionned result of Druet that $(I^p_{p,opt})$ is false if $p > 2$, $n > 3p - 2$, and the scalar curvature is positive somewhere, we get that g has to be Ricci flat. Then, by the above claim, we get that g is flat. The last assertion that M is covered by a torus follows from Bieberbach's Theorem. This proves the theorem. □

More questions on sharp inequalities when priority is given to the sharp constant are to compute or estimate the corresponding second constant when the sharp inequality is true, and also to decide whether or not the sharp inequality comes with extremal functions. For material on these questions we refer to Druet and Hebey [37].

9.4 The second best constant

Similar questions can be asked when priority is given to the second constant. The best second constant in (I^θ_p) is $B = V_g^{-\theta/n}$, where $V_g = \mathrm{Vol}_g(M)$ is the volume of (M, g). We say that the sharp form of (I^θ_p) with respect to the second constant is valid if there exists $A \in \mathbb{R}$ such that for any $u \in H^p_1(M)$,

$$\|u\|^\theta_{\frac{pn}{n-p}} \le A\|\nabla u\|^\theta_p + V_g^{-\theta/n}\|u\|^\theta_p \qquad (J^\theta_{p,opt})$$

It is easily seen that if $(J^\theta_{p,opt})$ is valid for some $\theta_0 \in [1, p]$, then $(J^\theta_{p,opt})$ is also valid for all $\theta \in [1, \theta_0]$. Combining results by Bakry [12], Druet (see Hebey [47]), and Hebey [47], it holds that if $p \le 2$, then $(J^\theta_{p,opt})$ is valid for any $\theta \in [1, p]$, and that if $p > 2$, then $(J^\theta_{p,opt})$ is valid if and only if $\theta \le 2$. When $p = 2$, $(J^2_{2,opt})$ holds true, and explicit inequalities can be found in Ilias [52]. In particular, as shown by Ilias [52], if there exists $k > 0$ such that $Ric_g \ge (n-1)kg$, one gets that

$$\left(\int_M |u|^{\frac{2n}{n-2}} dv_g\right)^{\frac{n-2}{n}} \le \frac{4}{n(n-2)kV_g^{2/n}} \int_M |\nabla u|^2 dv_g + V_g^{-2/n} \int_M u^2 dv_g$$

for all $u \in H^2_1(M)$. We prove in what follows that $(J^p_{p,opt})$ is false when $p > 2$. For that purpose, let $u \in C^\infty(M)$ be some nonconstant function. For $t > 2$ real, and $\epsilon > 0$, we define

$$\varphi_t(\epsilon) = \int_M |1 + \epsilon u|^t \, dv_g$$

Clearly, one has that

$$\varphi_t(\epsilon) = V_g + t\left(\int_M u \, dv_g\right)\epsilon + \frac{t(t-1)}{2}\left(\int_M u^2 \, dv_g\right)\epsilon^2 + o(\epsilon^2)$$

Hence,

$$\int_M |1 + \epsilon u|^p \, dv_g = V_g + p\left(\int_M u \, dv_g\right)\epsilon + \frac{p(p-1)}{2}\left(\int_M u^2 \, dv_g\right)\epsilon^2 + o(\epsilon^2)$$

and

$$\left(\int_M |1+\epsilon u|^q \, dv_g\right)^{p/q} = V_g^{p/q} + pV_g^{\frac{p}{q}-1}\left(\int_M u \, dv_g\right)\epsilon + \frac{p(q-1)}{2}V_g^{\frac{p}{q}-1}\left(\int_M u^2 \, dv_g\right)\epsilon^2 + \frac{p(p-q)}{2}V_g^{\frac{p}{q}-2}\left(\int_M u \, dv_g\right)^2\epsilon^2 + o(\epsilon^2)$$

where $q = np/(n-p)$. Suppose now that $(J^p_{p,opt})$ is valid. Noting that for $p > 2$,

$$\int_M |\nabla(1+\epsilon u)|^p \, dv_g = o(\epsilon^2)$$

one would get that for any $\epsilon > 0$,

$$V_g^{p/q} + pV_g^{\frac{p}{q}-1}\left(\int_M u \, dv_g\right)\epsilon + \frac{p(q-1)}{2}V_g^{\frac{p}{q}-1}\left(\int_M u^2 \, dv_g\right)\epsilon^2 + \frac{p(p-q)}{2}V_g^{\frac{p}{q}-2}\left(\int_M u \, dv_g\right)^2\epsilon^2 \le V_g^{1-\frac{p}{n}} + pV_g^{-\frac{p}{n}}\left(\int_M u \, dv_g\right)\epsilon + \frac{p(p-1)}{2}V_g^{-\frac{p}{n}}\left(\int_M u^2 \, dv_g\right)\epsilon^2 + o(\epsilon^2)$$

But $\frac{p}{q} = 1 - \frac{p}{n}$ so that, as one can easily check, such an inequality implies that

$$(q-1)\int_M u^2 \, dv_g \le (q-p)\frac{1}{V_g}\left(\int_M u \, dv_g\right)^2 + (p-1)\int_M u^2 \, dv_g$$

This means again that

$$V_g \int_M u^2 \, dv_g \le \left(\int_M u \, dv_g\right)^2$$

which is impossible as soon as u is nonconstant. This ends the proof that $(J^p_{p,opt})$ is not valid when $p > 2$.

9.5 The Nash inequality

The discussion on sharp Sobolev inequalities when priority is given to the first constant relies on the computation by Aubin and Talenti of the sharp constant in the Euclidean Sobolev inequality. Sharp constants for Euclidean Gagliardo-Nirenberg inequalities have been computed by Cordero-Erausquin, Nazaret, and Villani [24] and by Del Pino and Dolbeault [29]. We refer also to Beckner [14] for a related reference. In the case of the Nash inequality, the computation of the sharp constant is due to Carlen and Loss [18]. We let

$$C_n = \frac{(n+2)^{(n+2)/n}}{2^{2/n} n \lambda_1^N(\mathcal{B})|\mathcal{B}|^{2/n}}, \tag{26}$$

where $|\mathcal{B}|$ denotes the volume of the unit ball $\mathcal{B}$ in $\mathbb{R}^n$, and $\lambda_1^N(\mathcal{B})$ denotes the first nonzero Neumann eigenvalue of the Laplacian on radial functions on $\mathcal{B}$. The Carlen and Loss result [18] states as follows.

Theorem 9.6 *For any smooth function u with compact support in $\mathbb{R}^n$,*

$$\left(\int_{R^n} u^2\, dv_\xi\right)^{1+\frac{2}{n}} \le C_n \left(\int_{R^n} |\nabla u|^2\, dv_\xi\right)\left(\int_{R^n} |u| dv_\xi\right)^{\frac{4}{n}}$$

where C_n is as in (26). Moreover, C_n is the sharp constant in the inequality.

Proof. Let us first prove that the inequality of the theorem does hold. It suffices to establish this inequality for nonnegative, radially symmetric, decreasing functions. Let u be such a function. For $r > 0$ arbitrary, let

$$v(x) = \begin{cases} u(x) & \text{if } |x| \le r \\ 0 & \text{if } |x| > r \end{cases} \quad \text{and} \quad w(x) = \begin{cases} 0 & \text{if } |x| \le r \\ u(x) & \text{if } |x| > r \end{cases}$$

Clearly, $\|u\|_2^2 = \|v\|_2^2 + \|w\|_2^2$, and since u is radially decreasing,

$$w(x) \le u(r) \le \frac{1}{|\mathcal{B}|r^n}\|v\|_1$$

In particular,

$$\|w\|_2^2 \le \frac{1}{|\mathcal{B}|r^n}\|v\|_1\|w\|_1$$

Let

$$\overline{v} = \frac{1}{|\mathcal{B}|r^n}\|v\|_1$$

be the average of v. One gets from the variational characterization of λ_1^N that

$$\begin{aligned}
\|v\|_2^2 &= \int_{B_r(0)} (v-\overline{v})^2\, dx + \int_{B_r(0)} \overline{v}^2\, dx \\
&\le \frac{1}{\lambda_1^N\big(B_r(0)\big)}\int_{B_r(0)} |\nabla v|^2\, dx + \frac{1}{|\mathcal{B}|r^n}\|v\|_1^2 \\
&= \frac{r^2}{\lambda_1^N(\mathcal{B})}\int_{B_r(0)} |\nabla v|^2\, dx + \frac{1}{|\mathcal{B}|r^n}\|v\|_1^2 \\
&\le \frac{r^2}{\lambda_1^N(\mathcal{B})}\|\nabla u\|_2^2 + \frac{1}{|\mathcal{B}|r^n}\|v\|_1^2
\end{aligned}$$

where $B_r(0)$ stands for the Euclidean ball of center 0 and radius r. According to what we said above, and noting that

$$\|u\|_1^2 \ge \|v\|_1\big(\|v\|_1 + \|w\|_1\big)$$

this leads to

$$\|u\|_2^2 \le \frac{r^2}{\lambda_1^N(\mathcal{B})}\|\nabla u\|_2^2 + \frac{1}{|\mathcal{B}|r^n}\|u\|_1^2 \tag{27}$$

The right-hand side in this inequality is minimized at

$$r_{\min} = \left(\frac{n\lambda_1^N(\mathcal{B})}{2|\mathcal{B}|}\right)^{1/(n+2)} \left(\frac{\|u\|_1}{\|\nabla u\|_2}\right)^{2/(n+2)} \tag{28}$$

As one can easily check, taking $r = r_{\min}$ in (27) gives the inequality of the theorem. To see that this inequality is sharp, let u_0 be some eigenfunction associated to $\lambda_1^N(\mathcal{B})$. Set

$$u(x) = \begin{cases} u_0(|x|) - u_0(1) & \text{if } |x| \le 1 \\ 0 & \text{if } |x| \ge 1 \end{cases}$$

Clearly, u saturates (27) with $r = 1$. For such a function, $r_{\min} = 1$. One then easily gets from (28) that u also saturates the Nash inequality we just got. This ends the proof of the theorem. □

In addition to Theorem 9.6, Carlen and Loss [18] also determined the cases of equality in the optimal Nash inequality. As in the above proof, let u_0 be some eigenfunction associated to $\lambda_1^N(\mathcal{B})$, and set

$$u(x) = \begin{cases} u_0(|x|) - u_0(1) & \text{if } |x| \le 1 \\ 0 & \text{if } |x| \ge 1 \end{cases}$$

Then $\tilde{u}$ is an extremum function for the optimal Nash inequality if and only if after a possible translation, scaling, and normalization, $\tilde{u} = u$. As one can easily check, a striking feature of this result is that all of the extremals have compact support. Another reference on the subject, where the asymptotically sharp form with respect to dimension of the Nash inequality is investigated, is Beckner [15]

10 Explicit sharp inequalities

Explicit sharp inequalities can be given on specific compact and complete manifolds. By the sharp inequality of second order we refer to the inequality of Theorem 9.3 which states that there exists $B > 0$ such that

$$\left(\int_M |u|^{\frac{2n}{n-2}}\, dv_g\right)^{\frac{n-2}{n}} \le K(n,2)^2 \int_M |\nabla u|_g^2\, dv_g + B\int_M u^2\, dv_g \tag{29}$$

for all $u \in H_1^2(M)$. By the work of Hebey and Vaugon [49, 50] we know that the inequality holds true on compact manifolds. Still by the work of Hebey and Vaugon [49, 50], the inequality holds also true on complete Riemannian manifold with positive injectivity radius and curvature bounded up to the order 1. For such manifolds, in specific cases, sharp explicit inequalities can be computed. We quote results from Hebey and Vaugon [48]. In the case of the projective space $\mathbb{P}^n(\mathbb{R})$ we can, for instance, prove that (29) holds true with

$$B \leq \frac{n+2}{(n-2)\omega_n^{2/n}},$$

where ω_n is the volume of the unit sphere $\mathbb{S}^n$. The result extends to other quotients of the sphere. Similarly, if $\mathbb{S}^1(T) \times \mathbb{S}^{n-1}$ represents the product of the circle of radius T and the unit sphere $\mathbb{S}^{n-1}$, with the standard metric, then we can also prove that (29) holds true with

$$B \leq \frac{1+(n-2)^2T^2}{n(n-2)T^2\omega_n^{2/n}}.$$

In the case of complete noncompact manifolds, the following inequalities can be proved to hold. The result is quoted from Hebey [47].

Proposition 10.1 *The optimal inequality (29) holds true with*

(1) $B = -\frac{1}{\omega_n^{2/n}}$ *for the hyperbolic space* $\mathbb{H}^n$

(2) $B = \frac{m-n}{(m+n)\omega_{m+n}^{2/(m+n)}}$ *for the product* $\mathbb{S}^m \times \mathbb{H}^n$, $m \geq 2$, $n \geq 2$

(3) $B = \frac{m-n+2}{(m+n-2)\omega_{m+n}^{2/(m+n)}}$ *for the product* $\mathbb{P}^m(\mathbb{R}) \times \mathbb{H}^n$, $m \geq 2$, $n \geq 2$

(4) $B = \frac{n-1}{(n+1)\omega_{n+1}^{2/(n+1)}}$ *for the product* $\mathbb{S}^n \times \mathbb{R}$, $n \geq 2$

(5) $B = \frac{n+1}{(n-1)\omega_{n+1}^{2/(n+1)}}$ *for the product* $\mathbb{P}^n(\mathbb{R}) \times \mathbb{R}$, $n \geq 2$

(6) $B = -\frac{n-1}{(n+1)\omega_{n+1}^{2/(n+1)}}$ *for the product* $\mathbb{H}^n \times \mathbb{R}$, $n \geq 2$.

Moreover, at least when the dimension of the manifold is greater than or equal to 4, *these values are the best possible for* $\mathbb{H}^n$, $\mathbb{S}^m \times \mathbb{H}^n$, $\mathbb{S}^n \times \mathbb{R}$, *and* $\mathbb{H}^n \times \mathbb{R}$.

To these results we should of course add the classical results that (29) holds true with $B = \omega_n^{-2/n}$ in the case of the sphere $\mathbb{S}^n$, and with $B = 0$ in the case of the Euclidean space $\mathbb{R}^n$.

11 The Cartan-Hadamard conjecture

By definition, a Cartan-Hadamard manifold is a complete, simply connected Riemannian manifold of nonpositive sectional curvature. For such manifolds, by the work of Hoffman and Spruck, (14) is valid. The Cartan-Hadamard conjecture, a longstanding conjecture in the mathematical literature, states that for Cartan-Hadamard manifolds, (14) holds true when $p = 1$ with the best possible value $C_1 = K(n, 1)$. In other words, the Cartan-Hadamard conjecture states that for Cartan-Hadamard n-dimensional manifolds, the sharp inequality

$$\left(\int_M |u|^{\frac{n}{n-1}}\, dv_g\right)^{\frac{n-1}{n}} \leq K(n,1)\int_M |\nabla u| dv_g \tag{30}$$

holds true, where $K(n,1)$ is as in (21). By the work of Federer and Fleming [40], this is equivalent to saying that for any smooth, bounded domain Ω on a Cartan-Hadamard n-dimensional manifold (M,g), the sharp isoperimetric inequality

$$\text{Area}_g(\partial\Omega) \geq \frac{1}{K(n,1)}\text{Vol}_g(\Omega)^{1-\frac{1}{n}} \tag{31}$$

holds true. The sharp isoperimetric inequality (31) holds true for the Euclidean space in arbitrary dimension. Moreover, in that case, equality holds if and only if Ω is a ball. In the 2-dimensional case of Euclidean space, we are back to the famous ancient formula which states that for 2-dimensional domains in the plane, with area A and length L for their boundary, $L^2 \geq 4\pi A$. The equivalence of (30) and (31) is valid for arbitrary complete manifolds and the following proposition of Federer and Fleming [40] holds true.

Proposition 11.1 *The sharp isoperimetric inequality (31) is valid if and only if the sharp functional inequality (30) is valid.*

Proof. We prove that

$$\inf_{u\in H_1^1(M)\backslash\{0\}} \frac{\int_M |\nabla u| dv_g}{\left(\int_M |u|^{n/(n-1)}\, dv_g\right)^{(n-1)/n}} = \inf_\Omega \frac{\text{Area}_g(\partial\Omega)}{\text{Vol}_g(\Omega)^{1-\frac{1}{n}}} \tag{32}$$

As a starting point, consider Ω a smooth bounded domain in (M,g). For sufficiently small $\epsilon > 0$, let u_ϵ be the function

$$u_\epsilon(x) = \begin{cases} 1 & \text{if } x \in \Omega \\ 1 - \frac{1}{\epsilon} d_g(x,\partial\Omega) & \text{if } x \in M\backslash\Omega, d_g(x,\partial\Omega) < \epsilon \\ 0 & \text{if } x \in M\backslash\Omega, d_g(x,\partial\Omega) \geq \epsilon \end{cases}$$

where d_g stands for the distance associated to g. The function u_ϵ is Lipschitz for all $\epsilon > 0$. Moreover,

$$\lim_{\epsilon\to 0} \int_M u_\epsilon^{n/(n-1)}\, dv_g = \text{Vol}_g(\Omega)$$

and

$$|\nabla u_\epsilon| = \begin{cases} \frac{1}{\epsilon} & \text{if } x \in M\backslash\overline{\Omega}, d_g(x,\partial\Omega) < \epsilon \\ 0 & \text{otherwise} \end{cases}$$

Hence,

$$\lim_{\epsilon\to 0} \int_M |\nabla u_\epsilon| dv_g = \lim_{\epsilon\to 0} \frac{1}{\epsilon}\text{Vol}_g\left(\left\{x \notin \Omega / d_g(x,\partial\Omega) < \epsilon\right\}\right) = \text{Area}_g(\partial\Omega)$$

and one gets that

$$\inf_u \frac{\int_M |\nabla u| dv_g}{\left(\int_M |u|^{n/(n-1)}\, dv_g\right)^{(n-1)/n}} \leq \inf_\Omega \frac{\text{Area}_g(\partial\Omega)}{\text{Vol}_g(\Omega)^{1-\frac{1}{n}}} \tag{33}$$

Let us now prove the opposite inequality:

$$\inf_u \frac{\int_M |\nabla u| dv_g}{\left(\int_M |u|^{n/(n-1)}\, dv_g\right)^{(n-1)/n}} \geq \inf_\Omega \frac{\mathrm{Area}_g(\partial\Omega)}{\mathrm{Vol}_g(\Omega)^{1-\frac{1}{n}}} \tag{34}$$

Given u smooth with compact support in M, let $\Omega(t) = \{x \,/\, |u(x)| > t\}$, and $V(t) = \mathrm{Vol}_g(\Omega(t))$, for $t \in R_u$, the regular values of u. The proof of (34) is based on the co-area formula which states that for $f \in L^1(\mathrm{Supp} u)$,

$$\int_M f|\nabla u| dv_g = \int_0^\infty \left(\int_{\Sigma_t} f\, d\sigma\right) dt$$

where $\Sigma_t = |u|^{-1}(t)$. By the co-area formula,

$$\int_M |\nabla u| dv_g \geq \left(\inf_\Omega \frac{\mathrm{Area}_g(\partial\Omega)}{\mathrm{Vol}_g(\Omega)^{1-\frac{1}{n}}}\right) \int_0^\infty V(t)^{1-\frac{1}{n}}\, dt$$

and

$$\begin{aligned}
\int_M |u|^{n/(n-1)}\, dv_g &= \int_M \left(\int_0^{|u|} \frac{n}{n-1} t^{1/(n-1)} dt\right) dv_g \\
&= \frac{n}{n-1} \int_0^\infty \left(\int_{\Omega(t)} dv_g\right) t^{1/(n-1)} dt \\
&= \frac{n}{n-1} \int_0^\infty t^{1/(n-1)} V(t) dt
\end{aligned}$$

In order to prove (34), it suffices then to prove that

$$\int_0^\infty V(t)^{1-\frac{1}{n}}\, dt \geq \left(\frac{n}{n-1} \int_0^\infty t^{1/(n-1)} V(t) dt\right)^{1-\frac{1}{n}} \tag{35}$$

To establish (35), set

$$F(s) = \int_0^s V(t)^{1-\frac{1}{n}}\, dt\,, \qquad G(s) = \left(\frac{n}{n-1} \int_0^s t^{1/(n-1)} V(t) dt\right)^{1-\frac{1}{n}}$$

One has that $F(0) = G(0)$, and since $V(s)$ is a decreasing function of s,

$$\begin{aligned}
G'(s) &= \frac{n-1}{n} \left(\frac{n}{n-1}\right)^{1-\frac{1}{n}} \left(\int_0^s t^{1/(n-1)} V(t) dt\right)^{-1/n} s^{1/(n-1)} V(s) \\
&\leq \left(\frac{n}{n-1}\right)^{-1/n} \left(\int_0^s t^{1/(n-1)}\, dt\right)^{-1/n} s^{1/(n-1)} V(s)^{1-\frac{1}{n}} \\
&= V(s)^{1-\frac{1}{n}} = F'(s)
\end{aligned}$$

Clearly, (35) easily follows. Hence, (34) is true, and then, by combining (33) and (34) we get that (32) is also true. This proves the proposition. □

The Cartan-Hadamard conjecture is a difficult conjecture which, up to now, has been proved to be true only for 2-, 3-, and 4-dimensional Cartan-Hadamard manifolds. Such results are given here without any proof, apart for the 4-dimensional case due to Croke [28] that we discuss. The 2-dimensional case is due to Weil [76]. The 3-dimensional case of the conjecture is due to Kleiner [54]. The 4-dimensional case of the conjecture is due to Croke [28]. In his proof Croke gets an explicit value for the constant which turns out to be the sharp constant when $n = 4$. For $n \geq 3$, let

$$C(n) = \frac{\omega_{n-2}^{n-2}}{\omega_{n-1}^{n-1}} \left(\int_0^{\pi/2} \cos^{n/(n-2)}(t) \sin^{n-2}(t) dt \right)^{n-2} \tag{36}$$

where ω_n denotes the volume of the standard unit sphere $(\mathbb{S}^n, h)$ of $\mathbb{R}^{n+1}$. As one can easily check, $C(4)^{1/4} = K(4,1)$. A combination of Croke [28], Kleiner [54], and Weil [76] results give the following.

Theorem 11.1 *The Cartan-Hadamard conjecture is true in dimensions* $n = 2, 3, 4$. *Moreover, if* (M, g) *is a* n*-dimensional Cartan-Hadamard manifold of dimension* $n \geq 5$, *then*

$$\frac{Area_g(\partial\Omega)}{Vol_g(\Omega)^{1-\frac{1}{n}}} \geq \frac{1}{C(n)^{\frac{1}{n}}} \tag{37}$$

for all smooth, bounded domain Ω *in* M, *and* $C(n)$ *is as in (36).*

Proof. We restrict ourselves to the proof of Croke [28] that (37) holds true in dimension $n \geq 3$. The main tool in the proof is a formula due to Santalo [67]. Let Ω be a smooth, bounded domain in M. Every geodesic ray in Ω minimizes length up to the point it hits the boundary. Let $\Pi : U\Omega \to \Omega$ represent the unit sphere bundle with the canonical (local product) measure. For $v \in U\Omega$, let γ_v be the geodesic with $\gamma_v'(0) = v$ and let $\xi^t(v)$ represent the geodesic flow, that is, $\xi^t(v) = \gamma_v'(t)$. For $v \in U\Omega$, we let

$$l(v) = \max\left\{ t \,/\, \gamma_v(t) \in \Omega \right\}$$

For $x \in \partial\Omega$, we define N_x as the inwardly pointing unit normal vector to $\partial\Omega$ at x. In addition, let $\Pi : U^+\partial\Omega \to \partial\Omega$ be the bundle of inwardly pointing vectors, that is,

$$U^+\partial\Omega = \left\{ u \in U\Omega \,/\, \Pi(u) \in \partial\Omega, \langle u, N_{\Pi(u)} \rangle > 0 \right\}$$

By Santalo's formula, one has that for all integrable functions f,

$$\int_{U\Omega} f(u) du = \int_{U^+\partial\Omega} \left(\int_0^{l(u)} f\left(\xi^t(u)\right) \cos(u) dt \right) du$$

where $\cos(u)$ represents $\langle u, N_{\Pi(u)} \rangle$, and the measure on $U^+\partial\Omega$ is the local product measure du where the measure of the fiber is that of the unit upper hemisphere. From this formula, one gets that

$$\mathrm{Vol}_g(\Omega) = \frac{1}{\omega_{n-1}} \int_{U^+\partial\Omega} l(u) \cos(u) du$$

Moreover, one can prove that

$$\int_{U^+\partial\Omega} \frac{l(u)^{n-1}}{\cos(\mathrm{ant}(u))}\, du \le \mathrm{Area}_g(\partial\Omega)^2$$

and that

$$\int_{U^+\partial\Omega} \cos^{\frac{1}{n-2}}(\mathrm{ant}(u)) \cos^{\frac{n-1}{n-2}}(u) du \le A(n)\mathrm{Area}_g(\partial\Omega)$$

where

$$A(n) = \omega_{n-2} \int_0^{\frac{\pi}{2}} \cos^{n/(n-2)}(t) \sin^{n-2}(t) dt$$

and $\mathrm{ant}(u) = -\gamma_u'(l(u))$. We refer to Croke [28] for such assertions. By Hölder's inequality, one has that

$$\begin{aligned}
\mathrm{Vol}_g(\Omega) &= \frac{1}{\omega_{n-1}} \int_{U^+\partial\Omega} l(u)\cos(u) du \\
&= \frac{1}{\omega_{n-1}} \int_{U^+\partial\Omega} \frac{l(u)}{\cos^{\frac{1}{n-1}}(\mathrm{ant}(u))} \cos^{\frac{1}{n-1}}(\mathrm{ant}(u)) \cos(u) du \\
&\le \frac{1}{\omega_{n-1}} \left(\int_{U^+\partial\Omega} \frac{l(u)^{n-1}}{\cos(\mathrm{ant}(u))}\, du \right)^{\frac{1}{n-1}} \\
&\quad \times \left(\int_{U^+\partial\Omega} \cos^{\frac{1}{n-2}}(\mathrm{ant}(u)) \cos^{\frac{n-1}{n-2}}(u) du \right)^{\frac{n-2}{n-1}}
\end{aligned}$$

Hence,

$$\mathrm{Vol}_g(\Omega) \le \frac{1}{\omega_{n-1}} \mathrm{Area}_g(\partial\Omega)^{\frac{2}{n-1}} A(n)^{\frac{n-2}{n-1}} \mathrm{Area}_g(\partial\Omega)^{\frac{n-2}{n-1}}$$

and one gets that

$$\frac{\mathrm{Area}_g(\partial\Omega)}{\mathrm{Vol}_g(\Omega)^{1-\frac{1}{n}}} \ge \frac{1}{C(n)^{\frac{1}{n}}}$$

This proves (37) holds true in dimension $n \ge 3$. □

Recent advances on the Cartan-Hadamard conjecture have been made by Druet [35]. In his work, Druet proves the conjecture for small domains under the sole assumption that the scalar curvature should be negative. Druet's result is more general. We refer to the original reference Druet [35] for more information. Related references are by Johnson and Morgan [53] and Yau [77]. We refer also to Druet [34, 36]. Now, one can ask what happens to the sharp optimal inequality if we do not ask the manifold to be of nonpositive curvature. Rigidity holds in that case. The following proposition has been extended by Ledoux [55] to the entire scale of Sobolev inequalities.

Proposition 11.2 *Let (M, g) be a smooth, complete n-dimensional Riemannian manifold with nonnegative Ricci curvature. Suppose that for any smooth, bounded domain Ω in M, the sharp isoperimetric inequality (31) holds true. Then (M, g) is isometric to the Euclidean space $(\mathbb{R}^n, \xi)$ of the same dimension.*

Proof. Let $V(s)$ be the volume of $B_s(x_0)$ with respect to g, where $B_s(x_0)$ stands for the ball of center x_0 and radius s. Then

$$\frac{dV(s)}{ds} = \text{Area}_g\left(\partial B_s(x_0)\right)$$

Setting $\Omega = B_s(x_0)$ in the isoperimetric inequality, we then get that

$$\frac{1}{K(n,1)} V(s)^{(n-1)/n} \leq \frac{dV(s)}{ds}$$

for all s. Integrating yields

$$V(s) \geq \frac{1}{n^n K(n,1)^n} s^n$$

and since $K(n,1) = \frac{1}{n}(Vol_\xi(B_1(0)))^{-1/n}$, one gets that for every x_0 and for every s,

$$\text{Vol}_g\left(B_s(x_0)\right) \geq \text{Vol}_\xi\left(B_s(0)\right)$$

where $B_s(0)$ is the ball of center 0 and radius s in the Euclidean space $(\mathbb{R}^n, \xi)$. Under the assumption that (M, g) has nonnegative Ricci curvature, one gets from Gromov's comparison theorem that for every x_0 and every s,

$$\text{Vol}_g\left(B_s(x_0)\right) \leq \text{Vol}_\xi\left(B_s(0)\right)$$

Hence, for every x_0 and every s,

$$\text{Vol}_g\left(B_s(x_0)\right) = \text{Vol}_\xi\left(B_s(0)\right)$$

and one gets from the case of equality in Bishop's comparison theorem that (M, g) is isometric to the Euclidean space $(\mathbb{R}^n, \xi)$. □

As a remark, H_1^1-spaces and isoperimetric inequalities are closely related to BV-spaces. Possible references on BV-spaces are Ambrosio, Fusco, and Pallara [4], Evans and Gariepy [39], and Giusti [44].

Bibliography

[1] R. A. Adams: *Sobolev spaces* (Pure and Applied Mathematics, Vol. 65. Academic Press, New York-London, 1975)

[2] D. R. Adams: A sharp inequality of J. Moser for higher order derivatives *Ann. of Math.* **128** (1988) 385–398

[3] Adimurthi, O. Druet: Blow up analysis in dimension 2 and a sharp form of Trudinger-Moser inequality *Comm. PDE's* **29** (2004) 295–322

[4] L. Ambrosio, N. Fusco, D. Pallara: *Functions of Bounded Variation and Free Discontinuity Problems* (Oxford University Press, 2000)

[5] T. Aubin: Espaces de Sobolev sur les variétés Riemanniennes *Bull. Sci. Math.* **100** (1976) 149–173

[6] T. Aubin: Problèmes isopérimétriques et espaces de Sobolev *C. R. Acad. Sci. Paris* **280** (1974) 347–371

[7] T. Aubin: Problèmes isopérimétriques et espaces de Sobolev *J. Differential Geometry* **11** (1976) 573–598

[8] T. Aubin: Equations différentielles non linéaires et problème de Yamabe concernant la courbure scalaire *J. Math. Pures Appl.* **55** (1976) 269–296

[9] T. Aubin: *Nonlinear analysis on manifolds and Monge-Ampère equations* (Grundlehren der mathematischen Wissenschaften, Springer-Verlag, 252, 1982)

[10] T. Aubin, O. Druet, E. Hebey: Best constants in Soboelv inequalities for compact manifolds of nonpositive curvature *C. R. Acad. Sci. Paris* **326** (1998) 1117–1121

[11] T. Aubin, Y-Y. Li: On the best Sobolev inequality *J. Math. Pures Appl.* **78** (1999) 353–387

[12] D. Bakry: L'hypercontractivité et son utilisation en théorie des semigroupes *In: Lectures on probability theory (Saint-Flour, 1992)* (Lecture Notes in Mathematics, 1581, Springer, Berlin, 1994) 1–114

[13] D. Bakry, T. Coulhon, M. Ledoux and L. Saloff-Coste: Sobolev inequalities in disguise *Indiana Univ. Math. J.* **44** (1995) 1033–1074

[14] W. Beckner: Geometric proof of Nash's inequality *Int. Math. Res. Notices* (1998) 67–72

[15] W. Beckner: Asymptotic estimates for Gagliardo-Nirenberg embedding constants *Potential Analysis* **17** (2002) 253–266

[16] G. A. Bliss: An integral inequality *J. London Math. Soc.* **5** (1930) 40–46

[17] M. Cantor: Sobolev inequalities for Riemannian bundles *Bull. Amer. math. Soc.* **80** (1974) 239–243

[18] E. A. Carlen, and M. Loss: Sharp constant in Nash's inequality *Internat. Math. Res. Notices* **7** (1993) 213–215

[19] G. Carron: Inégalités isopérimétriques de Faber-Krahn et conséquences *In: Sémin. Congr., 1, Soc. Math.* (France, Paris, 1996) 205–232

[20] G. Carron: *Inégalités isopérimétriques sur les variétés riemanniennes* (Thèse de doctorat, Université Joseph Fourier, Grenoble, 1994)

[21] S.-Y. A. Chang: The Moser-Trudinger inequality and applications to some problems in conformal geometry *In: Nonlinear partial differential equations in differential geometry* ((Park City, UT, 1992), IAS/Park City Math. Ser., 2, Amer. Math. Soc., Providence, RI, 1996) 65–125

[22] I. Chavel: *Riemannian geometry-a modern introduction* (Cambridge Tracts in Mathematics, 108. Cambridge University Press, Cambridge, 1993)

[23] S. Y. Cheng, S. T. Yau: Differential equations on Riemannian manifolds and their geometric applications *Comm. Pure Appl. Math.* **28** (1975) 333–354

[24] D. Cordero-Erausquin, B. Nazaret, C. Villani: A new approach to sharp Sobolev and Gagliardo-Nirenberg inequalities *Adv. Math.* **182** (2004) 307–332

[25] T. Coulhon: Espaces de Lipschitz et inégalités de Poincaré *J. Funct. Anal.* **136** (1996) 81–113

[26] T. Coulhon, M. Ledoux: Isopérimétrie, décroissance du noyau de la chaleur et transformations de Riesz: un contre-exemple *Ark. Mat.* **32** (1994) 63–77

[27] T. Coulhon, L. Saloff-Coste: Isopérimétrie pour les groupes et les variétés *Rev. Mat. Iberoamericana* **9** (1993) 293–314

[28] C. B. Croke: A sharp four-dimensional isoperimetric inequality *Comment. Math. Helv.* **59** (1984) 187–192

[29] M. Del Pino, J. Dolbeault: Best constants for Gagliardo-Nirenberg inequalities and applications to nonlinear diffusions *J. Math. Pures Appl.* **81** (2002) 847–875

[30] M. P. do Carmo: *Riemannian geometry* (Mathematics: Theory & Applications. Birkhauser Boston, Inc., Boston, MA, 1992)

[31] O. Druet: Generalized scalar curvature type equations on compact Riemannian manifolds *Proc. Roy. Soc. Edinburgh Sect. A* **130** (2000) 767–788

[32] O. Druet: The best constants problem in Sobolev inequalities *Math. Ann.* **314** (1999) 327–346

[33] O. Druet: Optimal Sobolev inequalities of arbitrary order on compact Riemannian manifolds *J. Funct. Anal.* **159** (1998) 217–242

[34] O. Druet: Isoperimetric inequalities on compact manifolds *Geom. Dedicata* **90** (2002) 217–236

[35] O. Druet: Sharp local isoperimetric inequalities involving the scalar curvature *Proc. A.M.S.* **130** (2002) 2351–2361

[36] O. Druet: Inégalités de Sobolev optimales et inégalités isopérimétriques sur les variétés *In: Sémin. Théor. Spectr. Géom.* (20, 2001/2002) 23–100

[37] O. Druet, E. Hebey: The AB program in geometric anaysis: sharp Sobolev inequalities and related problems *Memoirs of the AMS* **761** (2002) 98

[38] O. Druet, E. Hebey, F. Robert: *Blow-up theory for elliptic PDEs in Riemannian geometry* (Mathematical Notes, 45, Princeton University Press, Princeton, NJ, 2004)

[39] L. C. Evans, R. F. Gariepy: *Measure theory and fine properties of functions* (Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992)

[40] H. Federer, W. H. Fleming: Normal and integral currents *Ann. of Math.* **72** (1960) 458–520

[41] L. Fontana: Sharp borderline Sobolev inequalities on compact Riemannian manifolds *Comment. Math. Helv.* **68** (1993) 415–454

[42] E. Gagliardo: Proprietà di alcune classi di funzioni in più variabili *Ricerche Mat.* **7** (1958) 102–137

[43] S. Gallot, D. Hulin, J. Lafontaine: *Riemannian geometry* (Universitext. Springer-Verlag, Berlin, 2004)

[44] E. Giusti: *Minimal surfaces and functions of bounded variation* (Monographs in Mathematics, Birkhäuser, 1984)

[45] A. A. Grigor′yan: The existence of positive fundamental solutions of the Laplace equation on Riemannian manifolds *Mat. Sb. (N.S.)* **128(170)** (1985) 354–363; *Mat. USSR Sb.* **56** (1987) 349–358

[46] E. Hebey: *Introduction à l'analyse non-linéaire sur les variétés* (Diderot éditeur, Paris, 1997)

[47] E. Hebey: *Nonlinear analysis on manifolds: Sobolev spaces and inequalities* (Courant Lecture Notes in Mathematics, 5. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 1999)

[48] E. Hebey, M. Vaugon: Meilleures constantes dans le théorème d'inclusion de Sobolev et multiplicité pour les problèmes de Nirenberg et Yamabe *Indiana Univ. Math. J.* **41** (1992) 377–407

[49] E. Hebey, M. Vaugon: The best constant problem in the Sobolev embedding theorem for complete Riemannian manifolds *Duke Math. J.* **79** (1995) 235–279
[50] E. Hebey, M. Vaugon: Meilleures constantes dans le théorème d'inclusion de Sobolev *Ann. Inst. H. Poincaré Anal. Non Linéaire* **13** (1996) 57–93
[51] D. Hoffman, and J. Spruck: Sobolev and isoperimetric inequalities for Riemannian submanifolds *Comm. Pure Appl. Math.* **27** (1974) 715–727
[52] S. Ilias: Constantes explicites pour les inégalités de Sobolev sur les variétés riemanniennes compactes *Ann. Inst. Fourier (Grenoble)* **33** (1983) 151–165
[53] D. Johnson, F. Morgan: Some sharp isoperimetric theorems for Riemannian manifolds *Indiana Univ. Math. J.* **49** (2000) 1017–1041
[54] B. Kleiner: An isoperimetric comparison theorem *Invent. Math.* **108** (1992) 37–47
[55] M. Ledoux: On manifolds with nonnegative Ricci curvature and Sobolev inequalities *Comm. Anal. Geom.* **7** (1999) 347–353
[56] J. M. Lee, T. H. Parker: The Yamabe problem *Bull. Amer. Math. Soc. (N.S.)* **17** (1987) 37–91
[57] V. G. Maz′ja: *Sobolev spaces* (Springer Series in Soviet Mathematics. Springer, Berlin–New York, 1985)
[58] N. G. Meyers, J. Serrin: $H = W$ *Proc. Nat. Acad. Sci. U.S.A.* **51** (1964) 1055–1056
[59] C. B. Jr. Morrey: Functions of several variables and absolute continuity, II. *Duke Math. J.* **6** (1940) 187–215
[60] J. Moser: A sharp form of an inequality by N.Trudinger *Ind. U. Math. J.* **20** (1971) 1077–1092
[61] J. Nash: Continuity of solutions of parabolic and elliptic equations *Amer. J. Math.* **80** (1958) 931–954
[62] L. Nirenberg: On elliptic partial differential equations *Ann. Scuola Norm. Sup. Pisa* **13** (1959) 115–162
[63] S. I. Pohožaev: On the eigenfunctions of the equation $\Delta u + \lambda f(u) = 0$ *Soviet Mathematics Doklady* **6** (1965) 1408–14011
[64] E. Rodemich: *The Sobolev inequalities with best possible constants* (Analysis Seminar at California Institute of Technology, 1966)
[65] T. Sakai: *Riemannian geometry* (Translations of Mathematical Monographs, 149. American Mathematical Society, Providence, RI, 1996)
[66] L. Saloff-Coste: *Aspects of Sobolev type inequalities* (London Mathematical Society Lecture Notes Series 289, Cambridge University Press, Cambridge, 2002)
[67] L. A. Santalo: *Integral geometry and geometric probability* (Encyclopedia of Mathematics and Its Applications, Vol. 1. Addison-Wesley, Reading, Mass.–London–Amsterdam, 1976)
[68] S. L. Sobolev: Sur un théorème d'analyse fonctionnelle *Rec. Math. Moscou, n. Ser.* **4** (1938) 471–497
[69] M. Spivak: *A comprehensive introduction to differential geometry* (Second edition. Publish or Perish, Inc., Wilmington, Del., 1979)
[70] M. Struwe: Critical points of embeddings of $H_0^{1,n}$ into Orlicz spaces *Ann. Inst. H. Poincaré Anal. Non Linéaire* **5** (1988) 425–464
[71] G. Talenti: Best constant in Sobolev inequality *Ann. Mat. Pura Appl.* **110** (1976) 353–372

[72] N. Trudinger: On embeddings into Orlicz spaces and some applications *J. Math. Mech.* **17** (1977) 473–483

[73] N. Varopoulos: Small time Gaussian estimates of heat diffusion kernels. I. The semigroup technique *Bull. Sci. Math.* **113** (1989) 253–277

[74] N. Varopoulos: Potential theory and diffusion on Riemannian manifolds *In: Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II (Chicago, Ill., 1981)* (Wadsworth Math. Ser., Wadsworth, Belmont, Calif., 1983) 821–837

[75] N. Varopoulos: Hardy-Littlewood theory for semigroups *J. Funct. Anal.* **63** (1985) 240–260

[76] A. Weil: Sur les surfaces à courbure négative *C. R. Acad. Sci. Paris Sér. I Math.* **182** (1926) 4069–4071

[77] S. T. Yau: Isoperimetric constants and the first eigenvalue of a compact manifold *Ann. Sci. École Norm. Sup.* **8** (1975) 487–507

EMMANUEL HEBEY

Université de Cergy-Pontoise, Département de Mathématiques, Site de Saint-Martin, 2 avenue Adolphe Chauvin, 95302 Cergy-Pontoise cedex, France

E-mail: `Emmanuel.Hebey@math.u-cergy.fr`

FRÉDÉRIC ROBERT

Laboratoire J.A.Dieudonné, Université de Nice Sophia-Antipolis, Parc Valrose, 06108 Nice cedex 2, France

E-mail: `frobert@math.unice.fr`

This page intentionally left blank

Handbook of Global Analysis
Demeter Krupka and David Saunders

Harmonic maps

Dedicated to the memory of James Eells

Frédéric Hélein and John C. Wood

Contents

Introduction

The subject of harmonic maps is vast and has found many applications, and it would require a very long book to cover all aspects, even superficially. Hence, we have made a choice; in particular, highlighting the key questions of *existence*, *uniqueness* and *regularity* of harmonic maps between given manifolds. Thus we shall survey some of the main methods of global analysis for answering these questions.

We first consider relevant aspects of harmonic functions on Euclidean space; then we give a general introduction to harmonic maps. The core of our work is in Sections 3–6 where we present the analytical methods. We round of the article by describing how twistor theory and integrable systems can be used to construct many more harmonic maps. On the way, we mention harmonic morphisms: maps between Riemannian manifolds which preserve Laplace's equation; these turn out to be a particular class of harmonic maps and exhibit some properties dual to those of harmonic maps.

More information on harmonic maps can be found in the following articles and books; for generalities: [61, 62, 63, 219], analytical aspects: [21, 88, 103, 118, 131, 133, 135, 189, 204, 194], integrable systems methods: [73, 94, 117], applications to complex and Kähler geometry: [63, 135], harmonic morphisms: [7], and other topics: [64, 231].

1 Harmonic functions on Euclidean spaces

Harmonic functions on an open domain Ω of $\mathbb{R}^m$ are solutions of the *Laplace equation*

$$\Delta f = 0, \quad \text{where } \Delta := \frac{\partial^2}{(\partial x^1)^2} + \cdots + \frac{\partial^2}{(\partial x^m)^2} \qquad \big((x^1, \cdots, x^m) \in \Omega\big). \tag{1}$$

The operator Δ is called the *Laplace operator* or *Laplacian* after P.-S. Laplace. Equation (1) and the *Poisson equation*[1] $-\Delta f = g$ play a fundamental role in mathematical physics: the Laplacian occurs in *Newton's law of gravitation* (the gravitational potential U obeys the law $-\Delta U = -4\pi G\rho$, where ρ is the mass density), *electromagnetism* (the electric potential V is a solution of $-\varepsilon_0 \Delta V = \rho$, where ρ is the electric charge distribution), *fluid mechanics* (the right hand side term in the Navier–Stokes system $\frac{\partial u^i}{\partial t} + u^j \frac{\partial u^i}{\partial x^j} + \frac{\partial p}{\partial x^i} = \nu \Delta u^i$ models the effect of the viscosity), and the *heat equation* $\frac{\partial f}{\partial t} = \Delta f$.

The *fundamental solution* $G = G_m$ of the Laplacian is the solution of the Poisson equation $-\Delta G = \delta$ on $\mathbb{R}^m$, where δ is the Dirac mass at the origin, that has the mildest growth at infinity, i.e. $G_2(x) = (2\pi)^{-1} \log(1/r)$ if $m = 2$ and $G_m(x) = 1/\{(m-2)\,|S^{m-1}|\, r^{m-2}\}$ if[2] $m \geq 1$ and $m \neq 2$.

1.1 The Dirichlet principle

The harmonic functions are critical points (also called *extremals*) of the Dirichlet functional

$$E_\Omega(f) := \frac{1}{2} \int_\Omega \sum_{\alpha=1}^m \left(\frac{\partial f}{\partial x^\alpha}(x) \right)^2 d^m x = \frac{1}{2} \int_\Omega |df_x|^2 d^m x,$$

where $d^m x := dx^1 \cdots dx^m$. This comes from the fact that, for any smooth function g with compact support in Ω, the *first variation* $(\delta E_\Omega)_f(g) := \lim_{\varepsilon \to 0} \{E_\Omega(f + \varepsilon g) - E_\Omega(f)\}/\varepsilon$ reads

$$(\delta E_\Omega)_f(g) = \int_\Omega \sum_{\alpha=1}^m \frac{\partial f}{\partial x^\alpha} \frac{\partial g}{\partial x^\alpha} d^m x = \int_\Omega (-\Delta f) g \, d^m x. \tag{2}$$

This variational formulation (G. Green, 1833; K.F. Gauss, 1837; W. Thomson, 1847; B. Riemann, 1853) reveals that the Laplace operator depends on the (canonical) metric on $\mathbb{R}^m$, since $|df_x|$ is nothing but the Euclidean norm of $df_x \in (\mathbb{R}^m)^*$.

This leads to a strategy to solve the **Dirichlet problem**: given an open bounded subset Ω of $\mathbb{R}^m$ with smooth boundary $\partial\Omega$ and a continuous function $\gamma : \partial\Omega \longrightarrow \mathbb{R}$, find a continuous function $f : \overline{\Omega} \longrightarrow \mathbb{R}$, smooth in Ω, such that

$$\Delta f = 0 \quad \text{in } \Omega, \quad \text{and} \quad f = \gamma \quad \text{on } \partial\Omega. \tag{3}$$

The idea to solve (3), named the **Dirichlet principle** by Riemann or the **direct method of the calculus of variations**, is the following: we consider the class of functions $\mathcal{D}_\gamma(\Omega) :=$

[1] We prefer to put a minus sign in front of Δ, since the operator $-\Delta$ has many positivity properties.

[2] Here $|S^{m-1}| = 2\pi^{m/2}/\,\Gamma(m/2)$ is the $(m-1)$-dimensional Hausdorff measure of the unit sphere S^{m-1}.

$\{f \in \mathcal{C}^2(\Omega) \cap \mathcal{C}^0(\overline{\Omega}) |\, f = \gamma \text{ on } \partial\Omega\}$ and we look for a map $\underline{f} \in \mathcal{D}_\gamma(\Omega)$ which minimizes E_Ω among all maps in $\mathcal{D}_\gamma(\Omega)$. *If we can prove the existence of a such a minimizer* $\underline{f}$ in $\mathcal{D}_\gamma(\Omega)$, then by (2), $\underline{f}$ is a critical point of E_Ω and is a solution of the Dirichlet problem (3). The difficulty was to prove *the existence* of a minimizer. Riemann was confident that there was such a minimizer, although K. Weierstrass proved that the method proposed at that time had a gap and many people had given up with this formal idea. Then D. Hilbert proposed in 1900 to replace $\mathcal{D}_\gamma(\Omega)$ by a larger class and this led to a definitive solution formulated by H. Weyl in 1940 [223].

1.2 Existence of solutions to the Dirichlet problem

Several methods may be used to solve the Dirichlet problem including the 'balayage' method by H. Poincaré [173], and the use of sub- and super-solutions by O. Perron [166], see [90]. But the variational approach seems to be the most robust one to generalize to finding harmonic maps between manifolds.

The modern variational proof for the existence of solutions to (3) uses the *Sobolev space* $W^{1,2}(\Omega)$: the set of (classes of) functions f in $L^2(\Omega)$ whose derivatives $\partial f/\partial x^j$ in the distribution sense are in $L^2(\Omega)$. When endowed with the inner product $\langle f, g\rangle_{W^{1,2}} := \int_\Omega \big(fg + \langle df, dg\rangle\big) d^m x$ and norm $||f||^2_{W^{1,2}} := \langle f, f\rangle_{W^{1,2}}$, the space $W^{1,2}(\Omega)$ is a Hilbert space. An important technical point is that $\mathcal{C}^\infty(\overline{\Omega})$ is dense in $W^{1,2}(\Omega)$. Assuming that the boundary $\partial\Omega$ is smooth, there is a unique linear continuous operator defined on $W^{1,2}(\Omega)$ which extends the trace operator $f \longmapsto f|_{\partial\Omega}$ from $\mathcal{C}^\infty(\overline{\Omega})$ to $\mathcal{C}^\infty(\partial\Omega)$. Its image is the Hilbert space $W^{\frac{1}{2},2}(\partial\Omega)$ of (classes of) functions γ in $L^2(\partial\Omega)$ such that $\int_{\partial\Omega}\int_{\partial\Omega} \big(\gamma(x) - \gamma(y)\big)^2/|x-y|^m d\mu(x) d\mu(y) < +\infty$, where $d\mu$ denotes the measure on $\partial\Omega$. So the Dirichlet problem makes sense if the boundary data γ belongs to $W^{\frac{1}{2},2}(\partial\Omega)$, and if we look for f in $W^{1,2}(\Omega)$. Inspired by the Dirichlet principle we define the class $W^{1,2}_\gamma(\Omega) := \{f \in W^{1,2}(\Omega) |\, u|_{\partial\Omega} = \gamma\}$ and we look for a map $f \in W^{1,2}_\gamma(\Omega)$ which minimizes E_Ω: it will be a *weak solution* of the Dirichlet problem.

The solution of this problem when Ω is *bounded* comes from the following. First one chooses a map $f_\gamma \in W^{1,2}_\gamma(\Omega)$, so that $\forall f \in W^{1,2}_\gamma(\Omega)$, $f - f_\gamma \in W^{1,2}_0(\Omega)$. But since Ω is bounded, functions g in $W^{1,2}_0(\Omega)$ obey the *Poincaré inequality* $||g||_{W^{1,2}} \leq C_P ||dg||_{L^2}$. This implies the bound $||f||_{W^{1,2}} \leq ||f_\gamma||_{W^{1,2}} + C_P\sqrt{2E_\Omega(f_\gamma)}$ for any $f \in W^{1,2}_\gamma(\Omega)$. A consequence is that $||f||_{W^{1,2}}$ *is bounded as soon as* $E_\Omega(f)$ *is bounded*. Now we are ready to study a minimizing sequence $(f_k)_{k\in\mathbb{N}}$, i.e. a sequence in $W^{1,2}_\gamma(\Omega)$ such that

$$\lim_{k\to\infty} E_\Omega(f_k) = \inf_{W^{1,2}_\gamma(\Omega)} E_\Omega. \tag{4}$$

Because $E_\Omega(f_k)$ is obviously bounded, $||f_k||_{W^{1,2}}$ is also bounded, so that f_k takes values in a compact subset of $W^{1,2}_\gamma(\Omega)$ for the weak $W^{1,2}$-topology. Hence, because of the compactness[3] of the embedding $W^{1,2}(\Omega) \subset L^2(\Omega)$, we can assume that, after extracting a subsequence if necessary, there exists $\underline{f} \in W^{1,2}_\gamma(\Omega)$ such that $f_k \to \underline{f}$ *weakly* in $W^{1,2}$, strongly in L^2 and a.e. on Ω. We write $f_k = \underline{f} + g_k$, so that $g_k \to 0$ *weakly* in $W^{1,2}$, and from the identity $E_\Omega(f_k) = E_\Omega(\underline{f}) + E_\Omega(g_k) + \int_\Omega \langle d\underline{f}, dg_k\rangle$ we obtain

$$\limsup_{k\to\infty} E_\Omega(f_k) = E_\Omega(\underline{f}) + \limsup_{k\to\infty} E_\Omega(g_k). \tag{5}$$

[3]By the Rellich–Kondrakov theorem, valid here because Ω is bounded.

Hence $\limsup_{k\to\infty} E_\Omega(f_k) \geq E_\Omega(\underline{f})$, i.e. E_Ω is *lower semi-continuous*. Comparing (4) and (5) we obtain

$$\left(E_\Omega(\underline{f}) - \inf_{W^{1,2}_\gamma(\Omega)} E_\Omega\right) + \limsup_{k\to\infty} E_\Omega(g_k) = 0.$$

Both terms in this equation are non-negative, hence must vanish: this tells us that $\underline{f}$ is a minimizer of E_Ω in $W^{1,2}_\gamma(\Omega)$ and *a posteriori* that $g_k \to 0$ *strongly* in $W^{1,2}$, i.e. $f_k \to \underline{f}$ *strongly* in $W^{1,2}$.

Hence we have obtained a *weak solution* to the Dirichlet problem. It remains to show that this solution is *classical*, i.e. that $\underline{f}$ is smooth in Ω and that, if γ is continuous, then $\underline{f}$ is continuous on $\overline{\Omega}$ and agrees with γ on $\partial\Omega$. This is the *regularity problem*. Several methods are possible: one may for instance deduce the interior regularity from the identity $\underline{f} = \underline{f} * \chi_\rho$ which holds on $\{x \in \Omega | \, B(x,\rho) \subset \Omega\}$, where $\chi_\rho \in \mathcal{C}^\infty_c(\mathbb{R}^m)$ is rotationally symmetric, has support in $B(0,\rho)$ and satisfies $\int_{\mathbb{R}^m} \chi_\rho = 1$ and $*$ denotes the *convolution* operator given by $f * g(x) = \int_{\mathbb{R}^m} f(x-y)g(y)d^m y$. This identity is actually a version of the mean value property (see the next paragraph) valid for weak solutions.

1.3 The mean value property and the maximum principle

Let f be a harmonic function on an open subset Ω, $x_0 \in \Omega$ and $\rho > 0$ such that $B(x_0,\rho) \subset \Omega$. Stokes' theorem gives: $\forall r \in (0,\rho]$, $\int_{\partial B(x_0,r)} (\partial f/\partial r)\, d\mu(x) = \int_{B(x_0,r)} \Delta f d^m x = 0$, where $r = |x - x_0|$. It implies that $⨍_{\partial B(x_0,r)} f := 1/(|S^{m-1}|\, r^{m-1}) \int_{\partial B(x_0,r)} f \, d\mu(x)$ is independent of r. Hence, since f is continuous at x_0, we have $f(x_0) = ⨍_{\partial B(x_0,r)} f$. By averaging further over all spheres $\partial B(x_0,r)$ with $0 < r < \rho$, one deduces that $f(x_0) = ⨍_{B(x_0,r)} f$.

A similar argument works for superharmonic or subharmonic functions: a smooth function $f : \Omega \longrightarrow \mathbb{R}$ is *superharmonic* (resp. *subharmonic*) if and only if $-\Delta f \geq 0$ (resp. $-\Delta f \leq 0$). Then, if f superharmonic (resp. subharmonic) and $B(x_0,\rho) \subset \Omega$, we have $f(x_0) \geq ⨍_{B(x_0,r)} f$ (resp. $f(x_0) \leq ⨍_{B(x_0,\rho)} f$).

The mean value property implies the *maximum* and *minimum principles*: assume that Ω is open, bounded and connected and that f is harmonic on Ω, continuous on $\overline{\Omega}$ and that $x_0 \in \Omega$ is an *interior*(!) point where f is maximal, i.e., $\forall x \in \Omega$, $f(x) \leq f(x_0)$. Then we choose $B(x_0,\rho) \subset \Omega$ and, by the mean value property, $f(x_0) = ⨍_{B(x_0,\rho)} f$ or, equivalently, $\int_{B(x_0,\rho)} (f(x_0) - f(x))\, d^m x = 0$. But since f is maximal at x_0 the integrand in this last integral is non-negative and hence must vanish. Thus $f(x) = f(x_0)$ on $B(x_0,\rho)$. So we have shown that $(f|_\Omega)^{-1}(\sup_\Omega f) := \{x \in \Omega | \, f(x) = \sup_\Omega f\}$ is open. It is also closed because f is continuous. Hence since Ω is connected, **either** $(f|_\Omega)^{-1}(\sup_\Omega f) = \Omega$ **and** f **is constant, or** $(f|_\Omega)^{-1}(\sup_\Omega f) = \emptyset$**, which means that** $\sup_\Omega f$ **is achieved on the boundary** $\partial\Omega$ (since $\overline{\Omega}$ is compact). This is the *(strong) maximum principle*.

One sees that the preceding argument still works if we replace the property $f(x_0) = ⨍_{B(x_0,\rho)} f$ by $f(x_0) \leq ⨍_{B(x_0,\rho)} f$, i.e. if we only assume that f is *subharmonic*. Similarly the *minimum principle* works for *superharmonic* functions.

1.4 Uniqueness and minimality

The uniqueness of solutions to the Dirichlet problem can be obtained as a consequence of the maximum principle: let f_1 and f_2 be two solutions of the Dirichlet problem and let $f := f_2 - f_1$. Since f_1 agrees with f_2 on $\partial\Omega$, the trace of f on $\partial\Omega$ vanishes. But f is also harmonic, and hence satisfies the maximum principle: this implies that $\sup_\Omega f = \sup_{\partial\Omega} f = 0$, so $f \leq 0$ on Ω. Similarly, the minimum principle implies $f \geq 0$ on Ω. Hence $f = 0$, which means that f_1 coincides with f_2.

A straightforward consequence of this uniqueness result is that any solution f of (3) actually coincides with **the** minimizer of E_Ω in $W^{1,2}_\gamma(\Omega)$. One can recover this minimality property directly from the identity

$$\forall g \in W^{1,2}_\gamma(\Omega), \quad E_\Omega(g) = E_\Omega(g-f) + \int_\Omega \langle df, dg\rangle - E_\Omega(f).$$

On using Stokes' theorem twice, $\Delta f = 0$ and the fact that $f|_{\partial\Omega} = g|_{\partial\Omega}$, we obtain

$$\int_\Omega \langle df, dg\rangle = \int_\Omega \mathrm{div}(g\nabla f) = \int_{\partial\Omega} g\frac{\partial f}{\partial n} = \int_{\partial\Omega} f\frac{\partial f}{\partial n} = \int_\Omega \mathrm{div}(f\nabla f) = 2E_\Omega(f). \quad (6)$$

Hence $E_\Omega(g) = E_\Omega(g-f) + E_\Omega(f)$, which implies that f minimizes E_Ω in $W^{1,2}_\gamma(\Omega)$.

1.5 Relation with holomorphic functions

In dimension 2, harmonic functions are closely linked with holomorphic functions. Throughout this article, we shall use the identification $\mathbb{R}^2 \simeq \mathbb{C}$, $(x,y) \longmapsto x + iy$ and the operators

$$\frac{\partial}{\partial z} := \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right), \quad \frac{\partial}{\partial \overline{z}} := \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right).$$

If Ω is an open subset of $\mathbb{C}$, recall that a smooth function $\varphi : \Omega \longrightarrow \mathbb{C}$ is *holomorphic* (rep. *antiholomorphic*) if and only if $\partial\varphi/\partial\overline{z} = 0$ (resp. $\partial\varphi/\partial z = 0$). Then because of the identity $\partial^2/\partial z\partial\overline{z} = \partial^2/\partial\overline{z}\partial z = (1/4)\Delta$ it is clear that, if $\varphi : \Omega \longrightarrow \mathbb{C}$ is holomorphic or antiholomorphic, then $\mathrm{Re}\,\varphi$ and $\mathrm{Im}\,\varphi$ are harmonic functions. Conversely, if we are given a harmonic function $f : \Omega \longrightarrow \mathbb{R}$, then $\partial f/\partial z$ is holomorphic. Moreover if Ω is simply connected the holomorphic function φ defined by $\varphi(z) = 2\int_{z_0}^z \frac{\partial f}{\partial z}(\zeta)\, d\zeta$ satisfies $\frac{\partial\varphi}{\partial z} = 2\frac{\partial f}{\partial z}$ and $f = \mathrm{Re}\,\varphi + C$, where $C \in \mathbb{R}$ is a constant. The imaginary part of φ provides us with another harmonic function $g := \mathrm{Im}\,\varphi$, the *harmonic conjugate function* of f. Note that some representation formulas for harmonic functions in terms of holomorphic data have been found in dimension three (E.T. Whitakker [224]) and in dimension four (H. Bateman and R. Penrose [8, 165]).

2 Harmonic maps between Riemannian manifolds

2.1 Definition

Throughout the rest of this article, $\mathcal{M} = (\mathcal{M}, g)$ and $\mathcal{N} = (\mathcal{N}, h)$ will denote smooth Riemannian manifolds, without boundary unless otherwise indicated, of arbitrary (finite) dimensions m and n respectively. We denote their Levi-Civita connnections by ${}^g\nabla$ and ${}^h\nabla$ respectively. By an *(open) domain* of $\mathcal{M}$ we mean a non-empty connected open subset of $\mathcal{M}$; if a domain has compact closure, we shall call that closure a *compact domain*. We use the *Einstein summation convention* where summation over repeated subscript-superscript pairs is understood.

We define harmonic maps as the solution to a variational problem which generalizes that in Section 1 as follows. Let $\phi : (\mathcal{M}, g) \to (\mathcal{N}, h)$ be a smooth map. Let Ω be a domain of $\mathcal{M}$ with a piecewise $\mathcal{C}^1$ boundary $\partial\Omega$. The *energy* or *Dirichlet integral* of ϕ over Ω is defined by

$$E_\Omega(\phi) = \frac{1}{2}\int_\Omega |d\phi|^2\, \omega_g\,. \tag{7}$$

Here ω_g is the volume measure on $\mathcal{M}$ defined by the metric g, and $|d\phi|$ is the Hilbert–Schmidt norm of $d\phi$ given at each point $x \in \mathcal{M}$ by

$$|d\phi_x|^2 = \sum_{i=1}^m h_{\phi(x)}\big(d\phi_x(e_i), d\phi_x(e_i)\big) \tag{8}$$

where $\{e_i\}$ is an orthonormal basis for $T_x\mathcal{M}$. In local coordinates $(x^1, \ldots, x^m)$ on $\mathcal{M}$, $(y^1, \ldots, y^n)$ on $\mathcal{N}$,

$$|d\phi_x|^2 = g^{ij}(x)h_{\alpha\beta}(\phi(x))\,\phi_i^\alpha\phi_j^\beta \quad \text{and} \quad \omega_g = \sqrt{|g|}\,dx^1\cdots dx^m\,; \tag{9}$$

here ϕ_i^α denotes the partial derivative $\partial\phi^\alpha/\partial x^i$ where $\phi^\alpha := y^\alpha \circ \phi$, (g_{ij}) denotes the metric tensor on $\mathcal{M}$ with determinant $|g|$ and inverse (g^{ij}), and $(h_{\alpha\beta})$ denotes the metric tensor on $\mathcal{N}$.

By a *smooth (one-parameter) variation* $\Phi = \{\phi_t\}$ *of* ϕ we mean a smooth map $\Phi : \mathcal{M} \times (-\varepsilon, \varepsilon) \to \mathcal{N}$, $\Phi(x,t) = \phi_t(x)$, where $\varepsilon > 0$ and $\phi_0 = \phi$. We say that it is *supported in* Ω if $\phi_t = \phi\ \forall\, t$ on the complement of the interior of Ω. A smooth map $\phi : (\mathcal{M}, g) \to (\mathcal{N}, h)$ is called *harmonic* if it is a *critical point* (or *extremal*) of the energy integral, i.e., for all compact domains Ω and all smooth one-parameter variations $\{\phi_t\}$ of ϕ supported in Ω, the *first variation* $\frac{d}{dt}E_\Omega(\phi_t)\big|_{t=0}$ is zero. The first variation is given by

$$(\delta E_\Omega)_\phi(v) := \frac{d}{dt}E_\Omega(\phi_t)\Big|_{t=0} = -\int_{\mathcal{M}} \langle \tau(\phi), v\rangle\, \omega_g\,. \tag{10}$$

Here v denotes the *variation vector field* of $\{\phi_t\}$ defined by $v = \partial\phi_t/\partial t|_{t=0}$, $\langle\cdot,\cdot\rangle$ denotes the inner product on $\phi^{-1}T\mathcal{N}$ induced from the metric on $\mathcal{N}$, and $\tau(\phi)$ denotes the *tension field of* ϕ defined by

$$\tau(\phi) \;=\; \operatorname{Tr}\, {}^W\nabla d\phi = \sum_{i=1}^m {}^W\nabla d\phi(e_i, e_i) = \sum_{i=1}^m \big\{ {}^\phi\nabla_{e_i}\big(d\phi(e_i)\big) - d\phi({}^g\nabla_{e_i}e_i)\big\} \tag{11}$$

Here ${}^\phi\nabla$ the pull-back of the Levi-Civita connection on $\mathcal{N}$ to the bundle $\phi^{-1}T\mathcal{N}$, and ${}^W\nabla$ the tensor product connection on the bundle $W = T^*\mathcal{M} \otimes \phi^{-1}T\mathcal{N}$ induced from these connections. We see that the tension field is the trace of the *second fundamental form* of ϕ

defined by $\beta(\phi) = {}^W\nabla d\phi$, more explicitly, $\beta(\phi)(X,Y) = {}^\phi\nabla_X\big(d\phi(Y)\big) - d\phi({}^g\nabla_X Y)$ for any vector fields X, Y on $\mathcal{M}$. In local coordinates,

$$\tau(\phi)^\gamma = g^{ij}\left(\frac{\partial^2\phi^\gamma}{\partial x^i \partial x^j} - {}^g\Gamma^k_{ij}\frac{\partial\phi^\gamma}{\partial x^k} + {}^h\Gamma^\gamma_{\alpha\beta}(\phi)\frac{\partial\phi^\alpha}{\partial x^i}\frac{\partial\phi^\beta}{\partial x^j}\right) \tag{12}$$

$$= \Delta_g\phi^\gamma + g(\operatorname{grad}\phi^\alpha, \operatorname{grad}\phi^\beta)\, {}^h\Gamma^\gamma_{\alpha\beta}\,. \tag{13}$$

Here ${}^g\Gamma^k_{ij}$ and ${}^h\Gamma^\gamma_{\alpha\beta}$ denote the Christoffel symbols on $(\mathcal{M}, g)$ and $(\mathcal{N}, h)$, respectively, and Δ_g denotes the Laplace–Beltrami operator on functions $f : \mathcal{M} \to \mathbb{R}$ given by

$$\Delta_g f = \operatorname{Tr}\, {}^W\nabla df = \sum_{i=1}^m \big\{e_i\big(e_i(f)\big) - \big({}^g\nabla_{e_i}e_i\big)f\big\}, \tag{14}$$

or, in local coordinates,

$$\Delta_g f = \frac{1}{\sqrt{|g|}}\frac{\partial}{\partial x^i}\left(\sqrt{|g|}\, g^{ij}\frac{\partial f}{\partial x^j}\right) = g^{ij}\left(\frac{\partial^2 f}{\partial x^i\partial x^j} - {}^g\Gamma^k_{ij}\frac{\partial f}{\partial x^k}\right). \tag{15}$$

Note that $\tau(\phi)$ can be interpreted as the negative of the gradient at ϕ of the energy functional E on a suitable space of mappings, i.e., it points in the direction in which E decreases most rapidly [61, (3.5)]. In local coordinates, the *harmonic equation*

$$\tau(\phi) = 0 \tag{16}$$

is a *semilinear* second-order elliptic system of partial differential equations.

2.2 Examples

We list some important examples of harmonic maps. See, for example, [66, 61, 63, 7] for many more.

1. **Constant maps** $\phi : (\mathcal{M}, g) \to (\mathcal{N}, h)$ and **identity maps** $\mathrm{Id} : (\mathcal{M}, g) \to (\mathcal{M}, g)$ are clearly always harmonic maps

2. **Isometries** are harmonic maps. Further, composing a harmonic map with an isometry on its domain or codomain preserves harmonicity.

3. **Harmonic maps between Euclidean spaces**. A smooth map $\phi : A \to \mathbb{R}^n$ from an open subset A of $\mathbb{R}^m$ is harmonic if and only if each component is a harmonic function, as discussed in the first Section.

4. **Harmonic maps to a Euclidean space**. A smooth map $\phi : (\mathcal{M}, g) \to \mathbb{R}^n$ is harmonic if and only if each of its components is a *harmonic function* on $(\mathcal{M}, g)$, as in first chapter. See [46] for recent references.

5. **Harmonic maps to the circle** S^1 are given by integrating harmonic 1-forms with integral periods. Hence, when the domain $\mathcal{M}$ is compact, there are non-constant harmonic maps to the circle if and only if the first Betti number of $\mathcal{M}$ is non-zero. In fact, there is a harmonic map in every homotopy class (see, [7, Example 3.3.8]).

6. **Geodesics**. For a smooth curve, i.e. smooth map $\phi : A \to \mathcal{N}$ from an open subset A of $\mathbb{R}$ or from the circle S^1, the tension field is just the acceleration vector of the curve; hence ϕ *is harmonic if and only if it defines a geodesic* parametrized linearly (i.e., parametrized by a constant multiple of arc length). More generally, a map $\phi : \mathcal{M} \to \mathcal{N}$ is called *totally geodesic* if it maps linearly parametrized geodesics of $\mathcal{M}$ to linearly parametrized geodesics of $\mathcal{N}$, such maps are characterized by the vanishing of their second fundamental form. Since (11) exhibits the tension field as the trace of the second fundamental form, *totally geodesic maps are harmonic*.

7. **Isometric immersions** Let $\phi : (\mathcal{N}, h) \to (\mathcal{P}, k)$ be an isometric immersion. Then its second fundamental form $\beta(\phi)$ of ϕ has values in the normal space and coincides with the usual second fundamental form $A \in \Gamma(S^2T^*\mathcal{N} \otimes N\mathcal{N})$ of $\mathcal{N}$ as an (immersed) submanifold of $\mathcal{P}$ defined on vector fields X, Y on $\mathcal{M}$ by $A(X, Y) = -$ normal component of ${}^h\nabla_X Y$.[4] (Here, by $S^2T^*\mathcal{N}$ we denote the symmetrized tensor product of $T^*\mathcal{N}$ with itself and $N\mathcal{N}$ is the normal bundle of N in P.) In particular, the tension field $\tau(\phi)$ is m times the mean curvature of $\mathcal{M}$ in $\mathcal{N}$ so that ϕ *is harmonic if and only if* $\mathcal{M}$ *is a minimal submanifold of* $\mathcal{N}$.

8. **Compositions** The composition of two harmonic maps is not, in general, harmonic. In fact, the tension field of the composition of two smooth maps $\phi : (\mathcal{M}, g) \to (\mathcal{N}, h)$ and $f : (\mathcal{N}, h) \to (\mathcal{P}, k)$ is given by

$$\tau(f \circ \phi) = df(\tau(\phi)) + \beta(f)(d\phi, d\phi) = df(\tau(\phi)) + \sum_{i=1}^{m} \beta(f)(d\phi(e_i), d\phi(e_i)) \quad (17)$$

where $\{e_i\}$ is an orthonormal frame on $\mathcal{N}$. From this we see that *if ϕ is harmonic and f totally geodesic, then $f \circ \phi$ is harmonic*.

9. **Maps into submanifolds**. Suppose that $j : (\mathcal{N}, h) \to (\mathcal{P}, k)$ is an isometric immersion. Then, as above, its second fundamental form A has values in the normal space of $\mathcal{N}$ in P and so from the composition law just discussed, $\phi : (\mathcal{M}, g) \to (\mathcal{N}, h)$ is harmonic if and only if $\tau(j \circ \phi)$ is normal to M, and this holds if and only if

$$\tau(j \circ \phi) + \mathrm{Tr}\ A(d\phi, d\phi) = 0. \quad (18)$$

10. **Holomorphic maps**. By writing the tension field in complex coordinates, it is easy to see that *holomorphic (or antiholomorphic) maps $\phi : (\mathcal{M}, g, J^{\mathcal{M}}) \to (\mathcal{N}, h, J^{\mathcal{N}})$ between Kähler manifolds are harmonic* [66].[5]

11. **Maps between surfaces**. Let $\mathcal{M} = (\mathcal{M}^2, g)$ be a *surface*, i.e., two-dimensional Riemannian manifold. Assume it is oriented and let $J^{\mathcal{M}}$ be rotation by $+\pi/2$ on each tangent space. Then $(\mathcal{M}^2, g, J^{\mathcal{M}})$ defines a complex structure on $\mathcal{M}$ so that it becomes a *Riemann surface*; this structure is automatically Kähler. Let $\mathcal{N}$ be another oriented surface. Then from the last paragraph, we see that *any holomorphic or antiholomorphic map from $\mathcal{M}$ to $\mathcal{N}$ is harmonic*.

A smooth map $\phi : (\mathcal{M}, g) \to (\mathcal{N}, h)$ between Riemannian manifolds is called *weakly conformal* if its differential preserves angles at *regular points*—points where the differential is non-zero. Points where the differential is zero are called *branch points*. In local

[4] The minus sign is often omitted

[5] A. Lichnerowicz relaxes the conditions on $\mathcal{M}$ and $\mathcal{N}$ for which this is true; see, for example, [61] or [7, Chapter 8].

coordinates, a smooth map ϕ is weakly conformal if and only if there exists a function $\lambda : \mathcal{M} \to [0, \infty)$ such that

$$h_{\alpha\beta}\phi_i^\alpha \phi_j^\beta = \lambda^2 g_{ij} \tag{19}$$

Weakly conformal maps between surfaces are locally the same as holomorphic maps and so *weakly conformal maps of surfaces are harmonic*.

12. **Maps from surfaces**. (i) Let $\mathcal{M} = (\mathcal{M}^2, g)$ be a surface and let $\phi : \mathcal{M} \to \mathcal{N}$ be a smooth map to an arbitrary Riemannian manifold. Then the energy integral (7) is clearly invariant under conformal changes of the metric, and thus so is harmonicity of ϕ. To see this last invariance another way, let (x, y) be *conformal local coordinates*, i.e., coordinates on an open set of $\mathcal{M}$ in which $g = \mu^2(dx^2 + dy^2)$ for some real-valued function μ. Write $z = x + iy$. Then the harmonic equation reads

$$^\phi\nabla_{\partial/\partial\bar{z}}\frac{\partial\phi}{\partial z} \equiv {}^\phi\nabla_{\partial/\partial z}\frac{\partial\phi}{\partial\bar{z}} = 0. \tag{20}$$

If $\mathcal{M}$ is oriented, then we may take (x, y) to be oriented; the the coordinates $z = x + iy$ give $\mathcal{M}$ the complex structure of the last paragraph. Hence, harmonicity of a map from a Riemann surface is well defined.

Alternatively, from (17) we obtain the slightly more general statement that *the composition of a weakly conformal map $\phi : \mathcal{M} \to \mathcal{N}$ of surfaces with a harmonic map $f : \mathcal{N} \to \mathcal{P}$ from a surface to an arbitrary Riemannian manifold is harmonic*.

For any smooth map $\phi : \mathcal{M}^2 \to (\mathcal{N}, h)$ from an oriented surface, define the Hopf differential by

$$\begin{aligned}\mathcal{H} = (\phi^* h)^{(2,0)} &= h\Big(\frac{\partial\phi}{\partial z}, \frac{\partial\phi}{\partial z}\Big)dz^2 \\ &= \frac{1}{4}\left\{h\Big(\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial x}\Big) - h\Big(\frac{\partial\phi}{\partial y}, \frac{\partial\phi}{\partial y}\Big) + 2i\,h\Big(\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}\Big)\right\} dz^2\end{aligned} \tag{21}$$

Here we use the complex eigenspace decomposition $\phi^* h = (\phi^* h)^{(2,0)} + (\phi^* h)^{(1,1)} + (\phi^* h)^{(0,2)}$ under the action of $J^{\mathcal{M}}$ on quadratic forms on $T\mathcal{M}$. Note that (i) if *ϕ is harmonic, then $\mathcal{H}$ is a holomorphic quadratic differential*, i.e., a holomorphic section of $\otimes^2 T^*_{1,0}M$;[6] (ii) *ϕ is conformal if and only if $\mathcal{H}$ vanishes*. It follows that *any harmonic map from the 2-sphere is weakly conformal* [144, 88, 117]. Indeed, when $\mathcal{M}$ is the 2-sphere, $\otimes^2 T^*_{1,0}\mathcal{M}$ has negative degree so that any holomorphic section of it is zero.

13. **Minimal branched immersions**. For a weakly conformal map from a surface $(\mathcal{M}^2, g)$, comparing definitions shows that the tension field is a multiple of its mean curvature vector, so that *a weakly conformal map $\phi : (\mathcal{M}^2, g) \to (\mathcal{N}, h)$ is harmonic if and only if its image is minimal at regular points*; such maps are called *minimal branched immersions.* In suitable coordinates, the branch points have the form $z \mapsto \big(z^k + O(z^{k+1}), O(z^{k+1})\big)$ for some $k \in \{2, 3, \ldots\}$ [95].

Note also that the energy of a weakly conformal map $\phi : (\mathcal{M}^2, g) \to (\mathcal{N}, h)$ from a compact surface is equal to its *area*:

$$\mathcal{A}(\phi) = \int_{\mathcal{M}} |d\phi(e_1) \wedge d\phi(e_2)|\,\omega_g \qquad (\{e_1, e_2\} \text{ orthonormal frame}). \tag{22}$$

[6] This is an example of a *conservation law*, see §3.1 for more details and the generalization to higher dimensions.

14. **Harmonic morphisms** are a special sort of harmonic map; we turn to those now.

2.3 Harmonic morphisms

A continuous map $\phi : (\mathcal{M}, g) \to (\mathcal{N}, h)$ is called a *harmonic morphism* if, for every harmonic function $f : V \to \mathbb{R}$ defined on an open subset V of $\mathcal{N}$ with $\phi^{-1}(V)$ non-empty, the composition $f \circ \phi$ is harmonic on $\phi^{-1}(V)$. It follows that ϕ is smooth, since harmonic functions have that property, by a classical result of Schwartz [195, Chapter VI, Théorème XXIX]. Further, since any harmonic function on a real-analytic manifold is real analytic [168], harmonic morphisms between real-analytic Riemannian manifolds are, in fact, real analytic.

The subject of harmonic morphisms began with a paper of C. G. J. Jacobi [125], published in 1848. Jacobi investigated when complex-valued solutions to Laplace's equation on domains of Euclidean 3-space remain solutions under post-composition with holomorphic functions in the plane. It follows quickly that such solutions pull back locally defined harmonic functions to harmonic functions, i.e., are harmonic morphisms. A hundred years later came the axiomatic formulation of *Brelot harmonic space*. This is a topological space endowed with a sheaf of 'harmonic' functions characterized by a number of axioms. The morphisms of such spaces, i.e. mappings which pull back germs of harmonic functions to germs of harmonic functions, were confusingly called *harmonic maps* [48]; the term *harmonic morphisms* was coined by B. Fuglede [77].

To keep the number of references manageable, in the sequel we shall often refer to the book [7] which gives a systematic account of the subject, and which may be consulted for a list of original references.

A smooth map $\phi : (\mathcal{M}, g) \to (\mathcal{N}, h)$ is called *horizontally (weakly) conformal* (or *semiconformal*) if, for each $p \in \mathcal{M}$, *either*, (i) $d\phi_p = 0$, in which case we call p a *critical point*, or, (ii) $d\phi_p$ maps the *horizontal space* $\mathcal{H}_p = \{\ker(d\phi_p)\}^\perp$ conformally onto $T_{\phi(p)}\mathcal{N}$, i.e., $d\phi_p$ is surjective and there exists a number $\lambda(p) \neq 0$ such that

$$h\big(d\phi_p(X), d\phi_p(Y)\big) = \lambda(p)^2\, g(X, Y) \qquad (X, Y \in \mathcal{H}_p)\,,$$

in which case we call p a *regular point*. On setting $\lambda = 0$ at critical points, we obtain a continuous function $\lambda : \mathcal{M} \to [0, \infty)$ called the *dilation* of ϕ; note that λ^2 is smooth since it equals $|d\phi|^2/n$. In local coordinates, the condition for horizontal weak conformality is

$$g^{ij}\phi_i^\alpha\phi_j^\beta = \lambda^2 h^{\alpha\beta}\,. \tag{23}$$

Note that this condition is dual to condition (19) weak conformality, see also [7]. We have the following characterization [77, 124]: *a smooth map $\phi : \mathcal{M} \to \mathcal{N}$ between Riemannian manifolds is a harmonic morphism if and only if it is both harmonic and horizontally weakly conformal.* This is proved by (i) showing that there is a harmonic function $f : \mathcal{N} \supset V \to \mathbb{R}$ with any prescribed (traceless) 2-jet; see [7, §4.2]; (ii) applying the formula (17) for the tension field of the composition of ϕ with such harmonic functions f. It follows that *a non-constant harmonic morphism is **(i)** an open mapping, **(ii)** a submersion on a dense open set* — in fact the complement of this, the set of critical points, is a *polar set*.

Regarding the behaviour of a harmonic morphism at a critical point, the *symbol*, i.e. the first non-zero term of the Taylor expansion is a harmonic morphism between Euclidean spaces given by homogeneous polynomials; by studying these it follows that (i) *if* $\dim \mathcal{M} < 2\dim \mathcal{N} - 2$, *the harmonic morphism has no critical points, i.e., is submersive*; (ii) *if* $\dim \mathcal{M} = 2\dim \mathcal{N} - 2$, *the symbol is the cone on a Hopf map* [7, Theorem 5.7.3]. When $\dim \mathcal{M} = 3$ and $\dim \mathcal{N} = 2$, locally [7, Proposition 6.1.5], and often globally [7, Lemma 6.6.3], a harmonic morphism looks like a submersion followed by a holomorphic map of surfaces; the critical set is the union of geodesics. When $\dim \mathcal{M} = 4$ and $\dim \mathcal{N} = 3$, critical points are isolated and the harmonic morphism looks like the cone on the Hopf map $S^3 \to S^2$ [7, §12.1]. In both these cases, there are *global factorization theorems*. In other cases, little is known about the critical points.

The system (16, 23) for a harmonic morphism is, in general, *overdetermined*, so there are no general existence results. However, in many cases, we can establish existence or non-existence as we now detail.

1. **When** $\dim \mathcal{N} = 1$, the equation (23) is automatic, so that *a harmonic morphism is exactly a harmonic map*. If $\mathcal{N} = \mathbb{R}$, it is thus a harmonic function; for $\mathcal{N} = S^1$, see Example 5 of §2.2.

2. **When** $\dim \mathcal{M} = \dim \mathcal{N} = 2$, the equation (16) is implied by the equation (23), so that *the harmonic morphisms are precisely the weakly conformal maps*; see Example 9 of §2.2 for a discussion of such maps.

3. **When** $\dim \mathcal{N} = 2$ **and** $\dim \mathcal{M}$ **is arbitrary**, we have a number of special properties which are dual to those for (weakly conformal) harmonic maps *from* surfaces: (i) *conformal invariance in the codomain*: if we replace the metric on the codomain by a conformally equivalent metric, or post-compose the map with a (weakly) conformal map of surfaces, then it remains a harmonic morphism; (ii) a *variational characterization*: harmonic morphisms are the critical points of the energy when both the map and the metric on the horizontal space are varied, see [7, Corollary 4.3.14]; (iii) a non-constant map is a harmonic morphism if and only if it is horizontally weakly conformal and, at regular points, its fibres are minimal [6], i.e., at regular points, the fibres form a *conformal foliation by minimal submanifolds*.

4. **When** $\dim \mathcal{N} = 2$ **and** $\dim \mathcal{M} = 3$, if $\mathcal{M}$ has constant curvature, there are many harmonic morphisms locally given by a sort of *Weierstrass formula* [7, Chapter 6]. Globally, there are few, for example, when $\mathcal{M} = \mathbb{R}^3$, only orthogonal projection from $\mathbb{R}^3$ to $\mathbb{R}^2$ followed by a weakly conformal map. If $\mathcal{M}$ does not have constant curvature, the presence of a harmonic morphism implies some symmetry of the Ricci tensor and, locally, there can be at most two non-constant harmonic morphisms (up to post-composition with weakly conformal maps), and, none for most metrics including that of the Lie group Sol. As for *global* topological obstructions, a harmonic morphism from a compact 3-manifold gives it the structure of a *Seifert fibre space* [7, §10.3].

5. **When** $\dim \mathcal{N} = 2$ **and** $\dim \mathcal{M} = 4$, if $\mathcal{M}$ is Einstein, there is a *twistor correspondence* between harmonic morphisms to surfaces and *Hermitian structures on* $\mathcal{M}$. There are curvature obstructions for the local existence of such Hermitian structures. See [7, Chapter 7].

6. **When** $\dim \mathcal{N} = 2$ **and** $\mathcal{M}$ **is a symmetric space**, by finding suitably orthogonal families of complex-valued harmonic functions and composing these with holomorphic maps, Gudmundsson and collaborators construct harmonic morphisms from many compact and non-compact classical symmetric spaces [93], see also [7, §8.2].

7. **Riemannian submersions** *are harmonic, and so are harmonic morphisms, if and only if their fibres are minimal.* The *Hopf maps* from $S^3 \to S^2$, $S^7 \to S^4$, $S^{15} \to S^8$, $S^{2n+1} \to \mathbb{C}P^n$, $S^{4n+3} \to \mathcal{H}P^n$ are examples of such harmonic morphisms. See also [7, §4.5].

8. The natural projection of a **warped product** $\mathcal{M} = F \times_{f^2} \mathcal{N} \to \mathcal{N}$ onto its second factor is a horizontally conformal map with grad λ vertical, totally geodesic fibres and integrable horizontal distribution; in particular is a harmonic morphism. The *radial projections* $\mathbb{R}^m \setminus \{\mathbf{0}\} \to S^{m-1}$ $(m = 2, 3, \ldots)$, given by $\mathbf{x} \mapsto \mathbf{x}/|\mathbf{x}|$, are such maps. See also [7, §12.4].

9. When $\dim \mathcal{M} - \dim \mathcal{N} = 1$, i.e., the map $\phi : \mathcal{M} \to \mathcal{N}$ has **one-dimensional fibres**, R. Bryant [29] gives the following *normal form* for the metric g on the domain of a submersive harmonic morphism ϕ in terms of the pull-back $\phi^* h$ of the metric on the codomain and the dilation λ of the map, namely,

$$g = \lambda^{-2} \phi^* h + \lambda^{2n-4} \theta^2$$

where θ is a connection 1-form; thus *locally* such a harmonic morphism is a *principal S^1-bundle with S^1-connection*; this holds *globally* if the fibres are all compact, see [7, §10.5].

10. It follows that given a **Killing field** V (or isometric action) on $(\mathcal{M}, g)$, there are locally harmonic morphisms with fibres tangent to V.

By analysing the overdetermined system (16, 23) using *exterior differential systems*, Bryant [29] shows that any harmonic morphism with one-dimensional fibres from a space form is of warped product type or comes from a Killing field (this has been generalized to Einstein manifolds by R. Pantilie and Pantilie & Wood, see [7, Chapter 12]). It follows that *the only harmonic morphisms from Euclidean spheres with one-dimensional fibres are the Hopf maps $S^{2n+1} \to \mathbb{C}P^n$.*

11. There are **topological restrictions** on the existence of harmonic morphisms, for example, since harmonic morphisms preserve the harmonicity of 1-forms, Eells and Lemare showed that *the Betti number of the domain cannot be less than that of the codomain*, see [7, Proposition 4.3.11]. Pantilie and Wood show that the *Euler characteristic* must vanish for a harmonic morphism with fibres of dimension one from a compact domain of dimension not equal to 4. In particular *there is no non-constant harmonic morphism from a sphere S^{2n} $(n \neq 2)$ to a Riemannian manifold of dimension $2n - 1$, whatever the metrics.* Further the Pontryagin numbers and the signature vanish, see [7, §12.1].

12. **When** $\dim \mathcal{M} = 4$, *the Euler characteristic is even and equals the number of critical points of the harmonic morphism*, so that we cannot rule out the existence of a harmonic morphism from S^4. By Bryant's result in item 8 above, there is no harmonic morphism from the *Euclidean* 4-sphere with one-dimensional fibres; however, there is one if the metric on S^4 is changed by a suitable conformal factor. This map is given by suspending the Hopf map, first finding a suspension which is horizontally conformal, then changing the metric conformally on the domain to 'render' it harmonic. At both stages, the problem is reduced to solving an ordinary differential equation for the suspension function with suitable boundary values, and the method applies to find many more harmonic morphisms, see [7, Chapter 13].

13. Finally note that J.-Y. Chen shows that **stable harmonic maps** *from compact Riemannian manifolds to S^2 are all harmonic morphisms.* This is shown by calculating the

second variation and showing that its non-negativity forces the map to be horizontally weakly conformal, see [7, §8.7].

3 Weakly harmonic maps and Sobolev spaces between manifolds

3.1 Weakly harmonic maps

An extension of the Dirichlet principle or, more generally, the use of variational methods requires the introduction of a class of distributional maps endowed with a topology which is sufficiently coarse to ensure the *compactness* of sequences of maps which we hope will converge to a solution. On the other hand, the energy functional should be defined on this class and we should be able to make sense of its Euler–Lagrange equation (16). These two requirements are somewhat in conflict, and will lead us to model the class of maps on the Sobolev space $W^{1,2}(\mathcal{M})$. But that will force us to work with weak solutions of (16), i.e., *weakly harmonic maps*. However, as soon as $m := \dim \mathcal{M} \geq 2$, a map $f \in W^{1,2}(\mathcal{M})$ is not continuous in general. Hence, even if $W^{1,2}(\mathcal{M}, \mathcal{N})$ makes sense, there is no reason for a map $\phi \in W^{1,2}(\mathcal{M}, \mathcal{N})$ to take values in any open subset, in general. This makes it difficult to study ϕ by using local charts on the *target manifold* $\mathcal{N}$. Today[7] most authors avoid these difficulties by using the Nash–Moser embedding theorem (see, for example, [91]) as follows. In the following **we shall assume that** $\mathcal{N}$ **is compact**. Then there exist an *isometric embbeding* $j : (\mathcal{N}, g) \longrightarrow (\mathbb{R}^N, \langle \cdot, \cdot \rangle)$. And we define (temporarily), for any open subset $\Omega \subset \mathcal{M}$,

$$W_j^{1,2}(\Omega, \mathcal{N}) := \{u \in W^{1,2}(\Omega, \mathbb{R}^N) |\ u(x) \in j(\mathcal{N}) \text{ a.e.}\}. \tag{24}$$

On this set the *energy* or *Dirichlet functional* defined by (7) now reads

$$E_\Omega(u) := \frac{1}{2} \int_\Omega g^{ij}(x) \left\langle \frac{\partial u}{\partial x^j}, \frac{\partial u}{\partial x^i} \right\rangle \omega_g.$$

But, if we assume that $\mathcal{M}$ is also compact, then for any two isometric embeddings j_1, j_2, the spaces $W_{j_1}^{1,2}(\mathcal{M}, \mathcal{N})$ and $W_{j_2}^{1,2}(\mathcal{M}, \mathcal{N})$ are homeomorphic and $E_\Omega(j_2 \circ j_1^{-1} \circ u) = E_\Omega(u)$. Hence we simply[8] write $W^{1,2}(\mathcal{M}, \mathcal{N}) := W_j^{1,2}(\mathcal{M}, \mathcal{N})$.

Weakly harmonic maps

In order to define weakly harmonic maps as extremals of $E_\mathcal{M}$ we have to specify which infinitesimal deformations of a map $u \in W^{1,2}(\mathcal{M}, \mathcal{N})$ we will consider. Consider a neighbourhood $\mathcal{V}$ of $\mathcal{N}$ in $\mathbb{R}^N$ such that the projection map $P : \mathcal{V} \longrightarrow \mathcal{N}$ which sends each $y \in \mathcal{V}$ to the nearest point in $\mathcal{N}$ is well defined and smooth[9]. Now let $u \in W^{1,2}(\mathcal{M}, \mathcal{N})$. For any map $v \in W^{1,2}(\mathcal{M}, \mathbb{R}^N) \cap L^\infty(\mathcal{M}, \mathbb{R}^N)$ we observe that,

[7] In his 1948 paper [156], C. B. Morrey had to work *without* the Nash–Moser theorem which was not yet proved.

[8] In the case where $\mathcal{M}$ is not compact, we may not have $W_{j_1}^{1,2}(\mathcal{M}, \mathcal{N}) \simeq W_{j_2}^{1,2}(\mathcal{M}, \mathcal{N})$ because the L^2 norms of $j_1 \circ \phi$ and $j_2 \circ \phi$ may be different (indeed, one of the two norms may be bounded whereas the other one may be infinite). This suggests that perhaps a more satisfactory (but less used) definition of $W_j^{1,2}(\mathcal{M}, \mathcal{N})$ would be: *the set of measurable distributions on* $\mathcal{M}$ *with values in* $\mathbb{R}^N$ *such that* $du \in L^2(\mathcal{M})$ *and* $u(x) \in j(\mathcal{N})$ *a.e.*

[9] We may use other projection maps, not necessarily Euclidean projections, see [118].

for ε sufficiently small, $u + \varepsilon v \in \mathcal{V}$, so that $u_\varepsilon^v := P(u + \varepsilon v) \in W^{1,2}(\mathcal{M}, \mathcal{N})$. We set $\dot{u}_0^v := \lim_{\varepsilon \to 0} (u_\varepsilon^v - u)/\varepsilon = dP_u(v)$ a.e. and

$$(\delta E_{\mathcal{M}})_u(\dot{u}_0^v) := \lim_{\varepsilon \to 0} \frac{E_{\mathcal{M}}(u_\varepsilon) - E_{\mathcal{M}}(u)}{\varepsilon}.$$

(What is important in this definition is that $\varepsilon \longmapsto u_\varepsilon^v$ is a differentiable curve into $W^{1,2}(\mathcal{M}, \mathbb{R}^N)$ such that $\forall \varepsilon$, $u_\varepsilon^v \in W^{1,2}(\mathcal{M}, \mathcal{N})$, $du_\varepsilon^v/d\varepsilon \in W^{1,2}(\mathcal{M}, \mathbb{R}^N)$ and $u_0^v = u$.) And u is **weakly harmonic** if and only if $(\delta E_{\mathcal{M}})_u(\dot{u}_0^v) = 0$ for all $v \in W^{1,2} \cap L^\infty(\mathcal{M}, \mathbb{R}^N)$. Equivalently u is a *solution in the distribution sense* of a *system of N coupled scalar elliptic PDEs*, i.e. an $\mathbb{R}^N$*-valued elliptic PDE*

$$\Delta_g u + g^{ij}(x)\, A_{u(x)}\left(\frac{\partial u}{\partial x^i}, \frac{\partial u}{\partial x^j}\right) = 0 \tag{25}$$

where $A \in \Gamma(S^2 T^* \mathcal{N} \otimes N\mathcal{N})$ is the second fundamental form of the embedding j as in §2.2 ($N\mathcal{N}$ now denotes the normal bundle of $\mathcal{N}$ in $\mathbb{R}^N$);[10] one can check that this condition is independent of the embedding j [118]. Indeed, it is just the equation (18).

Example. ($\mathcal{N} = S^n$, *the unit sphere*) The n-dimensional sphere S^n is the submanifold $\{y \in \mathbb{R}^{n+1} \mid |y| = 1\}$, its metric is the pull-back of the standard Euclidean metric by the embedding $j : S^n \longrightarrow \mathbb{R}^{n+1}$. The second fundamental form of j is given by $A_y(X, Y) = \langle X, Y \rangle y$, so that the weakly harmonic maps are the maps in $W^{1,2}(\mathcal{M}, S^n)$ such that

$$\Delta_g u + |du|^2 u = 0 \quad \text{in a distribution sense.} \tag{26}$$

Remarks (i) In (25), $\Delta_g u \in W^{-1,2}(\mathcal{M}, \mathbb{R}^N)$ is defined in the distribution sense, the coefficients of $A_{u(x)}$ are in L^∞ because $\mathcal{N}$ is compact and so $g^{ij}(x) A_{u(x)}(\partial u/\partial x^i, \partial u/\partial x^j) \in L^1(\mathcal{M}, \mathbb{R}^N)$.

(ii) The system (25) is an example of a *semilinear elliptic system with a nonlinearity which is quadratic in the first derivatives*, for which a general regularity theory has been developed (see [143, 229, 121, 83]). This nonlinearity is the reason why most of analytical properties valid for harmonic functions are lost: *existence*, *regularity*, *uniqueness* and *minimality* may fail in general, unless some extra hypotheses are added.

(iii) A difficulty particular to this theory is that $W^{1,2}(\mathcal{M}, \mathcal{N})$ is not a $\mathcal{C}^1$-manifold. One can only say that $W^{1,2}(\mathcal{M}, \mathcal{N})$ is a Banach manifold, which is not separable if $m \geq 2$, and that $\mathcal{C}^0 \cap W^{1,2}(\mathcal{M}, \mathcal{N})$ is a closed separable submanifold of $W^{1,2}(\mathcal{M}, \mathcal{N})$ (see [31]). Moreover, $W^{1,2}(\mathcal{M}, \mathcal{N})$ does not have the same topology as $\mathcal{C}^0(\mathcal{M}, \mathcal{N})$ in general (see §3.2 and 3.3).

Minimizing maps

A map $u \in W^{1,2}(\mathcal{M}, \mathcal{N})$ is called an **energy minimizing map** if any map $v \in W^{1,2}(\mathcal{M}, \mathcal{N})$ which coincides with u outside a compact subset $K \subset \mathcal{M}$ has an energy greater than or equal to that of u, i.e. $E_{\mathcal{M}}(v) \geq E_{\mathcal{M}}(u)$. A weaker notion is that

[10] An equivalent formula for A is, as follows: if for any $y \in \mathcal{N}$, we denote by $P_y^\perp : \mathbb{R}^N \longrightarrow N_y\mathcal{N}$ the orthonormal projection, then A can be defined by $A_y(X, Y) := (D_X P_y^\perp)(Y)$, $\forall X, Y \in \Gamma(T\mathcal{N})$, where D is the (flat) Levi-Civita connection on $\mathbb{R}^N$.

$u \in W^{1,2}(\mathcal{M}, \mathcal{N})$ is called **locally energy minimizing** if, for any point $x \in \mathcal{M}$, there exists a neighbourhood $U \subset \mathcal{M}$ of x such that any map $v \in W^{1,2}(\mathcal{M}, \mathcal{N})$ which coincides with u outside a compact subset $K \subset U$ has an energy greater than or equal to that of u.

Stationary maps

The family $\{u_\varepsilon^v | \; v \in W^{1,2}(\mathcal{M}, \mathbb{R}^N) \cap L^\infty(\mathcal{M}, \mathbb{R}^N)\}$ of infinitesimal deformations used for the definition of a weakly harmonic map u does not contain some significant deformations. For example, consider *radial projection* $u_\odot \in W^{1,2}(B^3, S^2)$ defined by[11]

$$u_\odot(x) = x/|x|\,; \tag{27}$$

it seems natural to *move* the singularity of this map along some smooth path. For example we let $a \in \mathcal{C}^1((-1,1), B^3)$ parametrize such a path in B^3 such that $a(0) = 0$ and we consider the family of maps $u_\varepsilon \in W^{1,2}(B^3, S^2)$ defined by $u_\varepsilon(x) = (x - a(\varepsilon))/|x - a(\varepsilon)|$. Then $du_\varepsilon/d\varepsilon$ is **not** in $W^{1,2}(B^3, \mathbb{R}^3)$, and hence we cannot take this infinitesimal variation of $u_\odot$ into account for weakly harmonic maps. This is the reason for considering a second type of variation: we let $(\varphi_t)_{t\in I}$ (where $I \subset \mathbb{R}$ is some open interval which contains 0) be a $\mathcal{C}^1$ family of smooth diffeomorphisms $\varphi_t : \mathcal{M} \longrightarrow \mathcal{M}$ such that φ_0 is the identity. Then for any $u \in W^{1,2}(\mathcal{M}, \mathcal{N})$, $(u \circ \varphi_t)_{t\in I}$ is a $\mathcal{C}^1$ family of maps in $W^{1,2}(\mathcal{M}, \mathcal{N})$ such that $u \circ \varphi_0 = u$. Following [189] we say that u **is stationary if** (i) u **is weakly harmonic, and** (ii) **for any family of diffeomorphisms** $(\varphi_t)_{t\in I}$ **with** $\varphi_0 = \mathbf{Id}_\mathcal{M}$,

$$\lim_{t\to 0}\big(E_\mathcal{M}(u \circ \varphi_t) - E_\mathcal{M}(u)\big)/t = 0. \tag{28}$$

Note that, without loss of generality, we can assume that the diffeomorphisms φ_t have the form $\varphi_t = e^{tX}$, where X is a smooth tangent vector field with compact support on $\mathcal{M}$. Maps u which satisfies (28) can be characterized by the following local condition derived by P. Baird and J. Eells, and by A. I. Pluzhnikov independently [6, 170, 118]. Let us stress temporarily the dependence of the Dirichlet energy on the metric g on $\mathcal{M}$ by writing $E_\mathcal{M} = E_{(\mathcal{M},g)}$. Then we remark that, by the change of variable $\widetilde{x} = e^{tX}(x)$ in the Dirichlet integral, we have:

$$E_{(\mathcal{M},g)}(u \circ e^{tX}) = E_{(\mathcal{M},(e^{-tX})^*g)}(u \circ e^{tX} \circ e^{-tX}) = E_{(\mathcal{M},(e^{-tX})^*g)}(u),$$

where $(e^{-tX})^*g$ is the pull-back of the metric g by e^{-tX}. But we compute:

$$E_{(\mathcal{M},(e^{-tX})^*g)}(u) = E_{(\mathcal{M},g)}(u) + t\int_\mathcal{M} \big(L_X g^{ij}\big)\, S_{ij}(u)\, \omega_g + o(t),$$

where

$$S_{ij}(u) := \frac{1}{2}|du|_g^2 g_{ij} - (u^*h)_{ij} = \frac{1}{2} g^{kl}(x) \left\langle \frac{\partial u}{\partial x^k}, \frac{\partial u}{\partial x^l} \right\rangle g_{ij} - \left\langle \frac{\partial u}{\partial x^i}, \frac{\partial u}{\partial x^j} \right\rangle,$$

where $|du|_g^2 = g^{ij}(x)\,\langle \partial u/\partial x^i\,,\, \partial u/\partial x^j\rangle$, is called the **stress-energy tensor**. If $u \in W^{1,2}(\mathcal{M}, \mathcal{N})$, then its components are in $L^1(\mathcal{M})$. Hence condition (28) is equivalent

[11] For $a \in \mathbb{R}^m$ and $r > 0$ we write $B^m(a,r) := \{x \in \mathbb{R}^m |\; |x - a| < r\}$; also, for brevity write $B^m := B^m(0,1)$.

to the fact that $\int_{\mathcal{M}}\left(L_X g^{ij}\right) S_{ij}(u)\,\omega_g = 0$, for all smooth tangent vector fields X with compact support on $\mathcal{M}$. Moreover, since the stress-energy tensor is symmetric we have the identity $\left(2g^{ik}\nabla_k X^j + L_X g^{ij}\right) S_{ij}(u) = 0$, from which we can deduce by an integration by parts that u **satisfies** (28) **if and only if** $S_{ij}(u)$ **is covariantly divergence-free**, i.e.

$$\begin{aligned}&\forall j,\ \nabla_i S^i_j(u) = 0 \ \text{ in the distribution sense}\\ &\big(\text{where } S^i_j(u) := g^{ik}S_{kj}(u) \text{ and } \nabla_i := \nabla_{\partial/\partial x^i}\big).\end{aligned} \tag{29}$$

Remarks (i) If the metric g on $\mathcal{M}$ is Euclidean, i.e. if we can write $g_{ij} = \delta_{ij}$ in some coordinate system, then the covariant conservation law (29) becomes a system of m conservations laws.

(ii) if $m = 2$ then S_{ij} is trace free. Furthermore we can use conformal local coordinates $z = x^1 + ix^2$ on $\mathcal{M}$. Then if we identify S_{ij} with the quadratic form $S := S_{ij}dx^i dx^j$, we easily compute:

$$-2S = \text{Re}\left\{\left(\left|\frac{\partial u}{\partial x^1}\right|^2 - \left|\frac{\partial u}{\partial x^2}\right|^2 - 2i\left\langle\frac{\partial u}{\partial x^1},\frac{\partial u}{\partial x^2}\right\rangle\right)(dz)^2\right\} = 4\mathcal{H},$$

where $\mathcal{H}$ is the Hopf differential of u as defined in (21). We note that: (i) u **is conformal if and only if** $\mathcal{H}$ **or equivalently** S **vanishes** and (ii) **the stress-energy tensor is divergence free, i.e.** (29) **holds, if and only if** $\mathcal{H}$ **is holomorphic**.

Relationship between the different notions of critical points

It is easy to prove the inclusions:

$$\begin{aligned}&\{\text{minimizing maps}\} \subset \{\text{locally minimizing maps}\}\\ &\qquad\qquad \subset \{\text{stationary maps}\} \subset \{\text{weakly harmonic maps}\}\,;\end{aligned}$$

these inclusions are strict in general. For example, the identity map Id : $S^3 \longrightarrow S^3$ is *locally minimizing* (see §6.2) but *not globally minimizing* (see §3.3). The map $u^{(2)} \in W^{1,2}(B^3, S^2)$, defined by $u^{(2)}(x) = P^{-1}\circ Z^2 \circ P(x/|x|)$, where $P : S^2 \longrightarrow \mathbb{C}P = \mathbb{C}\cup\{\infty\} = \mathbb{R}^2\cup\{\infty\}$ defined by

$$P(y^1,y^2,y^3) = (y^1 + iy^2)/(1+y^3) \tag{30}$$

is the stereographic projection and $Z^2(z) = z^2$ is *stationary* but *not locally minimizing* (see §4.3 and [24]). The map $v_\lambda :\in W^{1,2}(B^3,S^2)$, defined by $v_\lambda(x) = P^{-1}\circ\lambda\circ P(x/|x|)$, where λ is the multiplication by some $\lambda\in\mathbb{C}^*$ is *weakly harmonic* but *not stationary if* $|\lambda| \neq 1$ (see [118], §1.4). However, *smooth harmonic maps* are stationary: one can check by a direct computation that, if u is a map of class $\mathcal{C}^2$, (25) implies (29).

3.2 The density of smooth maps in $W^{1,p}(\mathcal{M},\mathcal{N})$

In this section and the following, it may clarify the discussion to consider the more general family of spaces

$$W^{1,p}(\mathcal{M},\mathcal{N}) := \{u \in W^{1,p}(\mathcal{M},\mathbb{R}^N)|\ u(x)\in\mathcal{N} \text{ a.e.}\},$$

$W^{1,p}(\mathcal{M}) := \{u \in L^p(\mathcal{M}) | \; du \in L^p(\mathcal{M})\}$ and $W^{1,p}(\mathcal{M}, \mathbb{R}^N) := W^{1,p}(\mathcal{M}) \otimes \mathbb{R}^N$, where $1 \leq p < \infty$. An interesting functional on $W^{1,p}(\mathcal{M}, \mathcal{N})$ is the *p-energy*

$$E_{\mathcal{M}}^{(p)}(u) := \frac{1}{p} \int_{\mathcal{M}} \left(g^{ij}(x) \left\langle \frac{\partial u}{\partial x^i}, \frac{\partial u}{\partial x^j} \right\rangle \right)^{p/2} \omega_g. \tag{31}$$

For any Riemannian manifold $\mathcal{M}$ of dimension m and for any compact manifold $\mathcal{N}$ of dimension n, let us define

- $H_s^{1,p}(\mathcal{M}, \mathcal{N})$:= the closure of $\mathcal{C}^1(\mathcal{M}, \mathcal{N}) \cap W^{1,p}(\mathcal{M}, \mathcal{N})$ in the *strong* $W^{1,p}$-topology;
- $H_w^{1,p}(\mathcal{M}, \mathcal{N})$:= the closure of $\mathcal{C}^1(\mathcal{M}, \mathcal{N}) \cap W^{1,p}(\mathcal{M}, \mathcal{N})$ in the *sequential weak*[12] $W^{1,p}$-topology: a map $u \in W^{1,p}(\mathcal{M}, \mathcal{N})$ belongs to $H_w^{1,p}(\mathcal{M}, \mathcal{N})$ if and only if there exists a sequence $(v_k)_{k \in \mathbb{N}}$ of maps in $\mathcal{C}^1(\mathcal{M}, \mathcal{N}) \cap W^{1,p}(\mathcal{M}, \mathcal{N})$ such that v_k converges weakly to u as $k \to \infty$.

Note that we have always the inclusions

$$H_s^{1,p}(\mathcal{M}, \mathcal{N}) \subset H_w^{1,p}(\mathcal{M}, \mathcal{N}) \subset W^{1,p}(\mathcal{M}, \mathcal{N}). \tag{32}$$

The easy case $m \leq p$. We first observe that, *if* $p > m$, *the Sobolev embedding theorem implies that* $W^{1,p}(\mathcal{M}, \mathcal{N}) \subset \mathcal{C}^0(\mathcal{M}, \mathcal{N})$, so that by a standard regularization in $W^{1,p}(\mathcal{M}, \mathbb{R}^N)$ followed by a projection onto $\mathcal{N}$, one can prove easily that $H_s^{1,p}(\mathcal{M}, \mathcal{N})$ is dense in $W^{1,p}(\mathcal{M}, \mathcal{N})$. This result has been extended in [191] **to the critical exponent** $p = m$, by using the Poincaré inequality:[13] $⨍_{B^m(x,r)} |\varphi - ⨍_{B^m(x,r)} \varphi|^m \leq C \int_{B^m(x,r)} |d\varphi|^m$. In conclusion, **if** $p \in [m, \infty)$**, all inclusions in** (32) are equalities.

The hard case $1 \leq p < m$. One of the more instructive example is radial projection $u_\odot : B^m \to S^{m-1}$ given by (27). This map has a point singularity at 0, but is in $W^{1,p}(B^m, S^{m-1})$ if $p < m$. We shall see later that $u_\odot$ cannot be approximated by smooth maps with values in S^{m-1} for $m - 1 < p < m$. Variants of $u_\odot$ are the maps $u_\odot^s : B^m \longrightarrow S^{m-s-1}$, for $s \in \mathbb{N}$ such that $0 \leq s \leq m-1$, defined by $u_\odot^s(x, y) = x/|x|$, for $(x, y) \in \mathbb{R}^{m-s} \times \mathbb{R}^s$: this map is singular along the s-dimensional subspace $x = 0$.

Approximation by smooth maps with singularities. The following result by F. Bethuel [12, 102] shows that the structure of the singularities of the maps $u_\odot^s$ is some-

---

[12] The space $H_w^{1,p}(\mathcal{M}, \mathcal{N})$ plays an important role when using variational methods. For example, if we minimize the p-energy *among smooth maps*, the minimizing sequence converges weakly. Hence the weak solution that we obtain is naturally in $H_w^{1,p}(\mathcal{M}, \mathcal{N})$.

[13] Note that similar arguments show that *maps such that* $\lim_{r \to 0} E_{x,r}(u) = 0$ *for all* $x \in \mathcal{M}$ *can be approximated by smooth maps* (see §4.3 for a definition of $E_{x,r}$). This result is a key ingredient in the regularity theory for harmonic maps by R. Schoen and K. Uhlenbeck [190], see again §4.3. All these results fit in the framework of a theory of maps into manifolds with *vanishing mean oscillation*, developed by H. Brezis and L. Nirenberg [27]: for any locally integrable function f on $\mathbb{R}^m$, for any $x \in \mathbb{R}^m$ and any $r > 0$, set $f_{x,r} := ⨍_{B^m(x,r)} f$, then let $||f||_{BMO} := \sup_{x \in \mathbb{R}^m} \sup_{r>0} (⨍_{B^m(x,r)} |f - f_{x,r}|^p)^{1/p}$, for some $p \in [1, \infty)$. Then the space of functions of *bounded mean oscillation* (BMO) on $\mathbb{R}^m$ is the set of locally integrable functions f on $\mathbb{R}^m$ such that $||f||_{BMO}$ is bounded, and this definition does *not* depend on p [128]. The subspace of functions of *vanishing mean oscillation* (VMO) on $\mathbb{R}^m$ is composed of maps such that $\lim_{r \to 0} (⨍_{B^m(x,r)} |f - f_{x,r}|^p)^{1/p} = 0$ for any $x \in \mathbb{R}^m$ (see [128, 118]).

how generic. Let

$\mathcal{R}^{p,k}(\mathcal{M},\mathcal{N})$: the set of maps $u \in W^{1,p}(\mathcal{M},\mathcal{N})$ such that $\exists\, \Sigma_u \subset \mathcal{M}$ with $u \in \mathcal{C}^k(\mathcal{M}\setminus\Sigma_u,\mathcal{N})$, $\Sigma_u = \bigcup_{i=1}^r \Sigma_i$, Σ_i is a subset of a manifold of dimension $m-[p]-1$, $\partial\Sigma_i$ is $\mathcal{C}^k$

(note that, if $m-1 \le p < m$, each Σ_i is a point). Then

$$\textit{if } 1 < p < m, \quad \textit{then } \mathcal{R}^{p,k}(\mathcal{M},\mathcal{N}) \textit{ is dense in } W^{1,p}(\mathcal{M},\mathcal{N}). \tag{33}$$

Moreover, F. B. Hang and F. H. Lin [102] proved that the singular set Σ_u can be chosen as the $(m-[p]-1)$-skeleton of a smooth rectilinear cell decomposition.

The case of maps into the sphere. The idea of the proof of (33) in the case where $\mathcal{N} = S^n$ and $n \le p < n+1$ is the following (see also [17, 88]). Let $u \in W^{1,p}(\mathcal{M},S^n)$. Then by convolution with mollifiers we first produce a sequence of smooth maps $(u_\rho)_\rho$ which converges strongly to u as $\rho \to 0$, but *has values in* $B^{n+1}(0,1)$. However, for any $\varepsilon > 0$ the measure of $V_\rho^\varepsilon := u_\rho^{-1}(B^{n+1}(0,1-\varepsilon))$ tends to 0 as $\rho \to 0$. The main task is to compose the restriction $(u_\rho)|_{V_\rho^\varepsilon}$ of u_ρ to V_ρ^ε with a projection map from $B^{n+1}(0,1-\varepsilon)$ to its boundary in order to obtain a map into $B^{n+1}\setminus B^{n+1}(0,1-\varepsilon)$. The naive projection $x \longmapsto (1-\varepsilon)x/|x|$ fails because $u_\rho/|u_\rho|$ has infinite $W^{1,p}$-norm in general. The trick, inspired by [107], consists of using a different projection map $\Pi_a : x \longmapsto (1-\varepsilon)(x-a)/|x-a|$, where $a \in B^{n+1}(0,\frac{1}{2})$: by averaging over $a \in B^{n+1}(0,\frac{1}{2})$ and using Fubini's theorem one finds that there exists some a such that the $W^{1,p}$-norm of $(\Pi_a \circ u_\rho)|_{V_\rho^\varepsilon}$ is bounded in terms of the $W^{1,p}$-norm of $(u_\rho)|_{V_\rho^\varepsilon}$. Moreover, Sard's theorem ensures that for a generic a, $u_\rho^{-1}(a)$, i.e. the singular set of $\Pi_a \circ u_\rho$, is a smooth submanifold of codimension $n+1 = [p]+1$.

The property (33) shows that questions of density rely on approximating maps in $\mathcal{R}^{p,k}(\mathcal{M},\mathcal{N})$ by smooth maps. Again it is instructive to look at the example of the map $u_\odot \in W^{1,p}(B^m,S^{m-1})$: a way to approximate $u_\odot$ is *to move the topological singularity through a path joining the origin 0 to the boundary* ∂B^m. Consider such a path (for example, $[-1,0]\times\{0\}^{m-1} \subset \mathbb{R}^m$), then by modifying $u_\odot$ inside a small tube around this path in such a way that the topological degree on each sphere $S_r^{m-1} := \partial B^m(0,r)$ cancels, we obtain a continuous map into the sphere. For instance, for $\varepsilon > 0$ sufficiently small, we construct a map u_ε by replacing, for any $r \in [0,1]$, the restriction $u_\odot|_{S_r^{m-1}}$ of $u_\odot$ to S_r^{m-1} by its left composition with the map $T_{\lambda(r,\varepsilon)}^{-1} \circ U \circ T_{\lambda(r,\varepsilon)} : S^{m-1} \longrightarrow S^{m-1}$, where $U(y) = (|y^1|, y^2, \cdots, y^m)$ and $T_\lambda(y) = (\cosh\lambda + y^1\sinh\lambda)^{-1}(\sinh\lambda + y^1\cosh\lambda, y^2, \cdots, y^m)$ ($\lambda \in \mathbb{R}$) and we choose $\lambda(r,\varepsilon)$ in such a way that u_ε coincides with $u_\odot$ outside the tubular neighbourhood of the path of radius ε. Then, inside the small tube, $|du_\varepsilon| \le (C/\varepsilon)|du_\odot|$ so that the extra cost in $W^{1,p}$-norm of this modification is of order $\varepsilon^{m-1}/\varepsilon^p$ (note that ε^{m-1} controls the volume of the tube). We see that

(i) if $1 \le p < m-1$, this p-energy cost can be as small as we want;
(ii) if $p = m-1$, the p-energy cost does not tend to zero as $\varepsilon \to 0$ but is bounded;
(iii) if $m-1 < p < m$, the cost tends to ∞ as $\varepsilon \to 0$.

These heuristic considerations are behind a series of results proved by F. Bethuel [12] and summarized in the following table:

Inclusions	$H^{1,p}_s(B^m,S^{m-1}) \subset H^{1,p}_w(B^m,S^{m-1})$	$\subset W^{1,p}(B^m,S^{m-1})$
$1 \le p < m-1$	$=$	$=$
$p = m-1$	$\subsetneq$	$=$
$m-1 < p < m$	$=$	$\subsetneq$

In particular, we see that $p = m-1$ is another critical exponent: $u_\odot$ can be approximated by smooth maps weakly in $W^{1,m-1}$ but not strongly.

The fact that $H^{1,p}_s(B^m,S^{m-1}) \neq W^{1,p}(B^m,S^{m-1})$ for $m-1 \le p < m$ can be checked by using a degree argument. Here is a proof for $m = 3$ and $2 \le p < 3$. Let $\omega_{S^2} := j^*(y^1dy^2 \wedge dy^3 + y^2dy^3 \wedge dy^1 + y^3dy^1 \wedge dy^2)$ be the volume form on S^2 (j is the embedding $S^2 \subset \mathbb{R}^3$). Let $\chi \in \mathcal{C}^\infty(B^3,\mathbb{R})$ be a function which depends only on $r = |x|$, such that $\chi(1) = 0$ (i.e. $\chi = 0$ on ∂B^3) and $\chi(0) = -1$. Assume that there exists a sequence $(u_k)_{k\in\mathbb{N}}$ of functions in $\mathcal{C}^2(B^3,S^2)$ such that $u_k \to u_\odot$ strongly in $W^{1,2}(B^3,S^2)$. Then since $u_k^*\omega_{S^2}$ is quadratic in the first derivatives of u, $\int_{B^3} d\chi \wedge u_k^*\omega_{S^2}$ converges to $\int_{B^3} d\chi \wedge u_\odot^*\omega_{S^2} = \int_0^1 4\pi(d\chi/dr)dr = 4\pi$. On the other hand, since u_k is smooth, $d\left(u_k^*\omega_{S^2}\right) = u_k^*(d\omega_{S^2}) = 0$ and so

$$0 = \int_{\partial B^3} \chi\, u_k^*\omega_{S^2} = \int_{B^3} d\left(\chi\, u_k^*\omega_{S^2}\right) = \int_{B^3} d\chi \wedge u_k^*\omega_{S^2}.$$

Hence we also deduce that $\int_{B^3} d\chi \wedge u_k^*\omega_{S^2} \to 0$, a contradiction. In §5.4 another proof is given for $p = 2$.

In fact, a nice characterization of $H^{1,2}_s(B^3,S^2)$ was given by Bethuel [10] in terms of the pull-back of the volume form ω_{S^2} on S^2: **a map $u \in W^{1,2}(B^3,S^2)$ can be approximated by smooth maps in the strong $W^{1,2}$-topology if and only if** $d(u^*\omega_{S^2}) = 0$ (see also §5.4 for more results about $u^*\omega_{S^2}$.) This may be generalized to some situations (see [11]) but not all: indeed it is not clear whether such a cohomological criterion can be found to recognize maps in $H^{1,3}_s(B^4,S^2)$ — for example, the singular map defined by $h^{\mathbb{C}}_\odot(x) = H^{\mathbb{C}}(x/|x|)$, where $H^{\mathbb{C}} : S^3 \longrightarrow S^2$ is the Hopf fibration, is in $W^{1,3}(B^4,S^2)$ but not in $H^{1,3}_s(B^4,S^2)$ — see [110] for more details on this delicate situation.

The role of the topology of $\mathcal{M}$ and $\mathcal{N}$. We have seen that, when $\mathcal{N} = S^n$, the topology of $\mathcal{N}$ may cause obstructions to the density of smooth maps in $W^{1,p}(\mathcal{M},\mathcal{N})$. The first general statement in this direction is due to F. Bethuel and X. Zheng [17] and Bethuel [12] in terms of the $[p]$-th homotopy group of $\mathcal{N}$; namely, for $\mathcal{M} = B^m$ we have

$$\textit{if } 1 < p < m, \textit{ then} \quad H^{1,p}_s(B^m,\mathcal{N}) = W^{1,p}(B^m,\mathcal{N}) \quad \Longleftrightarrow \quad \pi_{[p]}(\mathcal{N}) = 0.$$

However, for an *arbitrary* manifold $\mathcal{M}$, the condition that $\pi_{[p]}(\mathcal{N}) = 0$ is *not sufficient* to ensure that $H^{1,p}_s(\mathcal{M},\mathcal{N}) = W^{1,p}(\mathcal{M},\mathcal{N})$, in general. This was pointed out in [102]. An example is the map $v_\odot \in W^{1,2}(\mathbb{R}P^4,\mathbb{R}P^3)$ defined by $v_\odot[x^0 : x^1 : x^2 : x^3 : x^4] = [x^1 : x^2 : x^3 : x^4]$, with a singularity at $[1:0:0:0:0]$; there is no way to remove this singularity[14], so, *there is no sequence of smooth maps converging weakly to $v_\odot$*. Hence $H^{1,2}_w(\mathbb{R}P^4,\mathbb{R}P^3) \neq W^{1,2}(\mathbb{R}P^4,\mathbb{R}P^3)$, although $\pi_2(\mathbb{R}P^3) = 0$. A result due to P. Hajłasz [96] is valid for an arbitrary manifold $\mathcal{M}$:

$$\textit{if } 1 \le p < m, \textit{ then} \quad \pi_1(\mathcal{N}) = \cdots = \pi_{[p]}(\mathcal{N}) = 0 \implies H^{1,p}_s(\mathcal{M},\mathcal{N}) = W^{1,p}(\mathcal{M},\mathcal{N}).$$

[14] In contrast with the map $u_\odot \in W^{1,2}(B^4,S^3)$ where the topological singularity can be moved to the boundary with an arbitrary low energy cost.

The general result is due to F. B. Hang and F. H. Lin [102] and, in the case where $\mathcal{M}$ has no boundary, is the following. First we say that $\mathcal{M}$ *satisfies the k-extension property with respect to $\mathcal{N}$* if, for any CW complex structure $(X^j)_{j\in\mathbb{N}}$ on $\mathcal{M}$ and for any $f \in \mathcal{C}^0(X^{k+1}, \mathcal{N})$, the restriction $f|_{X^k}$ of f on X^k has a continuous extension to $\mathcal{M}$. Then, if $1 < p < m$, we have [102]:

$$H_s^{1,p}(\mathcal{M},\mathcal{N}) = W^{1,p}(\mathcal{M},\mathcal{N}) \Longleftrightarrow \begin{cases} \pi_{[p]}(\mathcal{N}) = 0 \textit{ and } \mathcal{M} \textit{ satisfies} \\ \textit{the } [p-1]\textit{-extension property with respect to } \mathcal{N}. \end{cases}$$

The case when p is not an integer. The identity between $H_s^{1,p}(B^m, S^{m-1})$ and $H_w^{1,p}(B^m, S^{m-1})$ for $p \neq m-1$ is actually a particular case of a general phenomenon, as shown by Bethuel [12]: *for any domain $\mathcal{M} \subset \mathcal{M}$ and for any compact manifold $\mathcal{N}$,*

$$\textit{if } p > 1 \textit{ is } \textbf{not an integer}\textit{, then} \quad H_s^{1,p}(\mathcal{M},\mathcal{N}) = H_w^{1,p}(\mathcal{M},\mathcal{N}).$$

The case when p is an integer. The question left open is, in cases where $H_s^{1,p}(\mathcal{M},\mathcal{N}) \subsetneq W^{1,p}(\mathcal{M},\mathcal{N})$, to characterize the intermediate space $H_w^{1,p}(\mathcal{M},\mathcal{N})$. A first answer was given in [12] for maps into the sphere:

$$\textit{if } p \in \mathbb{N} \textit{ satisfies } p < m, \textit{ then} \quad H_s^{1,p}(B^m, S^p) \subsetneq H_w^{1,p}(B^m, S^p) = W^{1,p}(B^m, S^p).$$

A generalization was proved by P. Hajłasz in [96]:

$$\begin{aligned} &\textit{if } p \in \mathbb{N} \textit{ satisfies } p < m, \\ &\textit{then} \quad \pi_1(\mathcal{N}) = \cdots = \pi_{p-1}(\mathcal{N}) = 0 \quad \Longrightarrow \quad H_w^{1,p}(\mathcal{M},\mathcal{N}) = W^{1,p}(\mathcal{M},\mathcal{N}). \end{aligned}$$

And the following further result was obtained by M. R. Pakzad and T. Rivière [162]:

$$\textit{for } p = 2, \quad \pi_1(\mathcal{N}) = 0 \quad \Longrightarrow \quad H_w^{1,2}(\mathcal{M},\mathcal{N}) = W^{1,2}(\mathcal{M},\mathcal{N}).$$

For more general situations, assuming that $\mathcal{M}$ has no boundary, a *necessary* condition for a map to be in $H_w^{1,p}(\mathcal{M},\mathcal{N})$ was found by F. B. Hang and F. H. Lin in [102]: they proved that **if** $u \in H_w^{1,p}(\mathcal{M},\mathcal{N})$ **then** $u_{\sharp,[p]-1}(h)$ **is extendible to** $\mathcal{M}$ **with respect to** $\mathcal{N}$. The precise definition of $u_{\sharp,[p]-1}(h)$ is delicate: roughly speaking, by using ideas of B. White (see [226, 227] and §3.3), it is possible to define the homotopy class $u_{\sharp,[p]-1}(h)$ of the restriction of a map $u \in H_w^{1,p}(\mathcal{M},\mathcal{N})$ to a *generic* $([p]-1)$-skeleton of a rectilinear cell decomposition h of $\mathcal{M}$. Furthermore Hang and Lin in [102] *conjectured*[15] *that this condition is also a sufficient one*, i.e., that *if $u \in W^{1,p}(\mathcal{M},\mathcal{N})$ and $u_{\sharp,[p]-1}(h)$ is extendible to $\mathcal{M}$ with respect to $\mathcal{N}$, then $u \in H_w^{1,p}(\mathcal{M},\mathcal{N})$*. In [101] Hang proved that **this conjecture is true for** $p = 2$.

Note that in the special case $p = 1$, Hang proved that $H_s^{1,1}(\mathcal{M},\mathcal{N}) = H_w^{1,1}(\mathcal{M},\mathcal{N})$ [100].

3.3 The topology of $W^{1,p}(\mathcal{M},\mathcal{N})$

The motivation for understanding the topology of $W^{1,p}(\mathcal{M},\mathcal{N})$ is to adapt the *direct method* of the calculus of variations to find a harmonic map in a homotopy class of maps between $\mathcal{M}$ and $\mathcal{N}$, i.e., by minimizing the energy in this homotopy class.

[15] They proved this conjecture for maps in $\mathcal{R}_w^{k,p}(\mathcal{M},\mathcal{N})$.

Some difficulties are illustrated by the following question[16] [66]: *What is the infimum of the energy in the homotopy class of the identity map* Id : $S^m \longrightarrow S^m$ *?*

- if $m = 1$, Id is minimizing and all minimizers in its homotopy class are rotations.
- if $m \geq 3$, **the infimum of the energy is** 0. Indeed, consider, for example, the family of conformal Möbius maps $T_\lambda : S^m \longrightarrow S^m$ for $\lambda \in \mathbb{R}$ defined by $T_\lambda(y) = (\cosh\lambda + y^1 \sinh\lambda)^{-1}(\sinh\lambda + y^1\cosh\lambda, y^2, \cdots, y^m)$; for all $\lambda \in \mathbb{R}$, T_λ is homotopic to the identity (actually T_0 equals the identity map) but as λ goes to $+\infty$, $E_{S^m}(T_\lambda)$ tends to zero and T_λ converges *strongly* to a constant map.
- the intermediate case $m = 2$ corresponds to the *critical dimension*; then *all the maps* T_λ *have the same energy, are conformal harmonic, and minimize the energy in their homotopy class, but* T_λ *converges* weakly *to a constant map*[17] *as* $\lambda \to +\infty$. One then speaks of a *bubbling* phenomenon, see §5.3.

Prescribing the action on the first homotopy group. The first positive result in these directions was in the case $m = \dim \mathcal{M} = 2$ and $\partial\mathcal{M} = \emptyset$ studied by R. Schoen and S.T. Yau [193]. Let γ be a smooth immersed path in $\mathcal{M}$ and $u \in W^{1,2}(\mathcal{M}, \mathcal{N})$; in general the 'restriction' $u \circ \gamma$ of u to γ is not continuous (just in $W^{\frac{1}{2},2}$) but one can prove that, if we change γ to a *generic* path $\widetilde{\gamma}$ which is homotopic to γ, then $u \circ \widetilde{\gamma}$ *is* continuous.

(i) First, we use the following observation[18]: for any map $f \in W^{1,2}(S^1 \times (0,1), \mathbb{R}^N)$, the map $(\theta, s) \longmapsto |df(\theta, s)|^2$ is in $L^1(S^1 \times (0,1))$; hence by using the Fubini–Study theorem on $S^1 \times (0,1)$ one deduces that, for a.e. $s \in (0,1)$, the map $\theta \longmapsto |df(\theta, s)|^2$ belongs to $L^1(S^1)$, so that the restriction of f to $S^1 \times \{s\}$ is in $W^{1,2}(S^1) \subset \mathcal{C}^0(S^1)$. We apply this result to $f = u \circ \Gamma$, where $\Gamma \in \mathcal{C}^1(S^1 \times (0,1), \mathcal{M})$ parametrizes a strip composed of parallel paths $\gamma_s := \Gamma(\cdot, s)$ homotopic to the same path γ.

(ii) Second, if $s_1 < s_2$ are two values in $(0,1)$ such that $u \circ \gamma_{s_1}$ and $u \circ \gamma_{s_2}$ are continuous, then we can use the existence theorem of Morrey [156] to prove that there exists a smooth minimizing harmonic map $U : S^1 \times (s_1, s_2) \longrightarrow \mathcal{N}$ which agrees with $u \circ \Gamma$ on $\partial S^1 \times (s_1, s_2) = (S^1 \times \{s_2\}) \cup (S^1 \times \{s_1\})$. We deduce that $u \circ \gamma_{s_1}$ *and* $u \circ \gamma_{s_2}$ *are homotopic*.

This leads to the definition of the image by u of the homotopy class of γ: it is the homotopy class of $u \circ \gamma_s$, where γ_s is a path in the same homotopy class as γ, which is generic in the above sense. We can thus define the induced conjugacy class of homorphisms

$$u_{\sharp 1} : \pi_1(\mathcal{M}) \longrightarrow \pi_1(\mathcal{N}).$$

One can check, moreover, that *this homomorphism is preserved by weak convergence in* $W^{1,2}(\mathcal{M}, \mathcal{N})$, i.e., if v_k converges weakly to u in $W^{1,2}$ when $k \to +\infty$, and $\forall k \in \mathbb{N}$, $(v_k)_{\sharp 1} = v_{\sharp 1}$ for some $v \in \mathcal{C}^0(\mathcal{M}, \mathcal{N})$, then $u_{\sharp 1} = v_{\sharp 1}$. Eventually this leads to the

[16] The displayed facts were noticed by C. B. Morrey.

[17] This is a consequence of the following observations: on the one hand by using the standard compactness arguments we can extract a subsequence of $(v_k)_{k \in \mathbb{N}}$ which converges weakly in $W^{1,2}$ and a.e. to some limit v, but on the other hand it is clear that v_k converges a.e. (and more precisely pointwise on $S^m \setminus \{(-1, 0, \cdots, 0)\}$) to $(1, 0, \cdots, 0)$, so that $v = (1, 0, \cdots, 0)$. Since this argument works for *any subsequence* the full sequence $(v_\lambda)_{\lambda > 0}$ converges weakly to this constant.

[18] Which itself is the key ingredient of the classical Courant–Lebesgue lemma, see, for example, [88, 3.3.1].

following **existence result** of Schoen and Yau [193]: *assume that $\mathcal{M}$ is surface without boundary. Then, for any family $\gamma_1, \cdots, \gamma_k$ of loops in $\mathcal{M}$ and for any continuous map $v : \mathcal{M} \longrightarrow \mathcal{M}$, there exists a locally energy-minimizing harmonic map in the class of maps $u \in W^{1,2}(\mathcal{M}, \mathcal{N})$ such that $u_{\sharp 1}([\gamma_i]) = v_{\sharp 1}([\gamma_i])$, $\forall i = 1, \cdots, k$.* This result has been generalized to the case where the dimension of $\mathcal{M}$ is arbitrary by F. Burstall [30] and B. White [225].

Remarks (i) Note that, if $\pi_j(\mathcal{N}) = 0$ for $j \geq 2$, then the homotopy class of a continuous map u from $\mathcal{M}$ to $\mathcal{N}$ is completely characterized by the induced conjugacy class of homomorphisms $u_{\sharp 1} : \pi_1(\mathcal{M}) \longrightarrow \pi_1(\mathcal{N})$; thus, when $m = 2$, the existence result of Schoen and Yau amounts to minimizing the energy in a given homotopy class of continuous maps between $\mathcal{M}$ and $\mathcal{N}$ (recall that continuous maps are then dense in $W^{1,2}(\mathcal{M}, \mathcal{N})$).

(ii) The definition of $u_{\sharp 1} : \pi_1(\mathcal{M}) \longrightarrow \pi_1(\mathcal{N})$ **does not make sense** if $u \in W^{1.p}(\mathcal{M}, \mathcal{N})$ for $1 \leq p < 2$. Indeed, as in step (i), we still have that $u \circ \gamma_s$ is continuous for a generic s, but step (ii) does not work: the homotopy class of $u \circ \gamma_s$ can vary as s changes (see B. White [227] or J. Rubinstein and P. Sternberg [186]).

Defining the d-homotopy class. For any $d \in \mathbb{N}$, we say that *two maps $u, v \in \mathcal{C}^0(\mathcal{M}, \mathcal{N})$ are d-homotopic* and we write $u \sim_d v$ *if their restrictions to the d-skeleton of a triangulation of $\mathcal{M}$ are homotopic*. For any map $u \in \mathcal{C}^0(\mathcal{M}, \mathcal{N})$ we thus can define the *d-homotopy class* $[u]_d := \{v \in \mathcal{C}^0(\mathcal{M}, \mathcal{N}) |\ u \sim_d v\}$. Observe that **if** $u \sim_d v$ **then the induced homomorphisms** $u_{\sharp j}, v_{\sharp j} : \pi_j(\mathcal{M}) \longrightarrow \pi_j(\mathcal{N})$ **coincide for each** $1 \leq j \leq d$, so that this notion extends the previous one. Actually A.I. Pluzhnikov [171] and B. White [227] showed that *it is possible to define the d-homotopy class of a map u in $H^{1,p}_s$, $H^{1,p}_w$ or $W^{1,p}(\mathcal{M}, \mathcal{N})$ for certain ranges of values of d and p*. The following table summarizes the results proved in [227]. It gives, for each space $H^{1,p}_s$, $H^{1,p}_w$ or $W^{1,p}$, the values of d for which one can define the d-homotopy class of a map u in this space, and it specifies natural topologies which preserve this d-homotopy class:

Spaces	$H^{1,p}_s(\mathcal{M}, \mathcal{N})$	$H^{1,p}_w(\mathcal{M}, \mathcal{N})$	$W^{1,p}(\mathcal{M}, \mathcal{N})$
Values of d for which $[u]_d$ makes sense:	$\mathbb{N} \cap [1, p]$	$\mathbb{N} \cap [1, p)$	$\mathbb{N} \cap [1, p-1]$
Topology which preserves $[u]_d$:	strong $W^{1,p}$	weak $W^{1,p}$	weak $W^{1,p}$

The definition of $[u]_d$ for $u \in H^{1,p}_w(\mathcal{M}, \mathcal{N})$ when $d < p$ follows from the following result [171, 226]: *if $d \in \mathbb{N}$ and $d < p$, then $\forall K > 0$, $\exists \varepsilon > 0$, such that if u_1 and u_2 are two Lipschitz continuous maps such that $||u_1||_{W^{1,p}}, ||u_2||_{W^{1,p}} < K$ and $||u_1 - u_2||_{L^p} < \varepsilon$, then $u_1 \sim_d u_2$.* Hence one can define the d-homotopy class of a given $u \in H^{1,p}_w(\mathcal{M}, \mathcal{N})$ by using any sequence of Lipschitz continuous maps $(v_k)_{k\in\mathbb{N}}$ which converges weakly to u in $W^{1,p}$ and setting $[u]_d := [v_k]_d$ for k large enough. For $u \in H^{1,p}_s(\mathcal{M}, \mathcal{N})$, the previous argument applies also when defining $[u]_d$ if $d < p$; if $d = p$ we must use a further approximation argument.

In constrast, the definition of $[u]_d$ for $u \in W^{1,p}(\mathcal{M}, \mathcal{N})$ and $d \leq p - 1$ cannot be obtained by using approximations by smooth maps, but must be done directly. Here the idea consists of proving that the restriction of u on a *generic* d-skeleton is continuous and that the homotopy class of this restriction is independent of the d-skeleton, following a strategy similar to the result of Schoen and Yau. The details of the proof are, however, more involved.

The k-homotopy type helps to characterize the topology of the spaces $H_s^{1,p}$ and $W^{1,p}(\mathcal{M},\mathcal{N})$, as follows.

Connected components of $H_s^{1,p}(\mathcal{M},\mathcal{N})$. For any $u \in H_s^{1,p}(\mathcal{M},\mathcal{N})$ denote by $[u]_{H_s^{1,p}}$ its connected component in $H_s^{1,p}(\mathcal{M},\mathcal{N})$ for the strong $W^{1,p}$-topology. The classes $[u]_{H_s^{1,p}}$ have been characterized by A.I. Pluzhnikov [171] and B. White [226] as follows: *the connected components of $H_s^{1,p}(\mathcal{M},\mathcal{N})$ are exactly the $[p]$-homotopy classes inside $H_s^{1,p}(\mathcal{M},\mathcal{N})$*. In other words, for any $u \in H_s^{1,p}(\mathcal{M},\mathcal{N})$, $[u]_{H_s^{1,p}} = [u]_{[p]}$.

This has the following important consequence: *for any smooth map $v \in \mathcal{C}^1(\mathcal{M},\mathcal{N})$, the infimum of the p-energy among smooth maps in the homotopy class of v depends uniquely on the $[p]$-homotopy type of v.* A further result is: *for a smooth map v, $v \sim_{[p]} C$ (where C is a constant map) if and only if the infimum of the p-energy in $[v]_{[p]}$ is* 0 [171, 226]. Note that the limit of a minimizing sequence of the p-energy in a $[p]$-homotopy class $[v]_{[p]}$ may not be in $[v]_{[p]}$, but only in its closure for the sequential weak topology of $W^{1,p}$ in general. See the example with $\mathcal{M} = \mathcal{N} = S^m$, $v = \mathrm{Id}$ discussed at the beginning of this section.

Connected components of $W^{1,p}(\mathcal{M},\mathcal{N})$. For $u \in W^{1,p}(\mathcal{M},\mathcal{N})$ denote by $[u]_{W^{1,p}}$ its connected component. The study of the connected components of $W^{1,p}(\mathcal{M},\mathcal{N})$ was initiated by H. Brezis and Y. Li [25]. Complete answers were obtained by F. B. Hang and F. H. Lin [102] as follows:

(i) *The connected components of $W^{1,p}(\mathcal{M},\mathcal{N})$ are path-connected.* This is a consequence of the following: $\forall u \in W^{1,p}(\mathcal{M},\mathcal{N})$, $\exists \varepsilon > 0$ such that $\forall v \in W^{1,p}(\mathcal{M},\mathcal{N})$, if $||u - v||_{W^{1,p}} < \varepsilon$, then there exists $U \in \mathcal{C}^0([0,1], W^{1,p}(\mathcal{M},\mathcal{N}))$ such that $U(0,\cdot) = u$ and $U(1,\cdot) = v$. We write $u \sim_{W^{1,p}} v$ for this property.

(ii) **the connected components of** $W^{1,p}(\mathcal{M},\mathcal{N})$ **are exactly the** $([p]-1)$**-homotopy classes inside** $W^{1,p}(\mathcal{M},\mathcal{N})$, i.e. $\forall u, v, \in W^{1,p}(\mathcal{M},\mathcal{N})$, $u \sim_{W^{1,p}} v$ if and only if $u \sim_{[p]-1} v$.

(iii) *as p varies, the quotient space $W^{1,p}(\mathcal{M},\mathcal{N})/\sim_{W^{1,p}}$ changes only for integer values of p*, i.e. if $[p_1] = [p_2] < p_1 < p_2 < [p_1] + 1$, the map $\iota_{p_2,p_1} : W^{1,p_2}(\mathcal{M},\mathcal{N})/{\sim_{W^{1,p_2}}} \longrightarrow W^{1,p_1}(\mathcal{M},\mathcal{N})/{\sim_{W^{1,p_1}}}$ induced by the inclusion $W^{1,p_2}(\mathcal{M},\mathcal{N}) \subset W^{1,p_1}(\mathcal{M},\mathcal{N})$ is a bijection (this was conjectured in [25]).

Result (ii) has the following corollary: *a map $u \in W^{1,p}(\mathcal{M},\mathcal{N})$ is connected to a smooth map by a path if and only if $u_{\sharp,[p]-1}$ is extendible to $\mathcal{M}$ with respect to $\mathcal{N}$.* This implies, in particular, the results (also proved in [25]):

- if $\forall j \in \mathbb{N}$ such that $1 \le j \le [p] - 1$ we have $\pi_j(\mathcal{N}) = 0$, then $W^{1,p}(\mathcal{M},\mathcal{N})$ is path-connected;
- if $p < m$, then $W^{1,p}(S^m,\mathcal{N})$ is path-connected.

Concerning (iii), the change in the number of connected components of $W^{1,p}(\mathcal{M},\mathcal{N})$ when p varies can occur in two ways. Indeed, *as p decreases*, *either* connected components coalesce together — this is, for example, the case for $W^{1,p}(S^m, S^m)$: this space has different connected components classified by the topological degree if $p \ge m$ and is connected if $p < m$; *or*, contradicting a conjecture in [25], new connected components can appear — this is the case for $W^{1,p}(\mathbb{R}P^3, \mathbb{R}P^2)$: for $p \in (2,3)$ connected components appear, forming a subset of maps which cannot be connected by a path to a smooth map (and which hence cannot be approximated by smooth maps), see [21, 102].

The degree. If $\dim \mathcal{M} = \dim \mathcal{N}$, the homotopy classes of maps $\mathcal{M} \longrightarrow \mathcal{N}$ can sometimes be classified by the topological degree. This is the case if, for instance, $\mathcal{M}$ is *connected, oriented*[19] *and without boundary and if* $\mathcal{N} = S^m$ (by a theorem of H. Hopf)[20]. The degree for a map $u \in \mathcal{C}^1(\mathcal{M}, S^m)$ is then given by the formula $\deg u = \left(1/|S^m|\right)\int_{\mathcal{M}} \det(du)\omega_{\mathcal{M}} = \left(1/|S^m|\right)\int_{S^m} u^*\omega_{S^m}$. We give this formula explicitly for the case $p = 2$:

$$\deg u = \frac{1}{4\pi}\int_{\mathcal{M}} u^*\omega_{S^2} = \frac{1}{4\pi}\int_{\mathcal{M}} \left\langle u, \frac{\partial u}{\partial x} \times \frac{\partial u}{\partial y}\right\rangle dx\, dy,$$

where (x, y) are local conformal coordinates on $\mathcal{M}$. This functional, being quadratic in the first derivatives of u, has the following continuity properties:

(i) it is continuous on $\mathcal{C}^1(\mathcal{M}, S^2)$ for the *strong* and the *weak* $W^{1,p}$ topology for all $p > 2$, hence for $p > 2$ we can extend deg on $H_s^{1,p}(\mathcal{M}, S^2) = W^{1,p}(\mathcal{M}, S^2)$;

(ii) it is continuous on $\mathcal{C}^1(\mathcal{M}, S^2)$ for the *strong* **but not** for the *weak* $W^{1,p}$ topology for $p = 2$, hence since $H_s^{1,2}(\mathcal{M}, S^2) = W^{1,2}(\mathcal{M}, S^2)$ we can extend deg on $W^{1,2}(\mathcal{M}, S^2)$, but this functional is not continuous with respect to the weak topology;

(iii) it is **not** continuous on $\mathcal{C}^1(\mathcal{M}, S^2)$ for the *strong* or the *weak* $W^{1,p}$ topology for all $p < 2$.

In cases (i) and (ii) ($p \geq 2$), the degree functional takes integer values and, $\forall k \in \mathbb{N}$, $\deg^{-1}(k)$ is a connected component of $W^{1,p}(\mathcal{M}, S^2)$ for its strong topology. In case (i), the continuity for the weak topology follows from the fact that, on the one hand, for a sequence $(u_k)_{k\in N}$ which converges weakly to some u in $W^{1,p}(\mathcal{M}, S^2)$, $f_k := (\partial_x u_k) \times (\partial_y u_k)$ converges weakly in $L^{p/2}$ to $f := (\partial_x u) \times (\partial_y u)$, because of a phenomenon of *compensated compactness*, based on writing $f_k = \partial_x\left(u_k(\partial_y u_k)\right) - \partial_y\left(u_k(\partial_x u_k)\right)$ (see [159, 211]). On the other hand, by the Rellich–Kondrakov theorem, we can assume that $u_k \to u$ strongly in $L^{2p/p-2}$ and hence in $L^{p/p-2} = \left(L^{p/2}\right)^*$. It follows that the integral $\int_{\mathcal{M}}\langle u_k, f_k\rangle\omega_{\mathcal{M}}$ converges to $\int_{\mathcal{M}}\langle u, f\rangle\omega_{\mathcal{M}}$. This delicate argument breaks down[21] for $p = 2$: we still have that f_k converges in the weak-$\star$ topology of L^1, but we *cannot find*, in general, a subsequence of u_k which converges strongly in L^∞ (otherwise we would have an embedding of $W^{1,2}(\mathcal{M})$ in $\mathcal{C}^0(\mathcal{M}) \subset L^\infty(\mathcal{M})$!). Indeed, in the case where $\mathcal{M} = S^2$, the family of (degree 1) Möbius maps $(T_\lambda)_{\lambda\in\mathbb{R}}$ converges weakly to a constant map in $W^{1,2}(S^2, S^2)$ as $\lambda \to +\infty$ (a *bubbling* phenomenon, see §5.3). Lastly (iii) can seen by considering the family of maps $(u_t)_{t\in[0,1]}$ from S^2 to S^2 defined by $u_t(x) =$

[19] If $\mathcal{M}$ is connected, without boundary but *not oriented*, the homotopy classes are classified by the degree mod 2.

[20] But if $\mathcal{M}$ and $\mathcal{N}$ are spheres with different dimensions, this is not so, for example, maps from S^3 to S^2 are classified according to their Hopf degree, see [110].

[21] A rich interplay between cohomology and compensated compactness theory occurs here: for any smooth function $\psi \in \mathcal{C}^1(\mathcal{M})$ and any 2-form β on S^2 which is *exact*, i.e., $\beta = d\alpha$ for some 1-form α, the functional $u \longmapsto \int_{\mathcal{M}} \psi\, u^*\beta$ is continuous for the weak $W^{1,2}$ topology because of the relation $u^*\beta = d(u^*\alpha)$, so that a compensated compactness argument is possible; however, if β is *closed but not exact*, this argument does not work. See [98] for a detailed study of these phenomena.

$(x-ta)/|x-ta|$, where $a \in \mathbb{R}^3$ has $|a| = 2$; for $1 \le p < 2$, this defines a continuous path in $W^{1,p}(S^2, S^2)$, which connects the smooth map $u_\odot = u_0$ of degree 1 to the smooth map u_1 of degree 0 (see [27, 21]).

Lastly, in [26] H. Brezis, Y. Li, P. Mironescu and L. Nirenberg defined a notion of *degree for maps* $u \in W^{1,p}(S^n \times \Lambda^{m-n}, S^n)$*, where* $m \ge n$ *and* Λ^{m-n} *is an open connected subset of* $\mathbb{R}^{m-n}$*, assuming that* $p \ge n+1$ (note that, in the special case $m = n$, the condition $p \ge n$ is enough). In the case $n = 1$, we recover from this result the conclusions of [30, 227, 186]. Furthermore, *two maps* u *and* v *in* $W^{1,p}(S^n \times \Lambda^{m-n}, S^n)$ *are in the same connected component if and only if deg* f *= deg* g (see [26, 21]). See [27] for further results concerning the degree.

3.4 The trace of Sobolev maps

For any domain $\Omega \subset \mathbb{R}^m$ with smooth boundary and for any $p \in [1, +\infty)$, the trace operator $\mathrm{tr}\colon \mathcal{C}^1(\Omega, \mathbb{R}^N) \longrightarrow \mathcal{C}^1(\partial\Omega, \mathbb{R}^N)$ can be extended to a continuous and surjective operator $\mathrm{tr} : W^{1,p}(\Omega, \mathbb{R}^N) \longrightarrow W^{1-\frac{1}{p},p}(\partial\Omega, \mathbb{R}^N) := \big\{ g \in L^p(\partial\Omega, \mathbb{R}^N) |\ ||g||_{W^{1-\frac{1}{p},p}(\partial\Omega)} < +\infty \big\}$ if $p > 1$, where:

$$||g||_{W^{1-\frac{1}{p},p}(\partial\Omega)} := ||g||_{L^p(\partial\Omega)} + \left(\int_{\partial\Omega} \int_{\partial\Omega} \frac{|g(x) - g(y)|^p}{|x-y|^{p+m-2}}\, dx\, dy \right)^{1/p} .$$

(If $p = 1$, the image of the trace operator is $L^1(\partial\Omega, \mathbb{R}^N)$.) This definition can be extended to the case of a manifold $\mathcal{M}$ with a smooth boundary, by using local charts to define $W^{1-\frac{1}{p},p}(\partial\mathcal{M}, \mathbb{R}^N)$ and the trace operator $\mathrm{tr} : W^{1,p}(\mathcal{M}, \mathbb{R}^N) \longrightarrow W^{1-\frac{1}{p},p}(\partial\mathcal{M}, \mathbb{R}^N)$. Similarly the trace $\mathrm{tr}\, u$ of a map $u \in W^{1,p}(\mathcal{M}, \mathcal{N})$ is always contained in:

$$W^{1-\frac{1}{p},p}(\partial\mathcal{M}, \mathcal{N}) := \big\{ g \in W^{1-\frac{1}{p},p}(\partial\mathcal{M}, \mathbb{R}^N) |\ g(x) \in \mathcal{N},\ \text{for a.e. } x \in \partial\mathcal{M} \big\}.$$

However, the map $\mathrm{tr} : W^{1,p}(\mathcal{M}, \mathcal{N}) \longrightarrow W^{1-\frac{1}{p},p}(\partial\mathcal{M}, \mathcal{N})$ is not onto in general, i.e., it is not true in general that any map $g \in W^{1-\frac{1}{p},p}(\partial\mathcal{M}, \mathcal{N})$ is the trace of a map in $W^{1,p}(\mathcal{M}, \mathcal{N})$. Obstructions occur even for continuous maps: for instance, the trace operator $\mathrm{tr} : \mathcal{C}^1(B^m, \mathcal{N}) \longrightarrow \mathcal{C}^1(\partial B^m, \mathcal{N})$ is onto if and only if $\pi_{m-1}(\mathcal{N}) = 0$. In the following we define $T^p(\partial\mathcal{M}, \mathcal{N}) := \big\{ g \in W^{1-\frac{1}{p},p}(\partial\mathcal{M}, \mathcal{N}) |\ \exists u \in W^{1,p}(\mathcal{M}, \mathcal{N}) \text{ such that } u|_{\partial\mathcal{M}} = g \big\}$. The question whether $T^p(\partial\mathcal{M}, \mathcal{N}) = W^{1-\frac{1}{p},p}(\partial\mathcal{M}, \mathcal{N})$ for given $\mathcal{M}$, $\mathcal{N}$ and p is largely open. Here are some results:

- If $p \ge m$, F. Bethuel and F. Demengel [16] proved that $T^p(\partial\mathcal{M}, \mathcal{N}) = W^{1-\frac{1}{p},p}(\partial\mathcal{M}, \mathcal{N})$ **if and only if any continuous map** $g \in \mathcal{C}^0(\partial\mathcal{M}, \mathcal{N})$ **can be extended to a map** $u \in \mathcal{C}^0(\mathcal{M}, \mathcal{N})$.
- For $1 \le p < m$, R. Hardt and F. H. Lin [107] proved that

$$\text{if} \quad \pi_1(\mathcal{N}) = \cdots = \pi_{[p]-1}(\mathcal{N}) = 0, \quad \text{then} \quad T^p(\partial\mathcal{M}, \mathcal{N}) = W^{1-\frac{1}{p},p}(\partial\mathcal{M}, \mathcal{N}) .$$

- Conversely Bethuel and Demengel [16] proved that, *if* $1 \le p < m$*, then* $\pi_{[p]-1}(\mathcal{N}) = 0$ *is a necessary condition for having* $T^p(\partial\mathcal{M}, \mathcal{N}) =$

$W^{1-\frac{1}{p},p}(\partial\mathcal{M},\mathcal{N})$. Moreover, they proved that, *if* $1 < p < m$, *then, for any* $\mathcal{N}$ *such that* $\pi_j(\mathcal{N}) \neq 0$ *for some integer* $j \leq [p]-1$, *one can construct a manifold with boundary* $\mathcal{M}$ *such that* $T^p(\partial\mathcal{M},\mathcal{N}) \neq W^{1-\frac{1}{p},p}(\partial\mathcal{M},\mathcal{N})$.

Furthermore it is proved in [16] that, in the case where $\mathcal{M} = B^m$ and $\mathcal{N} = S^1$, *if* $3 \leq p < m$ *then* $T^p(\partial B^m, S^1) \neq W^{1-\frac{1}{p},p}(\partial B^m, S^1)$. For more results on fractional Sobolev spaces into S^1, see the report of P. Mironescu [155] or the papers [20, 183].

4 Regularity

4.1 Regularity of *continuous* weakly harmonic maps

Note that *as soon as we know that a (weakly) harmonic map* ϕ *is continuous*, then we can *localize* its image, i.e. by restricting ϕ to a sufficiently small ball in $\mathcal{M}$ *we can assume that the image of* ϕ *is contained in an arbitrary small subset of* $\mathcal{N}$ *with good convexity properties or with a convenient coordinate system*. Thus the main results concern the **higher regularity** of *continuous weakly harmonic maps*. The hard step here is to prove that the weak solution ϕ is Lischiptz continuous, i.e. that $d\phi$ is bounded a.e.[22]. This was proved by O. Ladyzhenskaya, N. Ural'tseva in [143] in a more general context, by using contributions of C. B. Morrey [157], a proof can be found in [135]. In [189], a proof is given in the case when the weakly harmonic map is *Hölder* continuous. Estimates of the Hölder norms of higher derivatives of ϕ in terms of $|d\phi|$ were obtained by J. Jost and H. Karcher [136] for harmonic maps with values in a geodesically convex ball: on such balls they construct and use *almost linear functions* (which are based on *harmonic coordinates*, in which the Hölder norm of Christoffel symbols are bounded in terms of the curvature).

4.2 Regularity results in dimension two

If dim $\mathcal{M} = 2$ and $\mathcal{N}$ can be embedded isometrically in some Euclidean space, **all weakly harmonic maps in** $W^{1,2}(\mathcal{M},\mathcal{N})$ **are continuous** and hence, by the results of §4.1, smooth. This was proved first for *minimizing maps* by C. B. Morrey [156] (see also [88, p. 304] for an exposition of the original proof of Morrey).

This was extended to *conformal weakly harmonic maps* by M. Grüter [92] (see also [133]). Grüter's proof works also for conformal weak solutions of the H*-system* $\Delta_g u + A(u)(du,du) = 2H(u)(\partial u/\partial x^1 \times \partial u/\partial x^2)$ in an oriented 3-dimensional manifold $\mathcal{N}$, where $H(u)$ is a L^∞ bounded function on $\mathcal{N}$. Conformal solutions to this problem parametrize surfaces with prescribed mean curvature H. The proof in [92] uses the conformality assumption in an essential way. Then R. Schoen [189] proved that *all stationary maps on a surface are smooth*. The proof is based on the following trick. Let $u \in W^{1,2}(\mathcal{M},\mathcal{N})$ be a stationary map; since the Hopf differential $\mathcal{H}$ is holomorphic (see §3.1), *either* it vanishes everywhere and then u is conformal and we apply directly the result of Grüter, *or* $\mathcal{H} = h(dz)^2$ vanishes only at isolated points. If so, outside the zeros of h

[22] Once we know that $d\phi \in L^\infty_{loc}$, it then follows from (25) that $\Delta\phi \in L^\infty_{loc}$, which implies by standard estimates on the inverse of the Laplacian (see [157], 6.2.5) that $\phi \in W^{2,p}_{loc}$, for all $p < \infty$. Hence we deduce that $\Delta\phi \in W^{1,p}_{loc}$ and hence that $\phi \in W^{3,p}_{loc}$ for all $p > 0$. We can then repeat this argument to show that $\phi \in W^{r,p}_{loc}, \forall r$ and so the smoothness of the solution follows (it is called a *bootstrap* argument).

we can locally define the harmonic function $f(z) := \mathrm{Re}\big(2i \int_{z_0}^{z} \sqrt{h(\zeta)} d\zeta\big)$. Then the map $U := (u, f)$ with values in $\mathcal{N} \times \mathbb{R}$ is weakly harmonic and conformal and hence is smooth. Thus u is smooth outside the zeros of h, and hence is smooth everywhere by the result of J. Sacks and K. Uhlenbeck [188] (see §5.3).

The regularity of weakly harmonic maps on a surface in *the general case* was proved by F. Hélein, first in the special case where $\mathcal{N} = S^n$ [113], and then in the case where $\mathcal{N}$ is an arbitrary compact Riemannian manifold without boundary [116]. The proof for $\mathcal{N} = S^n$ is simpler and relies on a previous work by H. Wente [221] on the solutions $X \in W^{1,2}(B^2, \mathbb{R}^3)$ on the unit ball[23] of $\mathbb{R}^2$ of the H-system

$$\Delta X = 2H \frac{\partial X}{\partial x} \times \frac{\partial X}{\partial y}\,, \tag{34}$$

for a constant $H \neq 0$. Wente proved that any weak solution of this system is continuous and hence, thanks again to the general theory of quasilinear elliptic systems, smooth. It is based on the special structure of (34) which reads, for example for the first component of X, $\Delta X^1 = 2H\{X^2, X^3\}$, where we introduce the notation

$$\{a, b\} := \frac{\partial a}{\partial x}\frac{\partial b}{\partial y} - \frac{\partial a}{\partial y}\frac{\partial b}{\partial x} \qquad \text{for } a, b \in W^{1,2}(\Omega), \quad \text{where } \Omega \subset \mathbb{R}^2.$$

Since $\{a, b\}$ is quadratic in the first derivatives of a and b, it sits naturally in $L^1(B^2)$. Also, we know from the standard theory of singular integrals that, for any function $f \in L^1(B^2)$, a solution ψ of $-\Delta\psi = f$ is necessarily in all spaces $L^p_{loc}(B^2)$, for $1 \leq p < \infty$, but fails to be in $L^\infty(B^2)$. Here the key result is that a solution φ of the equation $-\Delta\varphi = \{a, b\}$ on B^2 is slightly more regular; in particular, we can locally estimate the L^∞ norm of φ in terms of $||a||_{W^{1,2}}$ and $||b||_{W^{1,2}}$. This is due to the special structure of $\{a, b\}$, which is a *Jacobian determinant*, and is connected to the theory of compensated compactness [159, 211]. These properties were expressed by H. Brezis and J.-M. Coron [23] as a *Wente inequality*,

$$||\varphi||_{L^\infty} + ||d\varphi||_{L^2} \leq C||a||_{W^{1,2}}||a||_{W^{1,2}}\,, \tag{35}$$

valid for any solution φ of $-\Delta\varphi = \{a, b\}$ on B^2 which satisfies $\varphi = 0$ on ∂B^2. This inequality was subsequently extended to arbitrary surfaces and the best constants for estimating $||\varphi||_{L^\infty}$ or $||d\varphi||_{L^2}$ were found, see [118]. The point here is that, once we have (35), we can easily deduce, by approximating by smooth maps, that **solutions to** $-\Delta\varphi = \{a, b\}$ **are continuous.** Hence the result of Wente follows.

For harmonic maps the key observation is that a u *is weakly harmonic if and only if the following conservation laws hold*

$$d\left(\star(u^i du^j - u^j du^i)\right) = 0 \qquad \forall\, i, j \ \text{ such that } \quad 1 \leq i, j \leq n+1, \tag{36}$$

where $\star$ is the Hodge operator on B^2. This was remarked and exploited for evolution problems [41, 197]. One can either check (36) directly by using (26) or derive it as a consequence of Noether's theorem, due to the invariance of the Dirichlet functional under

[23] Since the regularity problem is local, and every ball in a Riemannian surface is conformally equivalent to the Euclidean ball B^2, there is no loss of generality in working on B^2.

the action of $SO(n+1)$ on $W^{1,2}(\Omega, S^n)$ [118]. From (36) we deduce that there exist maps $b^i_j \in W^{1,2}(\Omega)$ such that $db^i_j = - \star (u^i du^j - u^j du^i)$. Then we note that $\Delta b^i_j dx \wedge dy = d(\star(db^i_j)) = d(u^i du^j - u^j du^i) = 2\{u^i, u^j\} dx \wedge dy$ so that, by a Hodge decomposition of db^i_j and by using Wente inequality, we can deduce the continuity of u. This was the approach in [113]. A more direct proof[24] is the following: since $2\langle u, du\rangle = d(|u|^2) = 0$, we can rewrite the harmonic map equation (26) as

$$-\Delta u^i = u^i |du|^2 = \left(u^i \frac{\partial u_j}{\partial x} - u_j \frac{\partial u^i}{\partial x} \right) \frac{\partial u^j}{\partial x} + \left(u^i \frac{\partial u_j}{\partial y} - u_j \frac{\partial u^i}{\partial y} \right) \frac{\partial u^j}{\partial y} = \{b^i_j, u^j\}\,, \tag{37}$$

where, as usual we sum over repeated indices, $u_i := \delta_{ij} u^j$ and we have used the relation $db^i_j = - \star (u^i du^j - u^j du^i)$. Note that an alternative way to write (37) is

$$d(\star du^i) + db^i_j \wedge du^j = 0. \tag{38}$$

We deduce that u is continuous. This method can be extended without difficulty if we replace the target S^n by any homogeneous manifold $\mathcal{N}$, since then Noether's theorem provides us with the conservation laws that we need [114].

In the case where $\mathcal{N}$ has no symmetry we need to refine the results on the quantities $\{a, b\}$. In [45] R. Coifman, P.-L. Lions, Y. Meyer and S. Semmes proved that, **if** $a, b \in W^{1,2}(\mathbb{R}^2)$ **then** $\{a, b\}$ **belongs to the generalized Hardy space** $\mathcal{H}^1(\mathbb{R}^2)$. We do not give here the various and slightly complicated definitions of the Hardy space $\mathcal{H}^1(\mathbb{R}^m)$, which was introduced by E. Stein and G. Weiss [206], but just list useful properties of it:

a) $\mathcal{H}^1(\mathbb{R}^m)$ is a strict subspace of $L^1(\mathbb{R}^m)$;
b) any function φ on $\Omega \subset \mathbb{R}^m$ such that $\Delta\varphi = f$ on Ω, where $f \in \mathcal{H}^1(\mathbb{R}^m)$, belongs to $W^{2,1}(\Omega)$, i.e. its second partial derivatives are integrable [205];
c) **let** $\alpha \in W^{1,2}(\mathbb{R}^m)$ **and** β **be a closed (in the distribution sense)** $(m-1)$**-form on** $\mathbb{R}^m$ **with coefficients in** $L^2(\mathbb{R}^m)$**; then** $d\alpha \wedge \beta = f\, dx^1 \wedge \cdots \wedge dx^m$**, where** f **belongs to** $\mathcal{H}^1(\mathbb{R}^m)$ [45]. In particular, if $a, b \in W^{1,2}(\mathbb{R}^2)$ then $\{a, b\} \in \mathcal{H}^1(\mathbb{R}^2)$;
d) by a theorem of C. Fefferman [72], $\mathcal{H}^1(\mathbb{R}^m)$ is the dual space of VMO$(\mathbb{R}^m)$ and **the dual space of** $\mathcal{H}^1(\mathbb{R}^m)$ **is BMO**$(\mathbb{R}^m)$ (see footnote 13).

Now we come back to the regularity problem. We now assume that there exists a smooth section $\widetilde{e} := (\widetilde{e}_1, \ldots, \widetilde{e}_n)$ of the bundle $\mathcal{F}$ of orthonormal tangent frames on $\mathcal{N}$. Although there are topological obstructions, there are ways to reduce to this situation, see [118]. For any map $u \in W^{1,2}(B^2, \mathcal{N})$, consider the pull-back bundle $u^*\mathcal{F}$. To any $R \in W^{1,2}(B^2, SO(n))$ we associate the section $e := \widetilde{e} \circ u \cdot R$ of $u^*\mathcal{F}$ defined by $e_a := (\widetilde{e}_b \circ u) R^b_a$, and we minimize over all gauge transformations $R \in W^{1,2}(B^2, SO(n))$ the functional $F(e) := \frac{1}{4} \int_{B^2} \sum_{1 \leq a,b \leq n} |\omega^b_a|^2 dx^1 dx^2$, where $\omega^b_a := \langle de_a, e_b\rangle$. It is easy to show that the infimum is achieved for some *harmonic section* $\underline{e}$ of $u^*\mathcal{F}$ [118, Lemma 4.1.3]. The Euler–Lagrange equation satisfied by $\underline{e}$ can be written as a system of conservation laws (again a consequence of Noether's theorem):

$$d\left(\star \underline{\omega}^b_a\right) = 0 \quad \text{on}\ \Omega \quad \text{and} \quad \underline{\omega}^b_a(\partial_n) = 0 \quad \text{on}\ \partial\Omega\,, \tag{39}$$

[24] This was pointed out by P.-L Lions.

which is satisfied by its Maurer–Cartan form $\underline{\omega}_a^b := \langle d\underline{e}_a, \underline{e}_b \rangle$. Thanks to (39), we can construct maps $A_a^b \in W^{1,2}(B^2)$ such that $dA_a^b = \star \underline{\omega}_a^b$ on B^2 and $A_a^b = 0$ on ∂B^2. Then the key observation is that

$$\Delta A_a^b = \left\langle \frac{\partial \underline{e}_a}{\partial x}, \frac{\partial \underline{e}_b}{\partial y} \right\rangle - \left\langle \frac{\partial \underline{e}_a}{\partial y}, \frac{\partial \underline{e}_b}{\partial x} \right\rangle = \sum_{i=1}^N \{ \underline{e}_a^i, \underline{e}_b^i \}, \tag{40}$$

where $(\underline{e}_a^i(x))_{1 \le i \le N}$ are the coordinates of $\underline{e}_a(x) \in T_{u(x)}\mathcal{N} \subset \mathbb{R}^N$ in a fixed orthonormal basis of $\mathbb{R}^N$. Hence the right hand side of (40) coincides locally with some function in $\mathcal{H}^1(\mathbb{R}^2)$, thanks to property c) of Hardy spaces above. Hence by property b), the second derivatives of A_a^b are locally integrable. This property implies that the components of dA_a^b are in the *Lorentz space* $L^{2,1}$, a slight improvement on L^2 [118]. But since $dA_a^b = \star \underline{\omega}_a^b$, this improvement is valid also for the connection $\underline{\omega}_a^b$.

Lastly, consider a weakly harmonic map $u \in W^{1,2}(B^2, \mathcal{N})$ and write its Euler–Lagrange equation (25) in the moving frame $\underline{e}$: if we set $\alpha^a := \langle \partial u/\partial z, \underline{e}_a \rangle$ and $\theta_b^a := \underline{\omega}_a^b(\partial/\partial \overline{z})$, we obtain $\partial \alpha^a / \partial \overline{z} = \theta_b^a \alpha^b$. In this equation, α^a is in L^2 whereas, thanks to the choice of a *Coulomb moving frame* $\underline{e}$, the function θ_b^a is in $L^{2,1}$. This slight improvement turns out to be enough to conclude that u is Lipschitz continuous.

Recently T. Rivière [184] proved the regularity of all maps $u \in W^{1,2}(B^2, \mathcal{N})$ which are critical points of the functional $F(u) := \frac{1}{2} \int_{B^2} |du|^2 dxdy + \int_{B^2} u^*\omega$, where ω is a $\mathcal{C}^1$ differential 2-form on $\mathcal{N}$ such that the coefficients of $d\omega$ are in $L^\infty(\mathcal{N})$. This answers positively conjectures of E. Heinz and S. Hildebrandt. The method provides, in particular, an alternative proof of the regularity of weakly harmonic maps with values in an arbitrary manifold without Coulomb moving frames. Instead, it relies on constructing *conservation laws*, as for maps into the sphere, but *without* symmetry. First, let us try to imitate equation (38) for a weakly harmonic map into an arbitrary compact manifold $\mathcal{N}$. We let $A \in \Gamma(S^2T^*\mathcal{N} \otimes N\mathcal{N})$ be the second fundamental form of the embedding $\mathcal{M} \subset \mathbb{R}^N$. For $y \in \mathcal{N}$ denote by $A_{jk}^i(y)$ the components of A_y in a fixed orthonormal basis $(\epsilon_1, \cdots, \epsilon_N)$ of $\mathbb{R}^N$, i.e., $A_y(X,Y) = A_{jk}^i(y) X^j Y^k \epsilon_i$ $(X, Y \in T_y\mathcal{N})$. Then we can write the Euler–Lagrange equation (25) for u as

$$d(\star du^i) - (\star A_{kj}^i(u)\, du^k) \wedge du^j = 0.$$

But since A takes values in the normal bundle, we have $\sum_{j=1}^N A_{ki}^j(u) du^j = 0$, so that we can transform the previous equation into

$$d(\star du^i) - (\star \Omega_j^i) \wedge du^j = 0 \quad \text{where} \quad \Omega_j^i := A_{kj}^i(u)\, du^k - A_{ki}^j(u)\, du^k. \tag{41}$$

If we compare with (38), which can also be written $d(\star du^i) - (\star(u^i du_j - u_j du^i)) \wedge du^j = 0$, we see that (38) is a particular case of (41), where $\Omega_j^i = u^i du_j - u_j du^i$. The difference is that we do not have $d(\star \Omega_j^i) = 0$ in general. But *both forms are skew-symmetric in* (i, j). And that property is actually enough. The idea is to substitute for $\star du^i$ another quantity, of the form $A_j^i(\star du^i)$, where $A_j^i \in W^{1,2}(B^2)$. A computation using (41) shows that

$$d\left(A_j^i(\star du^j)\right) = -\star\left(dA_j^i - A_k^i \Omega_j^k\right) \wedge du^j.$$

Hence if we assume that we can find maps $A := (A_j^i)_{1 \le i,j \le N}$, $B := (B_j^i)_{1 \le i,j \le N} \in W^{1,2}(B^2, M(N, \mathbb{R}))$ such that A is invertible with a bounded inverse and

$$\star\left(dA_j^i - A_k^i \Omega_j^k\right) = dB_j^i, \tag{42}$$

then we obtain an equation similar to (38), i.e.,

$$d\left(A^i_j(\star du^j)\right) + dB^i_j \wedge du^j = d\left(A^i_j(\star du^j) + B^i_j du^j\right) = 0. \tag{43}$$

Then formulation (43) allows us to prove the continuity of u easily: we use the Hodge decomposition: $A^i_j du^j = dD^i_j - \star dE^i_j$ for some functions $D^i_j, E^i_j \in W^{1,2}(B^2)$, then we deduce $d(\star dE^i_j) = -dA^i_j \wedge du^j$, i.e. $-\Delta E^i_j = \{A^i_j, u^j\}$ from the definition of E^i_j and so we obtain $d(\star dD^i_j) = -dB^i_j \wedge du^j$, i.e. $-\Delta D^i_j = \{B^i_j, u^j\}$ from (43). Hence, from properties b) and c) of Hardy spaces, we deduce that the first derivatives of D^i_j and E^i_j are in the Lorentz space $L^{2,1}$; since A has a bounded inverse, it follows that the first derivatives of u are also in $L^{2,1}$. Thus u is continuous. To complete the proof one needs to prove the existence of A and B solving (42). For that purpose Rivière adapts a result of K. Uhlenbeck [217] to first prove the existence of some gauge transformation map $P \in W^{1,2}(B^2, SO(N))$ such that $\Omega^P := P^{-1}dP + P^{-1}\Omega P$ satisfies the Coulomb gauge condition $d(\star\Omega^P) = 0$. This implies, in particular, that $P^{-1}dP + P^{-1}\Omega P = \star d\xi$, for some map $\xi \in W^{1,2}(B^2, so(N))$. Then by putting $A := \widetilde{A}P^{-1}$, equation (42) reduces to $d\widetilde{A} - \widetilde{A}(\star d\xi) + (\star dB)P = 0$, a linear elliptic system in $\widetilde{A}$ and B, which can be solved by a fixed point argument.

4.3 Regularity results in dimension greater than two

Preliminary facts

If $m := \dim \mathcal{M} \geq 3$, weakly harmonic maps in $W^{1,2}(\mathcal{M}, \mathcal{N})$ will *not* be regular in general and may even be completely discontinuous as shown by the result of T. Rivière (see §5.4), unless $\mathcal{N}$ has some convexity properties (see §6.3). But *partial* regularity results hold for minimizing or stationary maps. Indeed we are able, in general, to prove that such maps are smooth outside a closed subset Σ that we will call the *singular set*. The size of Σ is estimated in terms of some Hausdorff dimension and corresponding Hausdorff measure. Fix some $s \in [0, m]$. For any covering of Σ by a countable union of balls $(B^m_j)_{j\in J}$ of radius r_j, consider the quantity $\mathcal{H}^s((B^m_j)_{j\in J}, \Sigma) := \alpha(s)\sum_{j\in J} r^s_j$, where $\alpha(s) = 2\pi^{\frac{s}{2}}/s\Gamma(\frac{s}{2})$: this measures *approximately* the s-dimensional volume of Σ. The **s-dimensional Hausdorff measure of** Σ is:

$$\mathcal{H}^s(\Sigma) := \sup_{\delta>0} \inf_{r_j<\delta} \mathcal{H}^s\left((B^m_j)_{j\in J}, \Sigma\right)$$

(in the infimum, $(B^m_j)_{j\in J}$ is such that $\Sigma \subset \cup_{j\in J} B^m_j$).

Then there exists some $d \in [0, m]$ such that $\forall s \in [0, d)$, $\mathcal{H}^s(\Sigma) = 0$ and $\forall s \in (d, m]$, $\mathcal{H}^s(\Sigma) = +\infty$. If $\mathcal{H}^d(\Sigma)$ is finite, d is called the **Hausdorff dimension of** Σ. In the special case when Σ is a smooth submanifold of dimension k, then $d = k$ and $\mathcal{H}^d(\Sigma)$ coincides with the d-dimensional volume of Σ.

Furthermore it is useful to analyze the first consequences of the Euler–Lagrange equation (25) and the conservation law for the stress-energy tensor (29) concerning the regularity of a weak solution $u \in W^{1,2}(\mathcal{M}, \mathcal{N})$. Equation (25) implies that the components of $\Delta_g u$ are in $L^1(\mathcal{M})$, from which one can deduce that the first derivatives of u are locally in L^p for $1 \leq p < m/(m-1)$, which has no interest. However, the conservation law (29) immediately provides the following strong improvement to the regularity of u.

The monotonicity formula. Given a map $u \in W^{1,2}(\mathcal{M}, \mathcal{N})$; to each $x \in \mathcal{M}$ and $r > 0$ such that the geodesic ball $B(x,r)$ is contained in $\mathcal{M}$, we associate the quantity

$$E_{x,r}(u) := \frac{1}{r^{m-2}} \int_{B(x,r)} |du|_g^2 \omega_g.$$

Now *let $B(a,r) \subset \mathcal{M}$ be a geodesic ball centred at a and of radius $r > 0$ such that the distance from a to its cut locus and to $\partial\mathcal{M}$ is greated than r. Then there exist constants C (depending on m) and Λ (depending on a bound of the curvature on $B(a,r)$) such that, if $u \in W^{1,2}(\mathcal{M}, \mathcal{N})$* ***satisfies the relation*** *(29), then for all $x \in B(a, r/2)$ the function $(0, r/2] \ni \rho \longmapsto e^{C\Lambda\rho} E_{x,\rho}(u)$ is non-decreasing* [231]. If the metric on $\mathcal{M}$ is *flat* this holds with $\Lambda = 0$ and this can be proved by integrating over $B^m(x_0, r)$ the relation $(\partial/\partial x^\alpha)\big((x^\beta - x_0^\beta) S_\beta^\alpha(u)\big) = S_\alpha^\alpha(u) = \frac{1}{2}(m-2)|du|_g^2$, a consequence of (29). We then get an identity from which we derive:

$$\text{for } 0 < r_1 < r_2, \quad E_{x,r_2}(u) - E_{x,r_1}(u) = \frac{2}{r^{m-2}} \int_{B^m(x,r_2) \setminus B^m(x,r_1)} \left| \frac{\partial u}{\partial n} \right|^2 d^m x \geq 0, \tag{44}$$

where $\partial u / \partial n$ denotes the normal derivative of u. The monotonicity formula has strong consequences; for simplicity, we expound these in the case where $(\mathcal{M}, g)$ is flat[25]. First, elementary geometric reasoning shows that, for $\gamma \in (0,1)$, $E_{x_0,r}(u)$ controls $E_{x,\gamma r}(u)$ for $x \in B^m(x_0, (1-\gamma)r)$ and hence, by (44), $E_{x_0,r}(u)$ controls $E_{x,\rho}(u)$ for $x \in B^m(x_0, (1-\gamma)r)$ and $\rho \leq \gamma r$. Then, by a Poincaré–Sobolev inequality:

$$\frac{1}{\rho^m} \int_{B^m(r,\rho)} |u - u_{x,\rho}|^2 \, dx \leq C \, E_{x,\rho}(u), \qquad \text{with } u_{x,\rho} := \frac{1}{|B^m(x,\rho)|} \int_{B^m(x,\rho)} u \, dx,$$

we deduce a bound on $\sup\big\{\rho^{-m} \int_{B^m(r,\rho)} |u - u_{x,\rho}|^2 \, dx \mid x \in B^m(x_0, (1-\gamma)r), \rho \leq \gamma r\big\}$, i.e., roughly speaking, on the local BMO-norm of u on $B^m(x_0, (1-\gamma)r)$. The BMO space (see footnote 13) contains all the spaces L^p, for $1 \leq p < \infty$, and hence is very close to L^∞. Thus this is an important gain of regularity.

The ε-regularity. Our task is to put together consequences of (25) and (29) in order to improve the preceding observations. The *(continuous) main step* in most regularity results consists *of showing that there exists some $\varepsilon_0 > 0$ such that for any weak solution u (for a suitable notion of 'weak'), if $E_{a,r}(u) < \varepsilon_0$, then, for $0 < \sigma < \rho$ such that ρ/r is sufficiently small and for $x \in \mathcal{M}$ close to a,*

$$E_{x,\sigma}(u) \leq C \left(\frac{\sigma}{\rho} \right)^\alpha E_{x,\rho}(u) \tag{45}$$

for some constants $C > 0$ and $\alpha > 0$. If this is true, we are in a position to apply the *Dirichlet growth theorem of Morrey* (see [157, 83]), which implies that u is Hölder continuous with exponent $\alpha/2$ in a neighbourhood of a. This method is the reason for the partial regularity: a covering argument shows that, if $\Sigma := \{a \in \mathcal{M} | \ \lim_{r \to 0} \inf E_{a,r}(u) \geq \varepsilon_0\}$

[25] Since in the regularity theory we are interested in the local properties of weak solutions, the effect of the curvature of $\mathcal{M}$ can be neglected.

had a non-vanishing $(m-2)$-dimensional measure, u would have infinite energy, hence $\mathcal{H}^{m-2}(\Sigma)=0$ by contradiction. The continuous main step itself can be achieved by proving a *discrete* version of it: *there exists some* $\varepsilon_0>0$ *and some* $\tau\in(0,1)$ *such that, for any weak solution* u *(here again we stay vague), if* $E_{x,r}(u)<\varepsilon_0$*, then*

$$E_{x,\tau r}(u)\leq\frac{1}{2}E_{x,r}(u). \tag{46}$$

Indeed, by using this result at several scales and concatenating them, one easily deduces (45).

A first attempt. We now describe in a naive way an attempt to prove the *discrete main step* (46). First, we observe that, if u is defined on $B^m(a,r)$, then the map $T_{a,r}u$ defined by $T_{a,r}u(x):=u(rx+a)$ is defined on $B^m:=B^m(0,1)$ and, furthermore, $E_{0,1}(T_{a,r}u)=E_{a,r}(u)$, which shows that one can work without loss of generality with a map $u\in W^{1,2}(B^m,\mathcal{N})$. So our aim is to prove that $E_{0,\tau}(u)\leq\frac{1}{2}E_{0,1}(u)$ for some $\tau>0$ under some smallness assumption on $E_{0,1}(u)$. We split $u=v+w$, where v agrees with u on ∂B^m and is harmonic with values in $\mathbb{R}^N\supset\mathcal{N}$, and w vanishes on ∂B^m and has $\Delta w=\Delta u=-A(u)(du,du)$. Then, for $\tau\in(0,1)$,

$$\begin{aligned}E_{0,\tau}(u)&=\frac{1}{\tau^{m-2}}\int_{B^m(0,\tau)}|du|^2d^mx\\&\leq\frac{2}{\tau^{m-2}}\int_{B^m(0,\tau)}|dv|^2d^mx+\frac{2}{\tau^{m-2}}\int_{B^m(0,\tau)}|dw|^2d^mx.\end{aligned}$$

We now estimate separately each term on the right hand side. On the one hand, since v is harmonic, $|dv|^2$ is a subharmonic function (see Section 1) and hence

$$\frac{2}{\tau^{m-2}}\int_{B^m(0,\tau)}|dv|^2d^mx\leq\frac{2}{\tau^{m-2}}\tau^m\int_{B^m(0,1)}|dv|^2d^mx\leq 2\tau^2E_{0,1}(u). \tag{47}$$

On the other hand, we have

$$\begin{aligned}\int_{B^m(0,\tau)}|dw|^2d^mx&\leq\int_{B^m(0,1)}|dw|^2d^mx\\&=\int_{\partial B^m(0,1)}\left\langle w,\frac{\partial w}{\partial n}\right\rangle d^mx-\int_{B^m(0,1)}\langle w,\Delta w\rangle\,d^mx,\end{aligned}$$

which implies, since $w=0$ on ∂B^m,

$$\frac{2}{\tau^{m-2}}\int_{B^m(0,\tau)}|dw|^2d^mx\leq\frac{2}{\tau^{m-2}}\int_{B^m(0,1)}\big\langle u-v,A(u)(du,du)\big\rangle\,d^mx. \tag{48}$$

We see from (47) that, by choosing τ sufficiently small, the contribution of v in $E_{0,\tau}(u)$ can be as small as we want in comparison to $E_{0,1}(u)$. Hence the difficulty in proving (46) lies in estimating the right-hand side of (48). We may write $\int_{B^m(0,1)}\langle u-v,A(u)(du,du)\rangle d^mx\leq C\sup_{B^m(0,1)}|u-v|\int_{B^m(0,1)}|du|^2d^mx=C\sup_{B^m(0,1)}|u-v|E_{0,1}(u)$ and, by using the maximum principle for v we can estimate $\sup_{B^m(0,1)}|u-v|$ in terms of a bound $\mathrm{osc}_{B^m(0,1)}u:=\sup_{x,y\in B^m(0,1)}|u(x)-u(y)|$ on the oscillation of u on $B^m(0,1)$. However, *we have no estimate on these oscillations* but only on the *mean oscillation*, hence our attempt *failed*. Anyway, we see that we are in a borderline situation since, again, an estimate in BMO space is close to an L^∞ estimate. The following partial regularity results can be obtained by filling this gap between BMO and L^∞.

Regularity of *minimizing* maps in dimension greater than two

For minimizing maps, partial regularity results were obtained by R. Schoen and K. Uhlenbeck [190] (and also by M. Giaquinta and E. Giusti [84, 85] under the assumption that the image is contained a a single coordinate chart): **let** $u \in W^{1,2}(\mathcal{M}, \mathcal{N})$ **be a minimizing weakly harmonic map, then there exists a closed singular set** $\Sigma \subset \mathcal{M}$ **such that** u **is Hölder continuous on** $\mathcal{M} \setminus \Sigma$ **and** $\mathcal{H}^{m-3}(\Sigma) < \infty$**.** This is proved in two steps:

(i) first, one shows that a minimizing map u is smooth outside a singular set Σ such that $\mathcal{H}^{m-2}(\Sigma) = 0$;

(ii) then, one shows that, near a point $x_0 \in \Sigma$, the minimizing map u behaves asymptotically like a homogeneous map, so that, in particular, the singular set looks asymptotically like a cone centred at x_0. This forces a reduction of the dimension of Σ.

Step (i) [190, 84] relies on the ideas expounded in the previous paragraph, since a minimizing map is automatically stationary. A key observation is that, if we have a local BMO bound on a stationary map u, then we can *approximate* u locally by a smooth map $u^{(h)}$ (where $h > 0$ is small) with values in $\mathbb{R}^N \supset \mathcal{N}$, and the estimate on the *mean* oscillation of u becomes an estimate on the oscillations of $u^{(h)}$. Thus the previous attempt works if we replace u by $u^{(h)}$ (with suitable adaptations), leading to an estimate of $E_{0,\tau}(u^{(h)})$ in terms of $E_{0,1}(u)$. Since, again, u has small local mean oscillation, we can compose $u^{(h)}$ with a projection onto $\mathcal{N}$ to get a smooth map u_h *with values in* $\mathcal{N}$ which approximates u, and then deduce an estimate for $E_{0,\tau}(u_h)$ in terms of $E_{0,1}(u)$. But we are interested in estimating $E_{0,\tau}(u)$, and here we use the fact that u is a *minimizer*: by a delicate gluing process we construct a test function U_h which agrees with u on $\partial B^m(0, 2\tau)$ and coincides with u_h in $B^m(0,\tau)$, and we obtain (46) by comparing the energy of u and the energy of U_h on suitable balls.

Step (ii) [190, 85] is inspired by a similar work by H. Federer [71]. It is based on the analysis of a *blow-up sequence* $(u_k)_{k\in\mathbb{N}}$ of minimizing maps centred at a point a in the singular set Σ. Each $u_k \in W^{1,2}(B^m, \mathcal{N})$ is defined by $u_k(x) := u(a + r_k x)$, for some decreasing sequence $(r_k)_{k\in\mathbb{N}}$ which converges to 0. It is not difficult to prove that, after extraction of a subsequence if necessary, $(u_k)_{k\in\mathbb{N}}$ converges weakly in $W^{1,2}$ to a map $\underline{u}_a \in W^{1,2}(B^3, \mathcal{N})$, called the *tangent map at* a. However, one can prove that, actually, $(u_k)_{k\in\mathbb{N}}$ converges *strongly* in $W^{1,2}$ to $\underline{u}_a$ and that $\underline{u}_a$ is weakly harmonic[26]. Hence we can pass to the limit in (44) and deduce that $\partial\underline{u}_a/\partial n = 0$, i.e., $\underline{u}_a$ is homogeneous.

Remarks (i) A variant of the proof of step (i) has been proposed by S. Luckhaus [151], with applications to a much larger class of functionals on maps with values in manifolds. Also, in the special case $\mathcal{N} = S^2$, simpler proofs are available: by R. Hardt, D. Kinderlehrer and F. H. Lin [105], and by Y. Chen and Lin [42].

(ii) In step (ii) it is not clear a priori whether the tangent map $\underline{u}_a$ at a singularity a depends on the choice of the blow-up sequence $(u_k)_{k\in\mathbb{N}}$. It is actually a deep and difficult question. L. Simon [198] (see also [199] for simplifications) proved that *if* $\mathcal{N}$ *is real analytic, for any map* $u \in W^{1,2}(\mathcal{M}, \mathcal{N})$ *which is a minimizer and is singular at* $a \in \mathcal{M}$*, if the tangent map* $\underline{u}_a$ *is smooth outside 0, then this tangent map is unique*. In constrast, B. White [228] constructed a harmonic map into a smooth *non-analytic* Riemannian manifold with a one-parameter *family of tangent maps* having an isolated singularity at the same

[26] $\underline{u}_a$ is actually minimizing, as shown by S. Luckhaus [152].

point, hence proving that the analyticity assumption in the result of Simon is crucial. See the survey by Hardt [103] for a discussion of these questions.

Reduction of the singular set. These results can be improved if we assume some further conditions on $\mathcal{N}$: for instance, if $\mathcal{N}$ is non-negatively curved or if the image of a minimizing map is contained in a geodesically convex ball, then minimizing maps are smooth (see §6.3). Optimal examples of such convex targets are the compact subsets of $S^n_+ := \{y \in \mathbb{R}^{n+1} | \; y^{n+1} > 0\}$. These examples are close to the borderline case where the target is $\overline{S^n_+} := \{y \in \mathbb{R}^{n+1} | \; y^{n+1} \geq 0\}$, since minimizing maps into $\overline{S^n_+}$ may not be smooth (see §6.2). In order to estimate the size of the critical set outside these situations, one possible approach is to try to classify the *minimizing tangent maps* $u \in W^{1,2}(B^m, \mathcal{N})$, i.e. maps of the form $u(x) = \psi(x/|x|)$, where $\psi : S^{m-1} \longrightarrow \mathcal{N}$. This relies on proving kinds of *Bernstein theorems* for *minimizing tangent maps* into $\mathcal{N}$. These questions have been investigated by R. Schoen and K. Uhlenbeck [192] and M. Giaquinta and J. Souček [89] in two cases:

(i) in the limit case, where $\mathcal{N} = \overline{S^n_+}$: **a minimizing map** $u \in W^{1,2}(\mathcal{M}, \overline{S^n_+})$ **is smooth if** $n \leq 6$ **and has a closed singular set of Hausdorff dimension less or equal to** $n - 7$ **for** $n \geq 7$ [192, 89]. This is based on results in [120, 126] (see also §6.3).

(ii) beyond the limit case, if $\mathcal{N} = S^n$: **a minimizing map** $u \in W^{1,2}(\mathcal{M}, S^n)$ **is smooth if** $m :=$ **dim** $\mathcal{M} \leq \overline{m}(n)$**, where** $\overline{m}(n)$ **is given by the following table** [192]:

n	2	3	4	5	6	7	8	9	$[10, \infty)$
$\overline{m}(n)$	2	3	3	3	4	4	5	5	6

See also §6.2. Lastly, *extra results on reduction of the singular set* were proved for *stationary* maps by F. H. Lin and, in particular, are valid for minimizing maps, see below.

The structure of the singular set. The singular set Σ has a simple structure in dimension 3, since then it is composed of isolated point. However, in higher dimensions, Σ has a positive Hausdorff dimension in general and the analysis of its regularity requires the use of techniques from geometric measure theory. For maps u in $W^{1,2}(B^4, S^2)$ R. Hardt and F. H. Lin [108] proved that *the singular set* Σ *of a minimizer in* $W^{1,2}(B^4, S^2)$ *with a smooth trace on* ∂B^4 *is the union of a finite set and of finitely many Hölder continous closed curves with only finitely many crossings.* For more general situations L. Simon [200] proved that *if* $\mathcal{N}$ *is compact and real analytic, for any minimizer* $u \in W^{1,2}(\mathcal{M}, \mathcal{N})$ *with singular set* Σ *and any ball* $\mathcal{B} \subset \mathcal{M}$, $\Sigma \cap \mathcal{B}$ *is the union of a finite pairwise disjoint collection of locally* $(m-3)$*-rectifiable locally compact sets.*[27] See [103] for a survey; see also the book of Simon [201].

Minimizing maps from the unit ball B^3 **to** S^2. H. Brezis, J.-M. Coron et E. H. Lieb [24] found further results in the special case $\mathcal{M} = B^3 \subset \mathbb{R}^3$ and $\mathcal{N} = S^2$. They prove that a *minimizing harmonic map can only have singularities of degree* ± 1; more precisely, **the only homogeneous minimizing maps** $B^3 \ni x \longmapsto \psi(x/|x|) \in S^2$ **are of the form** $\psi(x/|x|) = \pm Rx/|x|$**, where** $R \in SO(3)$ **is a rotation** (similar results holds for $\mathcal{N} = \mathbb{R}P^2$). The minimality of the radial projection $u_\odot(x) = x/|x|$ is obtained by establishing the lower bound $E_{B^3}(u) \geq E_{N^3}(u_\odot) = 4\pi$ for any minimizing map

[27]More can be said when all *Jacobi fields* along (i.e., infinitesimal deformations of) the harmonic maps are *integrable*, i.e., come from genuine deformations through harmonic maps, see [200, 201, 145].

$u \in W^{1,2}_{u_\odot}(B^3, S^2)$, by using the following idea. By the partial regularity result [190] any such map u is smooth outside a finite singular set $\{a_1, \cdots, a_p\}$ with respective degrees $\{d_1, \cdots, d_p\}$. Then, from the local inequality $\frac{1}{2}|du|^2 \geq |u^*\omega_{S^2}|$, which holds a.e., one deduces that

$$E_{B^3}(u) \geq \int_{B^3} |u^*\omega_{S^2}|\, dx^1dx^2dx^3 \geq \int_{B^3} d\zeta \wedge (u^*\omega_{S^2}) = \int_{\partial B^3} \zeta\, u^*\omega_{S^2} - \int_{B^3} \zeta\, d(u^*\omega_{S^2})$$

for all $\zeta \in Lip(\Omega)$ such that $|\nabla\zeta|_{L^\infty} \leq 1$. But the condition: $u = u_\odot$ on ∂B^3 implies that $\int_{\partial B^3} \zeta u^*\omega_{S^2} = \int_{\partial B^3} \zeta\omega_{S^2}$. Furthermore, by using $d(u^*\omega_{S^2}) = \sum_{i=1}^p d_i\delta_{a_i}$ (see also §5.4 and (53)), we finally get

$$E_{B^3}(u) \geq \sup_{\zeta \in Lip(\Omega), |\nabla\zeta|_{L^\infty} \leq 1} \Big(\int_{\partial B^3} \zeta\, \omega_{S^2} - \sum_{i=1}^p d_i\, \zeta(a_i) \Big).$$

Then the proof can be reduced to an optimization problem on the set of configurations of the type $\{(a_1, d_1), \cdots, (a_p, d_p)\}$, which can be solved by adapting a theorem of Birkhoff.

Still for the case of minimizing harmonic maps u from B^3 to S^2, F. Almgren and E. H. Lieb [4] found a bound on the number $N(u)$ of singularities of u: $N(u)$ is certainly not bounded in terms of its energy $E_{B^3}(u)$, but it is in terms of the energy of its trace on ∂B^3. Indeed, **there exists a universal constant** $C > 0$ **such that, for any** $\varphi \in W^{1,2}(\partial B^3, S^2)$,

for any $u \in W^{1,2}_\varphi(B^3, S^2)$ **which is a minimizer of** E_{B^3} **, we have** $N(u) \leq CE_{\partial B^3}(\varphi)$.

The precise value of C is not known but examples constructed in [4] show that we must have $C \geq 1/(4\pi)$. It is also shown that a similar result where the energy $E_{\partial B^3}(\varphi)$ is replaced by the area covered by φ *cannot hold*.

Minimizers of the relaxed energy. The regularity of the minimizers in $W^{1,2}(B^3, S^2)$ of the functional $E^\lambda_{B^3} = E_{B^3} + 4\lambda\pi L$ (see §5.4) has been investigated by H. Brezis and F. Bethuel [14] who proved that, **if** $\lambda \in [0, 1)$**, any minimizer of** $E^\lambda_{B^3}$ **is smooth on** $B^3 \setminus \Sigma$**, where** $\mathcal{H}^0(\Sigma) < \infty$**, i.e.** Σ **is a finite collection of points**. The case $\lambda = 1$ corresponds to the *relaxed energy* $E^{rel}_{B^3} = E_{B^3} + 4\pi L$, which is harder to deal with: the only partial regularity result that we know is due to M. Giaquinta, G. Modica and J. Souček [87, 88] who showed that *minimizers of* $E^{rel}_{B^3}$ *are smooth on* $B^3 \setminus \Sigma$*, where* $\mathcal{H}^1(\Sigma) < \infty$. It is a paradox that the regularity theory for minimizers of the relaxed energy, which was designed for producing continuous harmonic maps, is less understood than the theory of minimizers of the standard energy functional.

Minimizers of the p**-energy**. The previous results have been extended to minimizers of the p-energy in various cases by S. Luckhaus [151], R. Hardt and F. H. Lin [107], M. Fuchs [75, 76], and by F. H. Lin in the important paper [148].

Regularity of *stationary* maps in dimension greater than two

For stationary maps, we have the following partial regularity result: **let** $u \in W^{1,2}(\mathcal{M}, \mathcal{N})$ **be a stationary map; then there exists a closed singular set** $\Sigma \subset \mathcal{M}$ **such that** u **is Hölder continuous on** $\mathcal{M} \setminus \Sigma$ **and** $\mathcal{H}^{m-2}(\Sigma) = 0$**.** This was proved by L. C. Evans [70] in the case where $\mathcal{N} = S^n$ and by F. Bethuel [13] in the general case.

The proof of Evans [70, 118, 88] is based on the discovery that the attempt expounded above really works for maps into a sphere S^n. Recall that the difficulty was to estimate a quantity of the type $\int_B \langle u-v, A(u)(du,du)\rangle\, d^mx$ and that only the *mean* oscillation of $u-v$ can be estimated in terms of $E_{0,1}(u)$. However we can use the same observations as in dimension two, i.e. write the harmonic map equation in the form $d(\star du^i)+du^j\wedge\star(u^idu_j-u_jdu^i)=0$, and use the conservation law $d(\star(u^idu_j-u_jdu^i))=0$. This implies, by using the property c) of Hardy spaces, that $A^i(u)(du,du)d^mx=u^i|du|^2d^mx=du^j\wedge\star(u^idu_j-u_jdu^i)$ coincides locally with a function in the Hardy space $\mathcal{H}^1(\mathbb{R}^m)$. Thus, by property d) of Hardy spaces, we can estimate $\int_B\big\langle u-v, A(u)(du,du)\big\rangle d^mx$ as a *duality product between the (local) BMO norm of* $u-v$ *and the (local) Hardy norm of* $A(u)(du,du)$, and hence complete the proof.

The proof of F. Bethuel [13, 118] uses a *Coulomb moving frame* $(\underline{e}_1,\cdots,\underline{e}_n)$ as in [116]. The strategy is somewhat parallel to the proof of Evans, but the realization is much more delicate. The idea for estimating $|du|$ on a small ball consists of using a Hodge decomposition $\langle d\left(\zeta(u-u_{0,1})\right),\underline{e}_a\rangle = dw^a+\star dv^a$, where $u_{0,1}:=|B^m(a,r)|^{-1}\int_{B^m(a,r)}u$ and $\zeta\in\mathcal{C}^\infty_c(B^m(a,r))$ is a cut-off function. Then both terms in the decomposition are estimated separately. However, because the system is not as simple as in the case treated by Evans, we need to replace Morrey's rescaled energy $E_{a,r}(u)$ by $M_{a,r}(u) := \sup\big\{\rho^{1-m}\int_{B^m(x,\rho)}|du| \mid B^m(x,\rho)\subset B^m(a,r)\big\}$ (which also controls the local bounded mean oscillation of u).

Remarks (i) Several variants of the proof by Evans exist: one can avoid the use of the Fefferman–Stein theorem on the duality between $\mathcal{H}^1$ and BMO, as done by S. Chanillo [40], or even avoid the use of the Hardy space, as done by S.-Y. A. Chang, L. Wang and P. C. Yang [39].

(ii) Using the conservation laws discovered by T. Rivière in [184], Rivière and M. Struwe [185] derived a simplified proof of the result of Bethuel, without using Coulomb moving frames.

Reduction of the singular set. The question of whether $\mathcal{H}^{m-3}(\Sigma)$ is finite is still open. The reason is that the blow-up technique used by Schoen and Uhlenbeck does not work here, since we are not able to prove that, after extracting a subsequence if necessary, a blow-up sequence $u_k(x)=u(a+r_kx)$ at a point a converges *strongly* when $r_k\to 0$. Indeed we can only prove that it converges *weakly*. This leads to the more general question of understanding *a sequence* $(v_k)_{k\in\mathbb{N}}$ *of stationary maps which converges* weakly *to some limit* v: after extracting a subsequence is necessary, we can assume that the energy density $|dv_k|^2d^mx$ converges weakly in the sense of Radon measures to a non-negative Radon measure μ which can be decomposed as $\mu=|dv|^2+\nu$; the measure ν detects the defect of strong convergence, i.e. *the sequence converges strongly if and only if* $\nu=0$. By a careful analysis of such sequences, F. H. Lin [147] proved that *the singular support*[28] Γ *of* μ *is a rectifiable subset with a finite* $(m-2)$*-dimensional Hausdorff measure*. Moreover ν is supported by Γ and, more precisely, is equal to the $(m-2)$-dimensional measure supported by Γ times an $\mathcal{H}^{m-2}$-measurable density $\Theta(x)$. This result is optimal as shown by the following example: assume that there exists a non-trivial harmonic map $\phi: S^2\longrightarrow\mathcal{N}$ and, for any $\lambda\in\mathbb{R}$, let $u_\lambda\in\mathcal{C}^\infty(B^2\times B^{m-2},\mathcal{N})$ be defined by $u_\lambda(x,y)=\phi\circ P^{-1}(\lambda x)$, where $P: S^2\longrightarrow\mathbb{R}^2$ is the stereographic projection (30). Then each u_λ is stationary

[28] The singular set Γ also coincides with $\cap_{r>0}\{x\in B^m|\ \liminf_{k\to\infty}E_{x,r}(u_k)\geq\varepsilon_0\}$.

and $|du_\lambda|^2\, d^m x$ converges weakly to a Radon measure ν supported by $\{0\} \times B^{m-2}$ when $\lambda \to +\infty$. Moreover Lin and Rivière [149] proved that, *in the case where* $\mathcal{N} = S^n$*, for a.e. point* $x \in \Gamma$ *(in the sense of* $(m-2)$*-dimensional measure) the density* $\Theta(x)$ *is a finite sum of energies of harmonic maps from* S^2 *to* S^m (this result generalizes the identity (49) for maps of surfaces) and, in particular, if $\mathcal{N} = S^2$, $\Theta(x)$ is a integer multiple of 8π. For a general target manifold, a further result by Lin [147] is that, for a given $\mathcal{N}$, *any sequence of weakly converging stationary maps converges strongly (i.e. satisfies* $\nu = 0$*) if and only if there is no smooth non-constant harmonic map from* S^2 *to* $\mathcal{N}$. Applying his results to a blow-up sequence of stationary maps, Lin [147] proved that **if $\mathcal{N}$ does not carry any harmonic S^2, then the singular set Σ of a stationary map with values in $\mathcal{N}$ has Hausdorff dimension $s \leq m-4$. If, furthermore, $\mathcal{N}$ is real analytic, then Σ is s-rectifiable.** On the other hand a consequence of the work by Lin and Rivière [149] is that, *for a stationary map* u *into* S^2*, if* $\liminf_{k\to\infty} E_{x,r}(u) < 8\pi$*, then* u *is continuous at* x.

Stationary critical points of the p-energy. A notion similar to the notion of stationary maps for critical points of the p-energy makes sense, and the previous regularity results has been extended to this case by L. Mou and P. Yang [158].

5 Existence methods

5.1 Existence of harmonic maps by the direct method

The general strategy for proving existence of harmonic maps consists of choosing a non-empty class $\mathcal{E} \subset W^{1,p}(\mathcal{M},\mathcal{N})$ of maps which is defined, for example, by some Dirichlet boundary conditions or some topological constraints, and then to consider a sequence $(u_k)_{k\in\mathbb{N}}$ minimizing the energy $E_{\mathcal{M}}$ in $\mathcal{E}$. Here we assume for simplicity that $\mathcal{M}$ is compact. One can repeat the arguments given in Section 1 for the solution to the classical Dirichlet problem: since $\mathcal{E}$ is non-empty it contains maps of finite energy and so, in particular, the minimizing sequence has bounded energy. Thus, there is a subsequence $(\varphi(k))_{k\in\mathbb{N}} \subset (k)_{k\in\mathbb{N}}$ such that $\left(u_{\varphi(k)}\right)_{k\in\mathbb{N}}$ converges *weakly* in $W^{1,p}(\mathcal{M},\mathbb{R}^N)$ to some map $\underline{u} \in W^{1,p}(\mathcal{M},\mathbb{R}^N)$. An extra task is to check that $\underline{u}(x) \in \mathcal{N}$ a.e. This is a consequence of the fact that, because of the Rellich–Kondrakov theorem, the subsequence $(\varphi(k))_{k\in\mathbb{N}}$ converges *strongly* to $\underline{u}$ in $L^p(\mathcal{M},\mathbb{R}^N)$ for all $p < 2m/(m-2)$. Hence we can extract a further subsequence $(\varphi_1(k))_{k\in\mathbb{N}} \subset (\varphi(k))_{k\in\mathbb{N}}$ such that $\left(u_{\varphi_1(k)}\right)_{k\in\mathbb{N}}$ converges a.e. on $\mathcal{M}$ to $\underline{u}$, by a standard result of Lebesgue theory. This implies $\underline{u}(x) \in \mathcal{N}$ a.e. on $\mathcal{M}$. Hence $\underline{u} \in W^{1,2}(\mathcal{M},\mathcal{N})$. Then two cases can occur:

(i) **$\mathcal{E}$ is closed with respect to the weak topology of** $W^{1,2}(\mathcal{M},\mathcal{N})$. Then we know that $\underline{u} \in \mathcal{E}$ and, using the fact that $E_{\mathcal{M}}$ is lower semi-continuous for the weak $W^{1,2}$-topology as in the classical case (see §1), we prove that $\underline{u}$ is actually an energy minimizing map in $\mathcal{E}$, and so is weakly harmonic. In the special case when $\mathcal{M}$ is two-dimensional, the classical regularity result of C. B. Morrey [156] ensures that $\underline{u}$ is smooth. In higher dimensions, the minimizers are only partially regular, as shown by the regularity theory of R. Schoen and K. Uhlenbeck [190] (see §4.3).

(ii) **$\mathcal{E}$ is not closed with respect to the weak topology of** $W^{1,2}(\mathcal{M},\mathcal{N})$. Then no general argument guarantees that $\underline{u} \in \mathcal{E}$ or that $\underline{u}$ is an energy minimizer.

5.2 The direct method in a class of maps closed for the weak topology

The class $\mathcal{E}$ is closed with respect to the weak topology of $W^{1,2}(\mathcal{M},\mathcal{N})$ in the following situations:

1. $\mathcal{E}$ **is defined through Dirichlet boundary conditions**, because the trace operator given by tr : $W^{1,2}(\mathcal{M},\mathbb{R}^N) \longrightarrow W^{\frac{1}{2},2}(\partial\mathcal{M},\mathbb{R}^N)$ is continuous for the weak topologies[29]. The first application was the solution of the Plateau problem for a surface in a Riemannian manifold by C. B. Morrey [156].

2. $\mathcal{E}$ **is defined by prescribing the action of maps in** $W^{1,2}(\mathcal{M},\mathcal{N})$ **on** $\pi_1(\mathcal{M})$ (see also §3.3). The first application was the following result by L. Lemaire [144]: **let** $\mathcal{M}$ **and** $\mathcal{N}$ **be two Riemannian manifolds of dimension 2, with** $\partial\mathcal{N} = \emptyset$**, and assume that** genus $\mathcal{M} \geq 1$ **and** genus $\mathcal{N} \geq 1$**. Then any homotopy class of maps between** $\mathcal{M}$ **and** $\mathcal{N}$ **contains a minimizing harmonic representative**. In the proof of this result, the fundamental groups $\pi_1(\mathcal{M})$ and $\pi_1(\mathcal{N})$ are seen as the automorphisms groups of the universal covers $\widetilde{\mathcal{M}}$ and $\widetilde{\mathcal{N}}$ of $\mathcal{M}$ and $\mathcal{N}$, respectively. Then, to any homotopy class represented by a map $\varphi : \mathcal{M} \longrightarrow \mathcal{N}$, we associate the class of equivariant maps $\widetilde{u} : \widetilde{\mathcal{M}} \longrightarrow \widetilde{\mathcal{N}}$ such that $\forall\gamma \in \pi_1(\mathcal{M})$, $\widetilde{u}\circ\gamma = \varphi_{\sharp 1}(\gamma)\circ\widetilde{u}$, and we minimize the energy integral over a fundamental domain of $\widetilde{\mathcal{M}}$ in this class. This result was subsequentely generalized by R. Schoen and S. T. Yau [193] to the case when the dimension of the target $\mathcal{N}$ is arbitrary, and then to higher dimensions in [30, 225].

3. $\mathcal{E}$ **is a family of maps which are** ***equivariant*** **with respect to a symmetry group.** This means that we are given a group G which acts by isometries $x \longmapsto g\cdot x$ and $y \longmapsto g\cdot y$, $(x \in \mathcal{M},\ y \in \mathcal{N},\ g \in G)$ on $\mathcal{M}$ and $\mathcal{N}$, respectively, and then $\mathcal{E} := \{u : \mathcal{M} \longrightarrow \mathcal{N} |\ \forall g \in G, \forall x \in \mathcal{M}, u(g\cdot x) = g\cdot u(x)\}$. That a critical point under such a symmetry constraint (assuming some extra hypotheses) is also a critical point without the symmetry constraint is the content of a general principle by R. Palais [163]. For a discrete group this approach was used, for example, by L. Lemaire [144] to prove the existence of harmonic maps between a surface $\mathcal{M}$ without boundary of genus $g \geq 2$ and the sphere S^2 which are equivariant with respect to a finite group spanned by reflections with respect to planes in $\mathbb{R}^3$. For continuous groups, this principle is expounded in [64] and the regularity of equivariant minimizing maps is studied by A. Gastel [78]. Many applications concern the reduction of the harmonic map problem to an ODE [55, 64] or to a system in two variables [79, 80], see §5.5.

4. $\mathcal{N}$ **is a manifold with** ***non-positive curvature*****.** This improves strongly the behaviour of minimizing sequences (see §6.3). One instance is the following result [189, Theorem 2.12]: *assume that* $\mathcal{E}$ *is a homotopy class of maps between two compact manifolds* $\mathcal{M}$ *and* $\mathcal{N}$ *of arbitrary dimensions and that* $\mathcal{N}$ *has* non-positive curvature *and let* $v \in \mathcal{C}^3(\mathcal{M},\mathcal{N})$. *Then there exists a harmonic map* $u \in \mathcal{C}^2(\mathcal{M},\mathcal{N})$ *such that* $u = v$ *on* $\partial\mathcal{M}$ *and* u *is homotopic to* v *through maps with fixed values on* $\partial\mathcal{M}$.

5. $\mathcal{E}$ **is a class of diffeomorphisms** between two Riemannian surfaces $\mathcal{M}$ and $\mathcal{N}$: a result by J. Jost and R. Schoen [137, 131] asserts that if $\partial\mathcal{M} = \partial\mathcal{N} = \emptyset$, if $\mathcal{M}$ and $\mathcal{N}$ have the same genus and if $\varphi : \mathcal{M} \longrightarrow \mathcal{N}$ is a diffeomorphism, then *there exists a harmonic diffeomorphism* u *homotopic to* φ *which has the least energy among all diffeomorphisms*

[29] Any *linear* operator between Banach spaces continuous for the strong topologies is continuous for the weak topologies.

homotopic to φ. Actually, the difficulty here is not to get the existence of the minimizer u, but rather to prove that u is weakly harmonic, as not all first variations are allowed.

6. **The target has non-empty boundary.** Again this condition does not cause particular problems when finding a minimizer, but does when proving that this minimizer satisfies, at least weakly, the harmonic maps equation, since, as in the previous example, we are not allowed to use all first variations. However, if $\mathcal{B}$ is a HJW-convex ball of $\mathcal{N}$ (see §6.3 for the definition), and if, for example, $\partial\mathcal{M} \neq \emptyset$ and we fix a Dirichlet boundary condition with values in $\mathcal{B}$, then S. Hildebrandt, W. Jäger and K.-O. Widman [120] prove the existence of a minimizing solution ot the Dirichlet problem with values in $\mathcal{B}$ which is weakly harmonic (in particular the image of the minimizing map does not meet $\partial\mathcal{B}$). A variant of this result was proved by J. Jost [132] in the case dim $\mathcal{M} = 2$: if we fix a boundary condition with values in a sufficiently small ball $\mathcal{B} \subset \mathcal{N}$ and we minimize the energy with this Dirichlet boundary condition *among those maps with values in* $\mathcal{N}$, then the minimizer takes values in $\mathcal{B}$.

In the results [156] in 1. and [193] in 2., by further minimizing over all Dirichlet boundary conditions which parametrizes a Jordan curve in $\mathcal{N}$ in the case of [156], or the conformal structures of $\mathcal{M}$ in the case of [193], the minimizing harmonic map becomes a minimal branched immersion in the sense of §2.2.

5.3 The direct method in a class of maps not closed for the weak topology: case dim $\mathcal{M} = 2$

This case holds in situations where the definition of $\mathcal{E}$ relies partially or completely on the action of maps $u : \mathcal{M} \longrightarrow \mathcal{N}$ on $\pi_2(\mathcal{M})$ or on the degree of maps between two surfaces. See also §3.2 and 3.3.

- For example, consider the case when $\mathcal{M}$ is the 2-dimensional ball B^2 and $\mathcal{N}$ any manifold such that $\pi_2(\mathcal{N})$ is non-trivial, and choose a smooth map $\varphi : B^2 \longrightarrow \mathcal{N}$ which is constant on ∂B^2 and covers a (non-zero) generator of $\pi_2(\mathcal{N})$. Then, as observed in [144], there is no minimizer in the class of maps homotopic to φ which shares the same Dirichlet boundary condition. This is a consequence of the more general result that *any harmonic maps on a ball* B^m *which agrees with a constant on the boundary is a constant map*, proved[30] by L. Lemaire [144] for $m = 2$.
- J. Eells and J. C. Wood [67] proved that any harmonic map of a given degree d between two Riemannian surfaces $\mathcal{M}$ and $\mathcal{N}$ is holomorphic or antiholomorphic if genus $\mathcal{M} + |d$ genus $\mathcal{N}| > 0$. This implies, for example, that *there is no harmonic map of degree* ± 1 *from a* 2*-torus to a* 2*-sphere whatever metrics they are given*, since there is no holomorphic map of degree 1 from a torus to $\mathbb{C}P^1 = S^2$. Hence in particular *the minimum of the energy among degree 1 (or* -1*) maps between a torus and a sphere is not achieved*. This last conclusion remains true if we replace the torus by a higher genus surface, as shown by Lemaire [144] and K. Uhlenbeck independently: a minimizing sequence necessarily converges weakly to a constant map. Furthermore Y. Ge [82] showed that, after extracting a subsequence if necessary, the energy density of such a sequence concentrates at one point.

[30] For $m \geq 3$ this result was extended by J. C. Wood [230] and by H. Karcher and Wood [141].

Bubbles

The first general analysis of the situation when dim $\mathcal{M} = 2$ was done by J. Sacks and K. Uhlenbeck [188] who addressed the question of finding harmonic maps inside a homotopy class $\mathcal{E}$ of maps between a surface $\mathcal{M}$ without boundary and an arbitrary compact manifold $\mathcal{N}$. One of the reasons why $\mathcal{E}$ is not closed with respect to the weak topology, in general, is the conformal invariance of the Dirichlet energy and of the harmonic maps problem in two dimensions (see §2.2). For example, when $\mathcal{M} = S^2$, the group of conformal transformations of S^2 is the group of homographies $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] : z \longmapsto (az+b)/(cz+d)$ acting on S^2 through the stereographic projection (30). Using the action of this group, it is easy to produce minimizing sequences in a homotopy class $\mathcal{E}$ of maps $S^2 \longrightarrow \mathcal{N}$ whose weak limit escapes from the homotopy class (see §3.3). This instability of minimizing sequences can be cured as in [188] by working with the perturbed functional $E^\alpha_{\mathcal{M}}(u) := \int_{\mathcal{M}} (1+|du|^2)^\alpha \mu$, for $\alpha > 1$ which is not conformally invariant anymore (here $\mu := \omega_g / \int_{\mathcal{M}} \omega_g$ is an area 2-form of total integral 1), and then letting $\alpha \to 1$. However a more serious difficulty is the following: imagine that $\pi_2(\mathcal{N})$ has at least two generators γ_1, γ_2 and that, for instance, we know that there exist minimizing harmonic maps $u_1, u_2 : S^2 \longrightarrow \mathcal{N}$ where u_1 (resp. u_2) is a representative of γ_1 (resp. γ_2). Then it may happen that there is no minimizer in the class $\gamma_1 + \gamma_2$: indeed maps in a minimizing sequence could look asymptotically like a map covering the image of u_1 in a neighbourhood of some point $p_1 \in S^2$ and the image of u_2 in a neighbourhood of another point $p_2 \in S^2$ (two *bubbles*), all the other points of S^2 (inside a domain conformally equivalent to a long cylinder) being mapped harmonically to a geodesic connecting a point of $u_1(S^2)$ to a point $u_2(S^2)$ (a *neck*). Then the limit may be either u_1 or u_2 (up to the composition with some conformal map of S^2) or a constant map (mapping S^2 to a point of the geodesic), depending how randomly the instability effects of the conformal group acts. Again by replacing an arbitrary minimizing sequence by a sequence $(u^\alpha)_{\alpha>1}$ of minimizers of $E^\alpha_{\mathcal{M}}$ in $\mathcal{E}$ we can possibly avoid the instability effects of the conformal group, but *we cannot avoid the possible bubblings, i.e. prevent the limit u of $(u^\alpha)_{\alpha>1}$ as $\alpha \to 1$ escaping from $\mathcal{E}$ in general.*

J. Sacks and K. Uhlenbeck prove the following results [188]. They first establish that, if $\alpha > 1$, the functional $E^\alpha_{\mathcal{M}}$ achieves its minimum in each connected component of $W^{1,2\alpha}(\mathcal{M}, \mathcal{N})$ at a smooth map u_α which satisfies the (elliptic) Euler–Lagrange equation of $E^\alpha_{\mathcal{M}}$. Then they prove three basic results:

(i) **The main estimate**. There exists $\varepsilon > 0$ and $\alpha_0 > 1$ such that, for any geodesic ball $\mathcal{B} \subset \mathcal{M}$, any map $u : \mathcal{B} \longrightarrow \mathcal{N}$ with $E^2_{\mathcal{B}}(u) < \varepsilon$ which is a smooth critical point of $E^\alpha_{\mathcal{M}}$ for some $\alpha \in [1, \alpha_0)$, we have a uniform family of estimates $||du||_{W^{1,p}(\mathcal{B}')} \leq C(p, \mathcal{B}')||du||_{L^2(\mathcal{B})}$ for any $p \in (1, \infty)$ and any smaller disk $\mathcal{B}' \subset \mathcal{B}$.

(ii) **The removability of isolated singularities for weakly harmonic maps**. This says that, for any map $u \in W^{1,2}(\mathcal{M}, \mathcal{N})$, and any finite family of points $\{z_1, \cdots, z_k\} \subset \mathcal{M}$ such that u is smooth and harmonic on $\mathcal{M} \setminus \{z_1, \cdots, z_k\}$, there exists a smooth extension of u to $\mathcal{M}$ which is harmonic.

(iii) **An energy gap**. $\exists \varepsilon > 0$, $\exists \alpha_0 > 1$ such that for any map $u \in W^{1,2}(\mathcal{M}, \mathcal{N})$ which is a critical point of $E^\alpha_{\mathcal{M}}$ for some $\alpha \in [1, \alpha_0)$, if $E_{\mathcal{M}}(u) < \varepsilon$, then u is constant.

Note that the proofs of (ii) and (iii) use (i). Thanks to the *main estimate* and a covering argument, Sacks and Uhlenbeck prove that a subsequence of the family of $E^\alpha_{\mathcal{M}}$-minimizers

$(u_\alpha)_{\alpha>1}$ converges to some map $u \in W^{1,2}(\mathcal{M},\mathcal{N})$ in the weak $W^{1,2}$ topology and in $\mathcal{C}^1(\mathcal{M}\setminus\{z_1,\cdots,z_k\},\mathcal{N})$, where $\{z_1,\cdots,z_k\}$ is a finite collection of points of $\mathcal{M}$ where possible bubblings occur. Then, by the result of *removability of isolated singularities* (ii), we deduce that u extends to a smooth harmonic map. However nothing guarantees that this map is non-constant. On the other hand, an analysis of the behaviour of u_α near the bubbling points z_j reveals that, *if* $|du_\alpha|$ *is not bounded in a neighbourhood of* z_j, then we can find a subsequence of maps $v_{j,\alpha} : B^2(0,R_\alpha) \longrightarrow \mathcal{N}$ (where $\lim_{\alpha\to 1} R_\alpha = +\infty$), defined by $v_{j,\alpha}(x) = u_\alpha(\exp_{x_\alpha}(\lambda_\alpha x))$, where (x_α) is a sequence of points of $\mathcal{M}$ which converges to z_j and $\lim_{\alpha\to 1}\lambda_\alpha = 0$, such that for any ball $B^2(0,R)\subset\mathbb{R}^2$, the restriction of $v_{j,\alpha}$ to $B^2(0,R)$ converges in $\mathcal{C}^1(B^2(0,R))$ to the restriction to $B^2(0,R)$ of some map v_j as $\alpha\to 1$, and $v_j : \mathbb{R}^2 \longrightarrow \mathcal{N}$ is a harmonic map of finite energy. Since $\mathbb{R}^2$ is conformally equivalent to S^2 minus a point and thanks again to the *removability of isolated singularities* result, we can extend v_j to a harmonic map $S^2 \longrightarrow \mathcal{N}$ (moreover we know that any harmonic map on S^2 is conformal, i.e. holomorphic or antiholomorphic, see §2.2). Hence we can picture the limit of u_α as the collection of harmonic maps $u : \mathcal{M} \longrightarrow \mathcal{N}$ and $v_j : S^2 \longrightarrow \mathcal{N}$, for $1 \le j \le k$, with the extra (lost) information that the image of each map v_j is connected by a geodesic to the point $u(z_j)$ (a so-called *bubble tree*). We have moreover[31]:

$$E_{\mathcal{M}}(u) + \sum_{j=1}^{k} E_{S^2}(v_j) = \lim_{\alpha\to 1}\sup E(u_\alpha). \tag{49}$$

By using this analysis and the *energy gap* property, Sacks and Uhlenbeck deduce the following results:

- if $\pi_2(\mathcal{N}) = 0$, or if we minimize in a conjugacy class of homomorphisms $\pi_1(\mathcal{M}) \longrightarrow \pi_1(\mathcal{N})$, the maps u_α converge strongly to u, hence u is a minimizer of the energy in the same class as u_α. We thus recover the results of L. Lemaire [144] or R. Schoen and S. T. Yau [193]. Here the conclusion is achieved by constructing test maps $\widehat{u}_\alpha$ which coincide with u_α away from the bubbling points and with the weak limit u near the bubbling points: because of the topological hypotheses, $\widehat{u}_\alpha$ is in the same topological class as u_α and hence we can exploit the inequality $E_\alpha(u_\alpha) \le E_\alpha(\widehat{u}_\alpha)$.
- if $\pi_2(\mathcal{N}) \neq 0$, choose $\mathcal{M} = S^2$ and a non-trivial *free 2-homotopy class* Γ of $\mathcal{N}$, i.e. a connected component of $\mathcal{C}^1(S^2,\mathcal{N})$ which does not contain the constant maps. Then **either γ contains a minimizing harmonic map or, for all $\delta > 0$, there exists non-trivial free 2-homotopy classes Γ_1 and Γ_2 such that $\Gamma \subset \Gamma_1 + \Gamma_2$ and $\inf_{v\in\Gamma_1} E_{\mathcal{M}}(v) + \inf_{v\in\Gamma_2} E_{\mathcal{M}}(v) < \inf_{v\in\Gamma} E_{\mathcal{M}}(v) + \delta$.**
- **if $\pi_2(\mathcal{N}) \neq 0$ and $\mathcal{M} = S^2$, there exist a set of free homotopy classes $\Lambda = \{\Gamma_i |\, i \in I\} \subset \pi_0\mathcal{C}^1(S^2,\mathcal{N})$ which forms a generating set for $\pi_2(\mathcal{N})$ under the action[32] of $\pi_1(\mathcal{N})$ such that each $\Gamma_i \in \Lambda$ contains a minimizing harmonic map**.

Note that the last result implies that **there exists a non-trivial harmonic map $S^2 \longrightarrow \mathcal{N}$ as soon as** $\pi_2(\mathcal{N}) \neq 0$. The second result can be translated into the following: *if* $\pi_2(\mathcal{N}) \neq$

[31] Sacks and Uhlenbeck just proved the inequality $\le$ in (49); the equality in (49) was established by J. Jost [132] and T. H. Parker [164].

[32] The set $\pi_0\mathcal{C}^1(S^2,\mathcal{N})$ of free homotopy classes can be identified with the set of orbits of the natural action of $\pi_1(\mathcal{N})$ on $\pi_2(\mathcal{N})$.

0, $\mathcal{M} = S^2$ *and* Γ *is a non-trivial* free 2-homotopy class *of* $\mathcal{N}$, *then if there exists* $\delta > 0$ *such that, for any non-trivial free 2-homotopy classes* Γ_1 *and* Γ_2 *with* $\Gamma \subset \Gamma_1 + \Gamma_2$ *we have*

$$\inf_{v\in\Gamma} E_{\mathcal{M}}(v) \leq \inf_{v\in\Gamma_1} E_{\mathcal{M}}(v) + \inf_{v\in\Gamma_2} E_{\mathcal{M}}(v) - \delta\,, \tag{50}$$

then the minimum of $E_{\mathcal{M}}$ *is achieved in* Γ. This important property is connected with a similar observation made previously by T. Aubin for the Yamabe problem [5] and with further subsequent developments like the results by C. Taubes [212] for the Yang–Mills connections on a 4-dimensional manifold or the concentration compactness principle of P.-L. Lions [150].

Remarks (i) An alternative analysis with improvements to the understanding of the bubbling phenomenon have been obtained by J. Jost [130, 131, 133] by using a method reminiscent of the *balayage* technique of H. Poincaré.

(ii) Further refinements to the analysis of bubbling were made by T. H. Parker [164] by using the notion of *bubble tree*, which was introduced previously by Parker and J. Wolfson in the study of *pseudo-holomorphic curves*, and by W. Y. Ding and G. Tian [58]. The heat flow equation also provides another approach, which was used by M. Struwe to recover the theory of Sacks and Uhlenbeck (see below).

(iii) The influence of bubbling phenomena is not confined to harmonic maps of surfaces, but plays a major role in the existence theory of harmonic maps in higher dimensions, as expounded in §5.4, and in regularity theory (see the results on reduction of the singular set of stationary maps in §4.3).

Applications of the theory of bubbling

In some cases a precise analysis to decide whether (50) holds is possible: this was done first by H. Brezis and J.-M. Coron [22] and J. Jost [131] independently. We set $\mathcal{M} = B^2$, the unit ball in $\mathbb{R}^2$, and $\mathcal{N} = S^2$ and we let $\gamma \in T^2(\partial B^2, S^2) :=$ the set of maps $\gamma : \partial B^2 \longrightarrow S^2$ such that there exists $u \in W^{1,2}(B^2, S^2)$ with $u|_{\partial B^2} = \gamma$. Then the class $\mathcal{E} := W^{1,2}_\gamma(B^2, S^2) := \{u \in W^{1,2}(B^2, S^2)|\ u|_{\partial B^2} = \gamma\}$ is non-empty and closed for the weak $W^{1,2}$ topology. Hence application of the direct method provides us with a smooth harmonic map $\underline{u}$ which minimizes E_{B^2} in $\mathcal{E}$. We now consider the functional on $W^{1,2}_\gamma(B^2, S^2)$ defined by

$$Q(u) := \frac{1}{4\pi}\int_{B^2}\Big\langle u\,,\ \frac{\partial u}{\partial x}\times\frac{\partial u}{\partial y}\Big\rangle d^2x = \frac{1}{4\pi}\int_{B^2} u^*\omega_{S^2}.$$

We observe that Q takes discrete values on $W^{1,2}_\gamma(B^2, S^2)$, more precisely: for all $u \in W^{1,2}_\gamma(B^2, S^2)$, $Q(u) - Q(\underline{u}) \in \mathbb{Z}$. The geometric interpretation of this is that, if we consider the map $u\sharp\underline{u} : S^2 \longrightarrow S^2$ defined via the identification $\mathbb{C}\cup\{\infty\} \simeq S^2$ by setting $u\sharp\underline{u} = u$ on B^2 and $(u\sharp\underline{u})(z) = \underline{u}(z/|z|^2)$ on $\mathbb{C}\setminus B^2$, then $Q(u) - Q(\underline{u})$ is the degree of $u\sharp\underline{u}$. Then for any $k \in \mathbb{Z}$, the classes $\mathcal{E}_k := \{u \in \mathcal{E}|\ Q(u) - Q(\underline{u}) = k\}$ are the connected components of $\mathcal{E}$ for the strong $W^{1,2}$ topology. So they are the free 2-homotopy classes of S^2. But they are not closed for the weak topology; hence it is not clear whether $\inf_{\mathcal{E}_k} E_{B^2}$ is achieved. However, one can prove that, *if* γ *is not constant, then there exists some* $v \in \mathcal{E}$ *such that* $|Q(v) - Q(\underline{u})| = 1$ *and* $E_{B^2}(v) < E_{B^2}(\underline{u}) + 4\pi$. But since the minimum of the energy in any *non-trivial homotopy class* of maps $S^2 \longrightarrow S^2$ is greater or equal to 4π, this shows that (50) holds, hence it follows that **there is a harmonic map $\overline{u}$ which minimizes E_{B^2} in its homotopy class, the latter being either $\mathcal{E}_1$ or $\mathcal{E}_{-1}$.** Moreover, as proved in [22], in the case when γ is the restriction of the inverse $P^{-1} : \mathbb{R}^2 \longrightarrow S^2$ of stereographic projection (30) to $\partial B^2 \subset \mathbb{R}^2$, the constructed solutions $\underline{u}$ and $\overline{u}$ (which here are restrictions to B^2 of stereographic projections) are the only miminizers in their respective class and moreover there are no minimizers in the other classes.

The following generalization was obtained partially by A. Soyeur [203] and later completed by E. Kuwert [142] and J. Qing [174] independently, see also [88] for an exposition. We first associate two degrees d^- and d^+ to the boundary data $\gamma \in T^2(\partial B^2, S^2)$: if γ has a *holomorphic (resp. antiholomorphic) extension* u^+ (resp. u^-) inside B^2 with values in $S^2 \simeq \mathbb{C}P$ we let $d^+ := Q(u^+) - Q(\underline{u})$ (resp. $d^- := Q(u^-) - Q(\underline{u})$), if γ has no holomorphic (resp. no antiholomorphic) extension inside B^2, set $d^+ := +\infty$ (resp. $d^- := -\infty$). Note that we always have $d^- \leq d^+$, with equality if and only if γ is a constant. Then

(i) **for $k \in \mathbb{Z}$ which satisfies $k \in (-\infty, d^-) \cup (d^+, \infty)$, the minimum of E_{B^2} is never achieved in $\mathcal{E}_k$.** Furthermore if $k \in (-\infty, d^-] \cup [d^+, \infty)$, $\inf_{\mathcal{E}_k} E_{B^2} = \inf_{\mathcal{E}_{d^\pm}} E_{B^2} + 4\pi|k - d^\pm|$, where $d^\pm = d^-$ if $k \leq d^-$ and $d^\pm = d^+$ if $k \geq d^+$ (so that, in particular, (50) does not hold);

(ii) **for all $k \in \mathbb{Z}$ such that $d^- \leq k \leq d^+$, the minimum of E_{B^2} is achieved in $\mathcal{E}_k$.**

Similar results have been obtained by Qing [176] for maps with values in a Kähler manifold.

The heat flow

Observe that, in the method of J. Sacks and K. Uhlenbeck, the family of minimizers $(u_\alpha)_{\alpha>1}$ of $E^\alpha_{\mathcal{M}}$ produces particular minimizing sequences for $E_{\mathcal{M}}$ as $\alpha \to 1$. One of the advantages of this is that, not only does it help to balance the instability due to the action of the conformal group, but it also gives us some control of the tension field (25). Another natural way to control the tension field for a minimizing sequence is to consider the heat flow equation:

$$\frac{\partial u}{\partial t} = \Delta_g u + g^{ij} A_u\Big(\frac{\partial u}{\partial x^i}, \frac{\partial u}{\partial x^j}\Big) \quad \text{on } [0,T) \times \mathcal{M}. \tag{51}$$

The study of this equation was initiated in [66, 1, 99] in the case when the curvature of the target manifold $\mathcal{N}$ is non-positive (see §6.3). If we remove this hypothesis, the first results[33] were obtained by M. Struwe [207], for the case when dim $\mathcal{M} = 2$ and $\partial\mathcal{M} = \emptyset$:

[33] Following the result of Struwe [207], further results on the heat flow when dim $\mathcal{M} \geq 3$ and with no assumption on the curvature of $\mathcal{N}$ were obtained: the first existence results were obtained by Y. M. Chen [41] for $\mathcal{M}$ arbitrary and $\mathcal{N} = S^n$, and by Struwe [208] for $\mathcal{M} = \mathbb{R}^m$ and $\mathcal{N}$ an arbitrary compact manifold. By putting together their ideas, Chen and Struwe [43] obtained the following existence result: *for any map $u_0 \in W^{1,2}(\mathcal{M}, \mathcal{N})$, there exists a weak solution to the heat flow equation defined for all time and with Cauchy data u_0, i.e. coinciding with u_0 at $t = 0$. This solution is regular outside a singular set which has locally finite m-dimensional Hausdorff measure with respect to the parabolic metric.* Then J.-M. Coron and J.-M. Ghidaglia [50] produced the first examples of weak solutions which blow up at finite time, hence proving that there are no classical solutions in general and Coron [49] built an example of Cauchy data for which there are infinitely many weak solutions to the heat flow equation (actually the Cauchy data is a weakly harmonic map). Later on, similar blow-up and non-uniqueness results were proved for the heat flow on surfaces (see the next paragraph).

for any $u_0 \in W^{1,2}(\mathcal{M},\mathcal{N})$*, there exists a* **global weak** *solution* $u : [0,\infty) \times \mathcal{M} \longrightarrow \mathcal{N}$ *of the heat equation (51) which satisfies the energy decay estimate:*

$$E_{\mathcal{M}}(u(T,\cdot)) + \int_0^T \int_{\mathcal{M}} \left|\frac{\partial u}{\partial t}\right|^2 \omega_g dt \;\leq\; E_{\mathcal{M}}(u_0), \quad \forall T > 0 \tag{52}$$

and which is smooth outside finitely many singular points $(\overline{t}_j, \overline{x}_j)_{1\leq j\leq k}$*. The solution is unique in this class. Moreover, at each singularity* $(\overline{t}_j, \overline{x}_j)$*, a harmonic sphere* v_j *bubbles off, i.e. there exists a sequence* $(t_{\ell,j}, x_{\ell,j})_{\ell\in\mathbb{N}}$ *which converges to* $(\overline{t}_j, \overline{x}_j)$ *(with* $t_{\ell,j} < \overline{t}_j$*) such that* $u_{\ell,j}(x) := u(t_{\ell,j}, \exp_{x_{\ell,j}}(\lambda_{\ell,j}x))$ *converges to* v_j *in* $W^{2,2}_{loc}(\mathbb{R}^2,\mathcal{N})$*, where* $(\lambda_{\ell,j})_{\ell\in\mathbb{N}}$ *is a sequence of positive numbers such that* $\lim_{\ell\to\infty}\lambda_{\ell,j} = 0$*. The map* v_j *can then be extended to a smooth harmonic map* $S^2 \longrightarrow \mathcal{N}$*. Lastly, there exists a sequence* $(T_\ell)_{\ell\in\mathbb{N}}$ *of times such that* $\lim_{\ell\to\infty} T_\ell = \infty$ *and* $u(T_\ell,\cdot)$ *converges weakly in* $W^{1,2}$ *to a smooth harmonic map* $u_\infty : \mathcal{M} \longrightarrow \mathcal{N}$ *as* $\ell \to \infty$. This result was extended to the case when $\partial\mathcal{M} \neq \emptyset$ by K. C. Chang [37]. These results can be used to recover similar results to those of Sacks and Uhlenbeck, see for example the last chapter of the book of Struwe [209].

The question of whether the solutions to the heat flow equation in two dimensions really develop singularities remained open for some time until K. C. Chang, W. Ding, R. Ye [38] constructed an example of an initial condition $u_0 : S^2 \longrightarrow S^2$ for which the heat flow does blow up in finite time. Note that the inequality in the estimate (52) would be straighforward if the solution were smooth (just multiply the heat equation by u and integrate). Actually the left-hand side of (52) is smooth outside the blow up times $(\overline{t}_j)_{1\leq j\leq k}$. In [177] J. Qing proved that, at these bubbling points, the discontinuity of this left-hand side is *just* equal to minus the sum of the energies of the harmonic spheres v_j which separate, i.e., there is no energy loss in the necks connecting u_∞ (the 'body' map) to the v_j's (the 'bubble' maps). He further proved that, if at some time $\overline{t}$ there are p harmonic spheres $v_{j_1}, \cdots v_{j_p}$ bubbling off, then $\lambda_{\ell,i}/\lambda_{\ell,j} + \lambda_{\ell,j}/\lambda_{\ell,i} + |x_{\ell,j} - x_{\ell,i}|^2/(\lambda_{\ell,i}\lambda_{\ell,j}) \to \infty$ as $\ell \to \infty$ for $i, j \in \{j_1, \cdots, j_p\}$ such that $i \neq j$; roughly speaking, this means that each bubble decouples from the other ones in distance or in scale. The analysis of what is happening in the necks was further refined in [178, 215]. In [213] P. Topping proved that *if* $\mathcal{M} = \mathcal{N} = S^2$ *and if one assumes the hypothesis* (H): u_∞ and the v_j's are either all holomorphic or all antiholomorphic, *then* $u(t,\cdot)$ *converges uniformly in time as* $t \to \infty$ *strongly in* $L^p(S^2,\mathbb{R}^3)$ *and in* $W^{1,2}(S^2\setminus\{\overline{x}_1, \cdots, \overline{x}_k\})$. The latter result depends strongly on the fact that the target is S^2 (see [213]).

The uniqueness of weak solutions to (51) was proved by A. Freire [74], *under the further assumption that* $E_{\mathcal{M}}(u(t,\cdot))$ *is a non-increasing function of* t. But, in [214], P. Topping constructed solutions of the heat flow from a surface to S^2 *which are different* from Struwe's solution, hence proving the **non-uniqueness of weak solutions to equation** (51), **in general**. The point, however, is that Topping's solutions are obtained by *attaching bubbles*, i.e. have the reverse behaviour of Struwe's solutions, so that the energy $E_{\mathcal{M}}(u(t,\cdot))$ *increases* by a jump of 4π each time a bubble is attached.

Lastly, in [215], P. Topping performed a very fine analysis of *almost-harmonic* maps from S^2 to S^2, i.e. maps $u \in W^{1,2}(S^2, S^2)$ such that the L^2 norm of the tension field $\tau(u) = \Delta_{S^2}u + |du|^2u$ is small. Recall that, if $\tau(u) = 0$, then u is harmonic and hence either holomorphic or antiholomorphic, so that its energy is 4π times its degree in $\mathbb{Z}$. P. Topping proved that this quantization of the energy remains true for almost-harmonic maps and more precisely establishes the estimate: $|E_{S^2}(u) - 4\pi k| \le C||\tau(u)||^2_{L^2(S^2)}$ (for some $k \in \mathbb{Z}$), for all u in $W^{1,2}(S^2, S^2)$ except for some exceptional special cases. This allows him to recover the same conclusions as in [213] concerning the convergence in time and the uniqueness of the location of the singularities of the heat flow, but *without* assuming the hypothesis (H) above. These results are strong in the sense that an almost-harmonic map u may have, for example, a holomorphic body with anti-holomorphic bubbles attached, and then u is *not* close to a harmonic map in the $W^{1,2}(S^2)$ topology. To deal with such cases, Topping established an estimate asserting the existence of a repulsive effect between holomorphic and antiholomorphic components of a bubble tree.

5.4 The direct method in a class of maps not closed for the weak topology: case dim $\mathcal{M} \ge 3$

Some cases where the class $\mathcal{E} \subset W^{1,2}(\mathcal{M}, \mathcal{N})$ chosen for the minimization of $E_{\mathcal{M}}$ is *not* weakly closed have already been described in §3.3. We will here mainly discuss other situations, starting from the work of H. Brezis, J.-M. Coron and E. H. Lieb [24].

Prescribing singularities

We begin with an example. Let $\Omega \subset \mathbb{R}^3$ and $a \in \Omega$; we will choose a subset $\mathcal{E}$ of $\mathcal{C}^1(\Omega \setminus \{a\}, S^2) \cap W^{1,2}(\Omega, S^2)$. Note that $\mathcal{C}^1(\Omega \setminus \{a\}, S^2) \cap W^{1,2}(\Omega, S^2)$ contains the map u_a defined by $u_a(x) = (x - a)/|x - a|$ (which is even weakly harmonic). Moreover, for each sphere $S^2_{a,r} = \partial B^3(a, r)$ centred on a which is contained in $\overline{\Omega}$, the restriction of u_a to $S^2_{a,r}$ has degree 1. Let us fix

$$\mathcal{E} := \left\{u \in \mathcal{C}^1(\Omega \setminus \{a\}, S^2) \cap W^{1,2}(\Omega, S^2) \mid \deg u|_{S^2_{a,r}} = 1, \text{ for } S^2_{a,r} \subset \Omega\right\}.$$

Then, in some sense, the minimization of E_Ω in $\mathcal{E}$ extends the problem of minimizing the energy among maps between surfaces of a given degree (see §5.3). Indeed, as shown in [24], after the extraction of a subsequence if necessary, a minimizing sequence $(u_k)_{k\in\mathbb{N}}$ of E_Ω in $\mathcal{E}$ converges weakly to a constant map c, *in all cases except if* Ω *is a ball centred at* a. If we assume, for simplicity, that there exists an unique line segment $[a, b]$ which joins a to the nearest point in $\partial\Omega$ (i.e., such that $b \in \partial\Omega$ and $d(a, \partial\Omega) = |b-a|$) then u_k converges strongly to c on $\Omega \setminus V_\varepsilon[a, b]$, where $V_\varepsilon[a, b]$ is a neighbourhood of $[a, b]$. Furthermore, the restriction of u_k to a sphere $S^2_{a,r}$ will be almost constant outside the intersection of $S^2_{a,r}$ with $V_\varepsilon[a, b]$, whereas it will almost conformally cover the target S^2 on $S^2_{a,r} \cap V_\varepsilon[a, b]$. Hence a *line of bubbles* separates from u_k along $[a, b]$. Lastly, the infimum of the energy, $\inf_{u\in\mathcal{E}} E_\Omega(u)$, is precisely $4\pi|b-a|$, i.e. the area of S^2 times the length of the line segment. A similar situation, arises if we have a **dipole** as introduced in [24]. Here we assume that

$$\mathcal{E} := \left\{u \in \mathcal{C}^1(\Omega \setminus \{p, n\}, S^2) \cap W^{1,2}(\Omega, S^2) \mid \deg u|_{S^2_{p,r}} = 1,\right.$$
$$\left.\deg u|_{S^2_{n,r}} = -1 \text{ for } S^2_{p,r}, S^2_{n,r} \subset \Omega\right\},$$

where $p, n \in B^3$ are two distinct points. Then, a minimizing sequence in the class $\mathcal{E}$ converges to a constant outside a neighbourhood of the line segment $[p, n]$, and its energy concentrates along $[p, n]$.

Actually, a more general situation was considered in [24]: let $\{a_1, \cdots, a_p\} \subset \Omega \subset \mathbb{R}^3$ be an arbitrary finite collection of points of $\Omega \subset \mathbb{R}^3$ and $d_1, \cdots, d_p \in \mathbb{Z}$. Then set

$$\mathcal{E} := \big\{u \in \mathcal{C}^1(\Omega \setminus \{a_1, \cdots, a_p\}, S^2) \cap W^{1,2}(\Omega, S^2) \,|\, \forall\, i = 1, \cdots, p, \\ \deg u|_{S^2_{a_i,r}} = d_i, \text{ for } S^2_{a_i,r} \subset \Omega\big\}.$$

In order to describe the behaviour of a minimizing sequence in $\mathcal{E}$ we need to define the notion of a *minimal connection* as introduced in [24]. For simplicity, we will assume that $\Omega = B^3 := B^3(0,1)$ and that the total degree $Q := \sum_{i=1}^p d_i$ is *zero*. First, call the points a_i such that $d_i > 0$ (resp. $d_i < 0$) *positive* (resp. *negative*) (points a_i such that $d_i = 0$ do not play any role in the following, hence we can forget about them without loss of generality). We list the positive points with each a_i repeated d_i times and write this list as $p_1, \cdots, p_\kappa$. Likewise we list the negative points as $n_1, \cdots, n_{\kappa'}$. Note that $\kappa - \kappa' = Q = 0$. A *connection* C is then a pairing of the two lists $(p_1, n_{\sigma(1)}), \cdots, (p_\kappa, n_{\sigma(\kappa)})$, where σ is a permutation of $\{1, \ldots, \kappa\}$. The *length* of the connection C is $L(C) := \sum_{i=1}^\kappa d(p_i, n_{\sigma(i)})$. Lastly, the *length of the minimal connection* is: $L := \min_C L(C)$ and a *minimal connection* is a connection $\underline{C}$ (which may not be unique) such that $L(\underline{C}) = L$. Then **the infimum**[34] **of E_{B^3} on $\mathcal{E}$ is** $4\pi L$ and, if we exclude the case when $\{a_1, \cdots, a_p\} = \emptyset$ or $\{0\}$, we have:

- this infimum is never achieved and, after extraction of a subsequence if necessary, a minimizing sequence $(u_k)_{k\in\mathbb{N}}$ of E_{B^3} in $\mathcal{E}$ converges weakly to a constant map;
- again, after extraction of a subsequence if necessary, lines of bubbles separate from u_k along a minimal connection $\underline{C}$. More precisely, the energy density $\frac{1}{2}|du_k|^2$ converges weakly in Radon measures to a measure μ supported by a minimal connection: for all measurable $A \subset B^3$, $\mu(A) = 4\pi\mathcal{H}^1(A \cap \underline{C})$, where $\mathcal{H}^1$ is the 1-dimensional Hausdorff measure (see §4.3 for the definition).

Moreover, the locations and degrees of the singularities of a map $u \in W^{1,2}(\Omega, S^2) \cap \mathcal{C}^1(\Omega \setminus \{a_1, \cdots, a_p\}, S^2)$ can be detected by computing the differential of the 2-form $u^*\omega_{S^2}$ (see §3.2), because of the relation:

$$d(u^*\omega_{S^2}) = \Big(\sum_{i=1}^p d_i\delta_{a_i}\Big) dx^1 \wedge dx^2 \wedge dx^3, \quad \text{where } \delta_{a_i} \text{ is the Dirac mass at } a_i\,. \tag{53}$$

Note that the coefficients of $u^*\omega_{S^2}$ are in $L^1(\Omega)$ and equation (53) holds in the distribution sense, i.e., $\int_{\partial B^3} \zeta(u^*\omega_{S^2}) - \int_{B^3} d\zeta \wedge u^*\omega_{S^2} = \sum_{i=1}^p d_i\zeta(a_i)$, $\forall \zeta \in \mathcal{C}^\infty(\Omega)$. In fact, the latter relation makes sense even if ζ belongs to the set $Lip(\Omega)$ of Lipschitz continuous functions on Ω. This leads to an alternative (dual) formula[35] for the length of the minimal

[34] An alternative proof of the inequality $\inf_{u\in\mathcal{E}} E_{B^3}(u) \geq 4\pi L$ was given by F. Almgren, W. Browder and E. H. Lieb [3] by using the *coarea formula* $\int_{B^3}(J_2u)(x)dx^1dx^2dx^3 = \int_{y\in S^2} \mathcal{H}^1(u^{-1}(y))d\mathcal{H}^2(y)$, valid for a smooth map $u : B^3 \longrightarrow S^2$. Here $\mathcal{H}^2$ is the 2-dimensional Hausdorff measure on S^2, $\mathcal{H}^1$ is the 1-dimensional Hausdorff measure on a generic fibre $u^{-1}(w)$ of u and $(J_2u)(x)$ denotes the 2-dimensional Jacobian of u at x. Note that the coarea formula has been extended to Sobolev mappings between manifolds by P. Hajłasz [97], leading to another variant of the proof of the Brezis, Coron and Lieb result.

[35] Note that $Q = 0$ implies that $\int_{\partial B^3} u^*\omega_{S^2} = 0$, so that the maximum in (54) is finite.

connection:

$$\begin{aligned} L(u) &= \max_{\zeta \in Lip(\Omega),\, |\nabla\zeta|_{L^\infty} \le 1} \sum_{i=1}^{p} d_i \zeta(a_i) \\ &= \max_{\zeta \in Lip(\Omega),\, |\nabla\zeta|_{L^\infty} \le 1} \left\{ \int_{\partial B^3} \zeta(u^*\omega_{S^2}) - \int_{B^3} d\zeta \wedge u^*\omega_{S^2} \right\}. \end{aligned} \tag{54}$$

But **the right-hand side of** (54) **makes sense for an arbitrary map** $u \in W^{1,2}_\varphi(B^3, S^2)$**, and can be used to extend the definition of** $L(u)$ **to the whole of** $W^{1,2}_\varphi(B^3, S^2)$ **if the degree of** φ **is zero**. Moreover, it was proved by Bethuel, Brezis and Coron [15] that *the functional* $L : W^{1,2}_\varphi(B^3, S^2) \longrightarrow \mathbb{R}$ *is continuous for the strong* $W^{1,2}$ *topology.* Lastly, a result of Brezis and P. Mironescu [20, 21] asserts that, for any $u \in W^{1,2}(B^3, S^2)$ such that $u|_{\partial B^3}$ is a smooth map of degree 0, there exist two sequences $(p_1, p_2, \ldots)$ and $(n_1, n_2, \ldots)$ of points of B^3 such that

$$d(u^*\omega_{S^2}) = 4\pi \sum_{i=1}^{\infty} (\delta_{p_i} - \delta_{n_i}) \tag{55}$$

and $\sum_{i=1}^{\infty} |p_i - n_i| < \infty$. Then $L(u)$ is equal to the infimum of all sums $\sum_{i=1}^{\infty} |p_i - n_i|$ such that (55) holds.

Generalizations. Similar situations occur, for instance, if we work in $W^{1,n}(\mathcal{M}, S^n)$, where dim $\mathcal{M} \ge n+1$, and we try to minimize the n-energy among maps which are smooth outside a codimension $n+1$ submanifold Σ and which have prescribed degree around each connected componen of Σ. This case was first considered by F. Almgren, W. Browder and E. H. Lieb [3], who pointed out that the minimal connection has to be replaced by an n-area minimizing integral current. We refer to [88, Chapter 5] for subsequent developments.

The gap phenomenon

An important and surprising observation was made by R. Hardt and F. H. Lin [106] at about the same time: we still assume that $\mathcal{M} = B^3$ and $\mathcal{N} = S^2$ and we let $\varphi : \partial B^3 \longrightarrow S^2$ be a smooth map *of degree* 0. Then $\mathcal{C}^1_\varphi(B^3, S^2) := \{u \in \mathcal{C}^1(B^3, S^2)|\ u = \varphi \text{ on } \partial B^3\}$ is not empty and we may consider its closure $H^1_{\varphi,s}(B^3, S^2)$ in the strong $W^{1,2}$ topology. Another natural class is $W^{1,2}_\varphi(B^3, S^2) := \{u \in W^{1,2}(B^3, S^2)|\ u = \varphi \text{ on } \partial B^3\}$. Then it is proved in [106] that we can choose the boundary conditions φ such that:

$$\inf_{u \in \mathcal{C}^1_\varphi(B^3, S^2)} E_{B^3}(u) = \inf_{u \in H^1_{\varphi,s}(B^3, S^2)} E_{B^3}(u) > \inf_{u \in W^{1,2}_\varphi(B^3, S^2)} E_{B^3}(u)\,. \tag{56}$$

This implies that the inclusion $H^1_s(B^3, S^2) \subset W^{1,2}(B^3, S^2)$ is strict, as discussed in §3.2. The construction of φ relies on ideas close to the preceding discussion: imagine that we fix two *dipoles* of length $\ell > 0$, i.e. pairs of points (p_1, n_1) and (p_2, n_2) with opposite degrees ± 1, such that $|p_1 - n_1| = |p_2 - n_2| = \ell$. Place the points p_1 and n_1 very close to the *north* pole $(0, 0, 1)$ of ∂B^3, with p_1 *outside* B^3 but n_1 *inside* B^3, specifically, $p_1 = (0, 0, 1 + \ell/2)$ and $n_1 = (0, 0, 1 - \ell/2)$. Similarly, place p_2 and n_2 very close to the *south* pole: $p_2 = (0, 0, -1 + \ell/2)$ and $n_2 = (0, 0, -1 - \ell/2)$. This is all embedded in, say, $B^3(0, 2)$. Now consider how a sequence of maps $(v_k)_{k \in \mathbb{N}}$ in $W^{1,2}(B^3(0, 2), S^2)$ which minimizes $E_{B^2(0,2)}$ in the class of maps v such that $d(v^*\omega_{S^2}) = \delta_{p_1} + \delta_{p_2} - \delta_{n_1} - \delta_{n_2}$

would look: v_k is almost constant outside neighbourhoods of the line segments $[p_1, n_1]$ and $[p_2, n_2]$, and the restriction of v_k to a piece of surface cutting one of these segments transversally covers S^2 almost conformally. Then we take $\varphi = (v_k)|_{\partial B^2}$ for k large enough. We observe that

(i) the degree of φ is equal to the sum of the degrees of the singularities n_1 and p_2 enclosed by ∂B^3, i.e., 0 ;

(ii) $\inf_{u\in W^{1,2}_\varphi(B^3,S^2)} E_{B^3}(u)$ is certainly smaller than $E_{B^3}(v_k)$, which is of order $4\pi\ell$;

(iii) $\inf_{u\in \mathcal{C}^1_\varphi(B^3,S^2)} E_{B^3}(u)$ is of order 8π.

Hence, (56) follows by choosing ℓ sufficiently small. To prove (iii), we estimate the energy of any map $\psi \in \mathcal{C}^1_\varphi(B^3, S^2)$ from below as follows. For any $h \in (-1, 1)$, consider the disk D_h which is the intersection of B^3 with the plane $\{x^3 = h\}$ and the domain $H_h := B^3 \cap \{x^3 < h\}$: its boundary ∂H_h is the union of D_h and the spherical cap $S_h := (\partial B^3)\cap\{x^3 \leq h\}$. On the one hand, the restriction of ψ to S_h is almost constant except in a small neighbourhood of the south pole, where $\psi|_{S_h}$ covers almost all of S^2 with degree 1, and on ∂D_h, the map ψ is nearly constant. On the other hand, since ψ is continuous inside H_h, the degree of its restriction to ∂H_h is 0. These two facts imply that the restriction $\psi|_{D_h}$ should almost cover S^2 with degree -1. Hence $\int_{D_h} \frac{1}{2}|d\psi|^2 d^3x \geq |\int_{D_h} \psi^*\omega_{S^2}| \simeq 4\pi$. By integrating this inequality on $h \in (-1, 1)$ we obtain (iii).

The relaxed energy

Exploiting the fact that $H^1_{\varphi,w}(B^3, S^2) = W^{1,2}_\varphi(B^3, S^2)$ (see §3.2), i.e. that $\forall u \in W^{1,2}_\varphi(B^3, S^2)$ there exists a sequence $(v_k)_{k\in\mathbb{N}}$ of maps in $\mathcal{C}^1_\varphi(B^3, S^2)$ which converges *weakly* in $W^{1,2}$ to u, we can define the **relaxed energy** E^{rel}_Ω on $W^{1,2}_\varphi(B^3, S^2)$ by

$$E^{rel}_{B^3} := \inf\{\lim_{k\to\infty} \inf \int_{B^3} |dv_k|^2 dx^1dx^2dx^3 \mid v_k \in \mathcal{C}^1_\varphi(B^3, S^2), v_k \to u \text{ weakly in } W^{1,2}\}.$$

The following expression for $E^{rel}_{B^3}$, valid when the degree of φ is zero, was given by F. Bethuel, H. Brezis and J.-M. Coron [15]:

$$E^{rel}_{B^3}(u) = E_{B^3}(u) + 4\pi L(u),$$

where $L(u)$ is length of the the minimal connection associated to u defined by (54). A nice theory was built by M. Giaquinta, G. Modica and J. Souček [88] in order to picture geometrically the relaxed energy and, more generally many bubbling phenomena[36]. The relaxed energy satisfies the properties (i) $\forall\, u \in W^{1,2}_\varphi(B^3, S^2)$, *$E^{rel}_{B^3}(u) \geq E_{B^3}(u)$, with equality if* $u \in \mathcal{C}^1_\varphi(B^3, S^2)$; (ii) $\inf_{u\in W^{1,2}_\varphi(B^3,S^2)} E^{rel}_{B^3}(u) = \inf_{u\in\mathcal{C}^1_\varphi(B^3,S^2)} E_{B^3}(u)$.

Other interesting functionals provided by interpolating between the Dirichlet energy E_{B^3} and the relaxed energy $E^{rel}_{B^3}$ were considered in [15]: for $\lambda \in \mathbb{R}$ consider $E^\lambda_{B^3}(u) = E_{B^3}(u) + 4\lambda\pi L(u)$. Then first of all, $\forall \lambda \in \mathbb{R}$, *the critical points of* $E^\lambda_{B^3}$ *on* $W^{1,2}(B^3, S^2)$

[36]The basic idea is to represent a map u between manifolds by its graph, which, in the case that u is in a Sobolev space but not continuous, is a *Cartesian current*, i.e. a current in the sense of geometric measure theory which satisfies some special conditions. In the enlarged class of Cartesian currents, we can describe precisely what the weak limit of a minimizing sequence is, keeping track of the *necks* connecting the bubbles in two dimensions, or the *minimal connection* in three dimensions. See [88] for a complete exposition.

are weakly harmonic. Second, *for* $0 \leq \lambda \leq 1$, $E^{\lambda}_{B^3}$ *is lower semi-continuous*. This implies that, for $0 \leq \lambda \leq 1$, the direct method can be used successfully in order to minimize $E^{\lambda}_{B^3}$ in, say, $W^{1,2}_{\varphi}(B^3, S^2)$ in order to obtain a family of weakly harmonic maps with the same boundary conditions (see the §4.3 for partial regularity results). This shows the strong *non-uniqueness of solutions* for the Dirichlet problem for harmonic maps in dimensions larger than three.

Minimizing the energy among continuous maps

In view of properties (i) and (ii) of the relaxed energy functional $E^{rel}_{B^3}$, it is tempting to use it in order to answer the following question: *given smooth boundary data* $\varphi : \partial B^3 \longrightarrow S^2$ *of degree 0, is there a smooth harmonic map* $B^3 \longrightarrow S^2$ *extending* φ *?* One strategy might be to minimize $E^{rel}_{B^3}$ over $W^{1,2}_{\varphi}(B^3, S^2)$: if we think, for example, of boundary data φ leading to a gap phenomenon described before, and we compare the values of the *relaxed energy* for the smooth and for the singular maps that we can construct, we realize that the gain in energy from allowing singularities is exactly cancelled by the cost due to the length of the minimal connection. But these considerations are only heuristic up to now: for the moment the question of whether minimizers of the relaxed energy are smooth is completely open (see §4.3).

On the other hand a direct approach to the problem of *minimizing the energy functional* $E_{\mathcal{M}}$ *in a class* $\mathcal{E}$ *of smooth maps in a given homotopy class between two arbitrary compact manifolds without boundary* $\mathcal{M}$ *and* $\mathcal{N}$ has been addressed by F. H. Lin [148]. He proved that *if* $(u_k)_{k\in\mathbb{N}}$ *is a minimizing sequence in* $\mathcal{E}$, *then, after extracting a subsequence if necessary,* u_k *converges weakly in* $W^{1,2}(\mathcal{M},\mathcal{N})$ *to a weakly harmonic map* $u \in W^{1,2}(\mathcal{M},\mathcal{N})$ *and* $|du_k|^2 d^m x$ *converges weakly to the Radon measure* $\mu = |du|^2 + \nu$. *Moreover,* u *is smooth away from a closed, rectifiable set* Σ *and* $\mathcal{H}^{m-2}(\Sigma)$ *is bounded.* The non-negative Radon measure ν measures the defect of strong convergence: it is the product of the $(m-2)$-dimensional Hausdorff measure supported by Σ times a function Θ on Σ which is measurable with respect to the $(m-2)$-dimensional Hausdorff measure. Lastly, **for almost all** $x \in \Sigma$, $\Theta(x)$ **is equal to a finite sum of energies of harmonic non-constant maps from** S^2 **to** $\mathcal{N}$, so that he obtain a higher-dimensional analogue of the results of Sacks and Uhlenbeck discussed in §5.3. Compare also with the results on the reduction of the singular set of a stationary map by Lin and Rivière presented in §4.3.

Towards completely discontinuous weakly harmonic maps

A notion of *relative relaxed energy* was introduced by F. Bethuel, H. Brezis and J.-M. Coron [15] as follows. Again, we fix smooth boundary data $\varphi : \partial B^3 \longrightarrow S^2$ of degree zero and we first define our functional on $\mathcal{R}^{2,1}_{\varphi}(B^3, S^2)$, the set of maps $u \in W^{1,2}_{\varphi}(B^3, S^2)$ which are $\mathcal{C}^1$ outside a finite number of points (see §3.2). For a pair (u, v) of maps in $\mathcal{R}^{2,1}_{\varphi}(B^3, S^2)$ we define the *length* $L(u, v)$ *of the minimal connection of* u *relative to* v to be the length of the minimal connection connecting the singularities of u and the singularities of v, where the singularities of v are *counted with opposite degrees*. By using the definition of the length of a minimal connection given by the right-hand side of (54), $L(u, v)$ can be expressed as

$$L(u, v) = \max_{\zeta \in Lip(\Omega),\, |\nabla\zeta|_{L^\infty} \leq 1} \int_{\partial B^3} d\zeta \wedge (u^*\omega_{S^2} - v^*\omega_{S^2})\,. \tag{57}$$

Thanks to (57), the functional $L : \mathcal{R}^{2,1}_\varphi(B^3, S^2) \times \mathcal{R}^{2,1}_\varphi(B^3, S^2) \longrightarrow \mathbb{R}$ can be extended to a functional $L : W^{1,2}_\varphi(B^3, S^2) \times W^{1,2}_\varphi(B^3, S^2) \longrightarrow \mathbb{R}$. It is shown in [15] that this functional is continuous on $W^{1,2}_\varphi(B^3, S^2) \times W^{1,2}_\varphi(B^3, S^2)$ and that, for any fixed $v \in W^{1,2}_\varphi(B^3, S^2)$, the functional

$$F_{B^3,v}(u) := E_{B^3}(u) + 4\pi L(u, v)$$

is lower semi-continuous on $W^{1,2}_\varphi(B^3, S^2)$. Moreover, the critical points of $F_{B^3,v}$ are weakly harmonic. This has turned out to be a powerful tool for constructing singular weakly harmonic maps.

First, R. Hardt, F. H. Lin and C. Poon [109] constructed weakly harmonic maps with a finite, but arbitrary, number of prescribed singularities located on a line. In their construction, they first fix a map $v \in \mathcal{R}^{2,1}_\varphi(B^3, S^2)$ which is invariant by rotations around some axis and which has dipoles of singularities along the axis of symmetry. Then they minimize the relative relaxed energy $F_{B^3,v}$ among all maps $u \in W^{1,2}_\varphi(B^3, S^2)$ which are also rotationally symmetric, and they show that the singular set of the minimizer is the same as the singular set of v. This result was improved by F. Rivière [181] who considered a sequence $(v_k)_{k\in\mathbb{N}^*}$ of rotationally symmetric maps in $W^{1,2}_\varphi(B^3, S^2)$ having more and more singularities along the axis of symmetry and the corresponding sequence $(u_k)_{k\in\mathbb{N}^*}$ of minimizers for F_{B^3,v_k} among rotationally symmetric maps in $W^{1,2}_\varphi(B^3, S^2)$. He was able to prove that $(u_k)_{k\in\mathbb{N}^*}$ converges to a weakly harmonic map having a line of singularity. Lastly Rivière [182] proved that, **for any non-constant map $\varphi : \partial B^3 \longrightarrow S^2$, there exists a weakly harmonic map in $W^{1,2}_\varphi(B^3, S^2)$ which is discontinuous everywhere in B^3**. This result rests on the *construction of a dipole* lemma: for any smooth map $w : B^3(a, r) \longrightarrow S^2$ such that $dw(a) \neq 0$ and for any $\rho \in (0, r)$ there exists a pair of points (p, n) inside $B^3(a, \rho)$ and a map $\widetilde{w} \in W^{1,2}(B^3(a, r), S^2)$ which is smooth outside $\{p, n\}$, has a degree 1 singularity at p and a degree -1 singularity at n, coincides with w in $B^3(a, r) \setminus B^3(a, \rho)$, and which satisfies

$$E_{B^3(a,r)}(\widetilde{w}) < E_{B^3(a,r)}(w) + 4\pi|p - n|\,. \tag{58}$$

That the inequality in (58) is strict[37] is crucial, as in the 2-dimensional theory (see §5.3). A second main ingredient in the proof of Rivière is the construction of a sequence $(v_k)_{k\in\mathbb{N}^*}$ of maps in $\mathcal{R}^{2,1}_\varphi(B^3, S^2)$ having more and more singularities. Each map v_{k+1} is constructed from v_k by adding a dipole and using the *construction of a dipole* lemma in order to control the extra cost of energy by (58). The sequence $(v_k)_{k\in\mathbb{N}^*}$ also converges strongly to some completely discontinuous map $v \in W^{1,2}_\varphi(B^3, S^2)$. The last task is then to show that any minimizer of $F_{B^3,v}$ is completely discontinuous.

5.5 Other analytical methods for existence

Morse and Lusternik–Schnirelman theories

A general reference for the ideas in this paragraph is the book of M. Struwe [209]. One of the first applications of these variational methods, devoted to existence proofs of *non-minimal* critical points is the work by G. D. Birkhoff [18] which establishes the existence of closed geodesics on a surface of genus 0, i.e. the image of a harmonic map of a circle,

[37] Note that a weaker, non-strict, analogous inequality was already obtained in [10].

see §2.2. Extensions to higher-dimensional harmonic maps is rather difficult and most of the known results concern the case $m = 2$.

In [188] J. Sacks and K. Uhlenbeck addressed the study of both *minimizing* (see §5.3) and *non-minimizing* harmonic maps from a surface without boundary $\mathcal{M}$ to a compact manifold without boundary $\mathcal{N}$. As for minimizing maps, they first establish the existence of non-minimizing critical points of the functional $E^\alpha_{\mathcal{M}}$ (see §5.3) for $\alpha > 1$, and then study the behaviour of these critical points when $\alpha \to 1$. The Morse theory for critical points of $E^\alpha_{\mathcal{M}}$ has better properties when $\alpha > 1$, since this functional then satisfies the *Palais–Smale condition*[38]. Let $\Omega(\mathcal{M}, \mathcal{N})$ be the space of base point preserving (continuous) maps from $\mathcal{M}$ to $\mathcal{N}$ (i.e. we fix some points $x_0 \in \mathcal{M}$ and $y_0 \in \mathcal{N}$ and we consider maps which send x_0 to y_0). First, Sacks and Uhlenbeck proved that *if $\Omega(\mathcal{M}, \mathcal{N})$ is not contractible, then there exist non-trivial critical points of $E^\alpha_{\mathcal{M}}$ between $\mathcal{M}$ and $\mathcal{N}$*. This critical point is non-minimizing if $\mathcal{C}^0(\mathcal{M}, \mathcal{N})$ is connected. They noticed that the hypothesis that $\Omega(\mathcal{M}, \mathcal{N})$ is not contractible is satisfied, in particular, *if $\mathcal{M} = S^2$ and if the universal cover of $\mathcal{N}$ is not contractible*, since $\pi_{k+2}(\mathcal{N}) = \pi_k(\Omega(S^2, \mathcal{N}))$. Second, they considered a sequence of maps from S^2 to $\mathcal{N}$ which are critical points of $E^\alpha_{\mathcal{M}}$ for $\alpha > 1$, and study its convergence as $\alpha \to 1$. The analysis is similar to the case of minimizing maps, see §5.3. They concluded that, **if the universal cover of $\mathcal{N}$ is not contractible, there exists a non-trivial harmonic map from S^2 to $\mathcal{N}$.** These results were extended by J. Jost in [133] using a different approach. Similar results have been obtained by Jost and Struwe [138], with applications to the Plateau problem for surfaces of arbitrary genus. See [134] for a survey and the papers by G. F. Wang [220] and Y. Ge [82] for recent applications to maps on a surface of genus greater than one with values in S^2.

These methods can also be applied on surfaces with boundary to construct non-minimizing harmonic maps with prescribed Dirichlet boundary condition. An example is the construction of saddle-point harmonic maps from the unit disc to the sphere S^n for $n \geq 3$ by V. Benci and J.-M. Coron [9]. This was extended to maps from a planar domain bounded by several disks by W. Y. Ding [54]. Similar results has been obtained by J. Qing [175] for maps from the unit disc to S^2.

Gauss maps of constant mean curvature surfaces

An important motivation for studying harmonic maps into spheres or, more generally, into a Grassmannian, is the result by E. A. Ruh and J. Vilms [187] on a submanifold Σ of dimension m immersed in the Euclidean space $\mathbb{R}^{m+p}$ and its Gauss map $f : \Sigma \longrightarrow G_m(m+p)$ to the Grassmannian of oriented m-dimensional subspaces of $\mathbb{R}^{m+p}$; this asserts that *the covariant derivative of the mean curvature vector field is equal to the tension field of its Gauss map*. In particular, *an immersion in $\mathbb{R}^{m+p}$ has parallel mean curvature if and only if its Gauss map is harmonic*. Note that, if $m = 2$ and $m + p = 3$, then $G_2(3) \simeq S^2$. The consequences of this fact are numerous[39]. For example, any construction of a mean curvature surface in $\mathbb{R}^3$ provides us with a harmonic map from that surfaces to S^2: constant mean curvature surfaces of genus 1 (tori) were first constructed by H. Wente [222]

[38] The Palais–Smale condition reads: *for any sequence of maps $(u_k)_{k\in\mathbb{N}}$ such that $E^\alpha_{\mathcal{M}}(u_k)$ is bounded and $\left(\delta E^\alpha_{\mathcal{M}}\right)_{u_k}$ converges to 0, there is a subsequence which converges strongly*, see [209].

[39] In particular, the structure of the completely integrable system for harmonic maps from a surface to S^2 and for constant mean curvature surfaces in $\mathbb{R}^3$ coincide *locally*, see Chapter 7.

by using a delicate analysis of the sinh–Gordon equation[40], later on N. Kapouleas [139] constructed higher-genus surfaces. The method here relies on gluing together pieces of explicitly known constant mean curvature surfaces (actually, segments of Delaunay surfaces) to produce, first, an approximate solution and then, by a careful use of a fixed point theorem, an exact solution near the approximate one. Since the work of Kapouleas, a huge variety of constructions has been done by following this strategy, see for example [140, 153].

A recent related result is the construction by P. Collin and H. Rosenberg [47] of *a harmonic diffeomorphism from the plane* $\mathbb{R}^2$ *onto the hyperbolic disc* H^2. Note that E. Heinz proved in 1952 that *there is no harmonic diffeomorphism from the hyperbolic disc* H^2 *onto the Euclidean plane* $\mathbb{R}^2$, and it was conjectured by R. Schoen that symmetrically there is should be no harmonic diffeomorphism from $\mathbb{R}^2$ to H^2 — the result of Collin and Rosenberg contradicts this conjecture. The proof relies on constructing an entire minimal graph in the product $H^2 \times \mathbb{R}$ which has the same conformal structure as $\mathbb{R}^2$. Hence, the harmonic diffeomorphism is the restriction to this graph of the projection mapping $H^2 \times \mathbb{R} \longrightarrow H^2$.

Ordinary differential equations

Many interesting examples of harmonic maps can be constructed by using reduction techniques. One powerful construction is the *join* of two *eigenmaps* of spheres introduced by R. T. Smith [202]: a map $u : S^m \longrightarrow S^n$ is called an *eigenmap* if and only if it is a harmonic map with a constant energy density; given two eigenmaps $u_1 : S^{m_1} \longrightarrow S^{n_1}$ and $u_2 : S^{m_2} \longrightarrow S^{n_2}$, and a function $\alpha : [0, \pi/2] \longrightarrow [0, \pi/2]$ such that $\alpha(0) = 0$ and $\alpha(\pi/2) = \pi/2$, the *α-join of* u_1 *and* u_2 is the map $u_1 *_\alpha u_2 \longrightarrow S^{m_1+m_2+1} \longrightarrow S^{n_1+n_2+1}$ defined by $(u_1 *_\alpha u_2)(x_1 \sin s\, ,\, x_2 \cos s) = (u_1(x_1) \sin \alpha(s)\, ,\, u_2(x_2) \cos \alpha(s))$. The harmonic map equation on $u_1 *_\alpha u_2$ reduces to an ordinary differential equation for α which can be solved in many cases [202, 64, 167]. A similar ansatz is the *α-Hopf construction* [179] $\varphi : S^{m_1+m_2+1} \longrightarrow S^{n+1}$ on a harmonic bi-eigenmap $f : S^{m_1} \times S^{m_2} \to S^n$: φ defined by $\varphi(x_1 \sin s\, ,\, x_2 \cos s) = (f(x_1, x_2) \sin \alpha(s)\, ,\, \cos \alpha(s))$. This construction leads also to a family of new examples [64, 56, 57, 81]. Similar reductions to systems of equations in more variables have been done [79, 80][41].

6 Other analytical properties

6.1 Uniqueness of and restrictions on harmonic maps

Uniqueness of harmonic maps in a given class of maps does not hold in general. The main case where uniqueness holds, with general methods to prove it, is when the target manifold satisfies strong convexity properties (see §6.3). An example of a result outside this situation requires the *smallness of the scaled energy* $E_{x,r}$ (see §4.3) for maps from $B^3 \subset \mathbb{R}^3$ to a compact manifold $\mathcal{N}$: *There exist some* $\varepsilon_0 > 0$ *and a constant* $C = C(\mathcal{N})$ *such that, for any boundary data* $g \in W^{1,2}(\partial B^3, \mathcal{N})$ *such that* $E_{\partial B^3}(g) < \varepsilon_0$*, there is a unique weakly harmonic map* $u \in W^{1,2}_g(B^3, \mathcal{N})$ *such that* $\sup_{x_0 \in B^3, r>0} \{r^{-1} \int_{B^3(x_0,r)\cap B^3} |du|^2 d^3x\} <$

[40]Since the work by Wente, a full classification of constant mean curvature tori has been obtained by using methods of completely integrable systems, see Chapter 7.

[41]Harmonic *morphisms* can also be found by this method, see [7, Chapter 13].

$C\varepsilon_0$. This was proved by M. Struwe [210] by using the regularity techniques for stationary maps in dimension greater than 2 (see §4.3).

Other restrictions on harmonic maps occur in the case where $\mathcal{M}$ is a surface without boundary and rely on methods of complex analysis (as in the result of Eells and Wood [67], see §5.3) or on the use of twistor theory for maps from the 2-sphere and integrable systems theory for maps from tori (see Chapter 7). See also the non-existence results for harmonic maps on a manifold with a non-empty boundary which are constant on the boundary [144, 230, 141] in §5.3.

6.2 Minimality of harmonic maps

A natural question is the following. *Consider a weakly harmonic map $u \in W^{1,2}(\mathcal{M},\mathcal{N})$; then is u an energy minimizer?* If the answer is yes, one of the most efficient methods to prove it is to combine results on existence, regularity and uniqueness. Many such results are available if $\mathcal{N}$ has good convexity properties; these are expounded in §6.3. Here is an example by R. Schoen and K. Uhlenbeck [192] of a result which can be proved without these convexity assumptions. *Let $S^n_+ := \{y \in S^n \subset \mathbb{R}^{n+1} |\ y^{n+1} > 0\}$ and $u : \mathcal{M} \longrightarrow S^n_+$ be a smooth harmonic map, then u is an energy minimizer among maps from $\mathcal{M}$ to S^n.* The proof proceeds as follows: let $\Omega \subset \mathcal{M}$ be any bounded domain with smooth boundary and apply the existence theorem of S. Hildebrandt, W. Jäger and K.-O. Widman [120] which asserts that there exists a *smooth* least energy map $\widetilde{u}$ from Ω to S^n_+ which agrees with u in $\partial\Omega$. Then, by the uniqueness result of W. Jäger and H. Kaul [126], we actually have $\widetilde{u} = u$ on Ω. Hence, u is energy minimizing among maps with values in S^n_+. Now let $v \in W^{1,2}(\Omega, S^n)$ be a map which agrees with u on $\partial\Omega$ and let $v_+ := \left(v^1, \cdots, v^n, |v^{n+1}|\right)$. We observe that $v_+ \in W^{1,2}(\Omega, S^n)$, v_+ agrees with u on $\partial\Omega$ and $E_{\mathcal{M}}(v_+) = E_{\mathcal{M}}(v)$. Actually v_+ takes values in the closure $\overline{S^n_+}$ of S^n_+, but it is easy to produce a continuous family $(R_\varepsilon)_{\varepsilon \leq 0}$ of retraction maps $R_\varepsilon : \overline{S^n_+} \longrightarrow \overline{S^n_+}$ such that $R_0 = \mathrm{Id}$, the image of R_ε is contained in S^n_+ if $\varepsilon > 0$, and $\lim_{\varepsilon\to 0} E_\Omega(R_\varepsilon \circ v_+) = E_\Omega(v_+)$. Moreover since $u(\overline{\Omega})$ is compact in S^n_+, we can construct R_ε in such a way that $R_\varepsilon \circ v_+$ agrees with u on $\partial\Omega$. Hence, $\forall \varepsilon > 0$, $E_\Omega(R_\varepsilon \circ v_+) \geq E_\Omega(u)$ which gives $E_\Omega(v) = E_\Omega(v_+) \geq E_\Omega(u)$ on letting $\varepsilon \longrightarrow 0$; the result follows. By similar reasoning, Jäger and Kaul [127] proved also that, if $u_\ominus \in W^{1,2}(B^m, S^m)$ is the map defined by $u_\ominus(x) = (x/|x|, 0)$, the minimum in $W^{1,2}_{u_\ominus}(B^m, S^m)$ is achieved by (i) a smooth rotationally symmetric diffeomorphism from B^m to $\overline{S^m_+}$ if $1 \leq m \leq 6$, (ii) $u_\ominus$ if $7 \leq m$.

Another favorable circumstance for proving the minimality of a harmonic map is if the harmonic map is a *diffeomorphism*. In dimension two the following result was proved by J.-M. Coron and F. Hélein [52]. *Let $(\mathcal{M}, g)$ and $(\mathcal{N}, h)$ be two Riemannian surfaces, then any harmonic diffeomorphism $\underline{u}$ between $(\mathcal{M}, g)$ and $(\mathcal{N}, h)$ is an energy minimizer among maps in the same homotopy class and (if $\partial\mathcal{M} \neq \emptyset$) with the same boundary conditions.* The idea is that, thanks to the Hopf differential of $\underline{u}$, one can construct an isometric embedding $(\mathcal{N}, h) \subset (\mathcal{M}, h_1) \times (\mathcal{M}, h_2)$ with two natural projections $\pi_a : (\mathcal{N}, h) \longrightarrow (\mathcal{M}, h_a)$ (for $a = 1, 2$) such that $\pi_1 \circ \underline{u}$ is harmonic conformal and hence a minimizer and $\pi_2 \circ \underline{u}$ is harmonic into $(\mathcal{M}, h_2)$. However the curvature of $(\mathcal{M}, h_2)$ is non-positive[42]. Thus $\pi_2 \circ \underline{u}$

[42]This argument does not work if $\mathcal{M} \simeq \mathcal{N} \simeq S^2$ but in this case any harmonic map is conformal and hence minimizing.

is also a minimizer thanks to results in [2, 111] (see §6.3). Moreover $\underline{u}$ is the *unique minimizer* if there exists a metric g_2 on $\mathcal{M}$ of negative curvature which is conformal to g [52]. Coron and Hélein also proved the minimality of some rotationally symmetric harmonic diffeomorphisms in dimension greater than two. These results were extended by Hélein [112, 115], by using *null Lagrangians*[43].

Because of the partial regularity theory of R. Schoen and K. Uhlenbeck [190] (see §4.3), it is important to identify the *homogeneous* maps u in $W^{1,2}(B^m, \mathcal{N})$ which are minimizing (recall that u is homogeneous if it is of the form $u(x) = \psi(x/|x|)$), since the *minimizing tangent maps*, which model the behaviour of a minimizing map near a singularity, are homogenous. Most known results concern the map $u_\odot^s \in W^{1,2}(B^m, S^{m-s-1})$ defined by $u_\odot^s(x, y) = x/|x|$, for $(x, y) \in \mathbb{R}^{m-s} \times \mathbb{R}^s$ (having an s-dimensional singular set) and, in particular, radial projection $u_\odot := u_\odot^0 \in W^{1,2}(B^m, S^{m-1})$: **for any** $m \geq 3$ **and for any** $s \geq 0$**,** $u_\odot^s$ **is a minimizer.** Various proofs exist, depending on the values of m and s:

- for $s = 0$ and $m \geq 7$ by Jäger and Kaul [127], as a corollary of the previous results on $u_\ominus$;
- for $s = 0$ and $m = 3$ by H. Brezis, J.-M. Coron and E. H. Lieb [24] (see §4.3) and $u_\odot$ is the *unique* minimizer;
- for $s = 0$ and $m \geq 3$ by F. H. Lin [146];
- for $s \geq 0$ and $m \geq 3$ by J.-M. Coron and R. Gulliver [51] (the general case).

The method of Lin is very short and uses a comparison of the energy functional $E_{B^m}(u)$, for $u \in W^{1,2}_{u_\odot}(B^m, S^{m-1})$, with another functional $F(u) := \int_{B^m} u^*(d\beta) = \int_{B^m} d(u^*\beta)$, where β is the $(m-1)$-form on $B^m \times \mathbb{R}^3$ defined by $\beta := \sum_{1 \leq i < j \leq m} (-1)^{i+j+1} (y^i dy^j - y^j dy^i) \wedge dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge \widehat{dx^j} \wedge \cdots dx^m$. Write $u^*(d\beta) = \lambda(du)\, dx^1 \wedge \cdots \wedge dx^m$. First, *from the fact that* u *takes values in* S^{m-1} *a.e.*, we show that $\lambda(du) \leq (m-2)|du|^2$ a.e., with equality if $u = u_\odot$. Second, we obtain from Stokes' theorem,

$$\begin{aligned} 2(m-2)E_{B^m}(u) \geq \int_{B^m} d(u^*\beta) &= \int_{\partial B^m} u^*\beta \\ &= \int_{\partial B^m} u_\odot^*\beta = \int_{B^m} d(u_\odot^*\beta) = 2(m-2)E_{B^m}(u_\odot). \end{aligned}$$

The functional $\int_{B^m} u^*(d\beta)$ is an example of a *null Lagrangian*. Lin's method is similar to the use of *calibrations* for minimal surfaces and to the argument used in equation (6) for harmonic functions. The proof of Coron and Gulliver uses two ingredients: (i) a representation of the energy of a map u by an integral over the Grassmannian manifold $G_3(\mathbb{R}^{m-s})$ of 3-planes Y in $\mathbb{R}^{m-s}$ of the energies of $\pi_Y \circ u \in W^{1,2}(B^m, S^2_Y)$, where $\pi_Y : S^{m-s-1} \longrightarrow S^{m-s-1} \cap Y := S^2_Y$ is the natural 'radial' projection and (ii) the coarea formula [44]. They also studied the maps $h_\odot^{\mathbb{C}} \in W^{1,2}(B^4, S^2)$ and $h_\odot^{\mathcal{H}} \in W^{1,2}(B^8, S^4)$

[43] The results by Coron and Hélein [52] use methods inspired from the work of Coron and R. Gulliver [51], whereas the use of null Lagrangians for harmonic maps was introduced by F. H. Lin [146], see below.

[44] See footnote 34. A similar method was used by Hélein in his thesis for proving: let $\phi : D \to \mathcal{N}$ be a submersive harmonic morphism with connected fibres from a compact domain of $\mathbb{R}^3$ with smooth boundary to a Riemann surface, then ϕ is the unique energy minimizer amongst maps with the same boundary values. See also [52].

defined by $h^{\mathbb{C}}_{\odot}(x) = H^{\mathbb{C}}(x/|x|)$ and $h^{\mathcal{H}}_{\odot}(x) = H^{\mathcal{H}}(x/|x|)$, where $H^{\mathbb{C}} : S^3 \to S^2$ and $H^{\mathcal{H}} : S^7 \to S^4$ are the *complex* and *quaternionic Hopf fibrations* (see §2.3), respectively, and they proved by similar methods that $h^{\mathbb{C}}_{\odot}$ and $h^{\mathcal{H}}_{\odot}$ are minimizing.

6.3 Analytic properties according to the geometric structure of $\mathcal{N}$

The target manifold $(\mathcal{N}, h)$ has non-positive Riemannian curvature

In this case, the harmonic map problem has many good convexity properties.

Existence. The first existence result was obtained by J. Eells and J. Sampson [66], and S.I. Al'ber [1] independently through the study of the heat equation $\partial\phi/\partial t = \tau(\phi)$ for a map $\phi : [0, \infty) \times \mathcal{M} \longrightarrow \mathcal{N}$, where $\partial\mathcal{M} = \emptyset$, with the Cauchy condition $\phi(0, \cdot) = \phi_0$ where $\phi_0 : \mathcal{M} \longrightarrow \mathcal{N}$ is a smooth map: if $\mathcal{M}$ and $\mathcal{N}$ are compact there always exists a finite time solution (i.e. defined on $[0, T] \times \mathcal{M}$), but *if* $(\mathcal{N}, h)$ *has non-positive curvature, this solution can be extended for all time*. Moreover the solution $\phi(t, \cdot)$ *converges*[45] *to a smooth harmonic map* $\underline{\phi}$ when $t \to +\infty$, which is *homotopic to* ϕ_0. When the boundary $\partial\mathcal{M}$ is non-empty and a Dirichlet condition $\phi(t, \cdot) = g$ on $\partial\mathcal{M}$ is imposed, these results were extended by R. Hamilton [99]. The existence conclusion can be recovered by using the Leray–Schauder degree theory [119], the maximum principle [130] or the direct method (see [216, 189] and §5.2).

Regularity. *Weakly harmonic maps into a non-positively curved manifold are smooth* and, moreover, the existence of convex functions on $\mathcal{N}$ allows *higher regularity estimates*: these are consequences of more general results, see §4.1 and below.

Minimality. The harmonic map $\underline{\phi}$ constructed in [66, 1] or [99] is actually energy minimizing [2, 111]. This follows by using the first and the second variation formulae for $E_{\mathcal{M}}$ given in [66]; this implies, in particular, the following identity [2]: let $\phi, \phi_0 : \mathcal{M} \longrightarrow \mathcal{N}$ be two smooth maps, and let $\Phi : [0, 1] \times \mathcal{M} \longrightarrow \mathcal{N}$ be a *geodesic homotopy between* ϕ_0 *and* ϕ, i.e. a smooth homotopy such that $\Phi(0, \cdot) = \phi_0$ and $\Phi(1, \cdot) = \phi$ and, for each fixed $x \in \mathcal{M}$, $s \longmapsto \Phi(s, x)$ is a geodesic; then, if ϕ_0 is harmonic we have

$$E_{\mathcal{M}}(\phi) - E_{\mathcal{M}}(\phi_0) = \int_0^1 d\sigma \int_0^\sigma ds \int_{\mathcal{M}} \Big\{ |\nabla_{\frac{\partial}{\partial s}} d\phi|^2 - g^{ij}\, {}^h R_{\alpha\beta\gamma\delta}(\phi) \frac{\partial \phi^\alpha}{\partial s} \phi_i^\beta \frac{\partial \phi^\gamma}{\partial s} \phi_j^\delta \Big\} \omega_g. \tag{59}$$

Hence if ${}^h R$ is non-positive the right hand side is nonnegative and this implies that any harmonic map is the minimizer in its homotopy class.

Uniqueness. Actually, each homotopy class contains, in most cases,[46] **only one** harmonic map: this was shown by P. Hartman [111] and S.I. Al'ber [2] independently and can be deduced from (59), see also [189]. An alternative method is possible if $\partial\mathcal{M} \neq \emptyset$: if $\mathcal{N}$ is *simply connected*[47] we can use the squared distance function $d^2 : \mathcal{N} \times \mathcal{N} \longrightarrow [0, \infty)$, which is a *strictly convex function* [135], see below.

[45] Eells and Sampson [66] established that $\phi(t, \cdot)$ *sub*converges to $\underline{\phi}$, but Hartman [111] proved that it actually converges.

[46] If $\partial\mathcal{M} = \emptyset$ non uniqueness can occur in two cases: we may have two different *constant* harmonic maps or two different harmonic maps which parametrize the same geodesic.

[47] Then any pair of points $p, q \in \mathcal{N}$ can be joined by a unique geodesic [135], and so $(\mathcal{N}, h)$ is convex.

Other properties. The *Bochner identity for harmonic maps* proved in [66],

$$\frac{1}{2}\Delta_g|d\phi|^2 = |\nabla d\phi|^2 - g^{ij}g^{kl}\,{}^hR_{\alpha\beta\gamma\delta}(\phi)\phi_i^\alpha\phi_k^\beta\phi_j^\gamma\phi_l^\delta + g^{ij}\,{}^gRic(\phi_i,\phi_j)\,, \tag{60}$$

is particularly useful if ${}^hR$ is non-positive and $\mathcal{M}$ is compact, since it then implies [66]:

$$-\Delta_g|d\phi|^2 \leq C|d\phi|^2, \quad \text{so, in particular } |d\phi|^2 \text{ is subharmonic.} \tag{61}$$

This inequality can be used to prove: (i) Liouville-type theorems [66]; (ii) the compactness in the $\mathcal{C}^k$-topology of the set $\mathcal{H}_\Lambda$ of maps $u \in \mathcal{C}^\infty(\mathcal{M},\mathcal{N})$ such that u is harmonic and $E_{\mathcal{M}}(u) < \Lambda$ (see [189]).

The target manifold has weaker convexity properties

The case when there exists a convex function on $\mathcal{N}$. Such functions are abundant on simply connected non-positively curved manifolds, but they also exist on any sufficiently small geodesic ball in $\mathcal{N}$. The basic observation is that *the composition of any harmonic map with a convex function is subharmonic and hence obeys the maximum principle* [124]. For instance, if the squared distance function $d^2 : \mathcal{N}\times\mathcal{N} \longrightarrow [0,\infty)$ exists and is *convex*, we can compose it with a pair $(u_0,u_1) : \mathcal{M} \longrightarrow \mathcal{N}\times\mathcal{N}$ of harmonic maps which agree on $\partial\mathcal{M} \neq \emptyset$ to prove the **uniqueness** result that $u_1 = u_2$, see, for example, [135]. Even more [86], if $g : \mathcal{N} \longrightarrow [0,\infty)$ is *bounded* and *strictly convex*, then for any $\mathcal{C}^2$ harmonic map $\phi : \mathcal{M} \longrightarrow \mathcal{N}$, we have

$$c_1|d\phi|^2 \leq \Delta(g\circ\phi), \quad \text{where } c_1 > 0. \tag{62}$$

Use of inequality (62) together with the monotonicity inequality (see §4.3) leads to the **local estimate** $\sup_{B(a,r/2)}|d\phi|^2 \leq Cr^{-n}\int_{B(a,r)}|d\phi|^2$, see [86, 189]. This can be used as the starting point for higher order estimates, see [86, 135].

The case when the image of ϕ **is contained in a geodesically convex ball**. The *optimal regularity result for weakly harmonic maps* with this kind of hypothesis is due to S. Hildebrandt, W. Jäger and K.-O. Widman [120]. We will say that a domain $\mathcal{B}\subset\mathcal{N}$ is an *HJW-convex ball* (after Hildebrandt, Jäger and Widman) if $\mathcal{B}$ is a geodesic ball $B(p_0,R)\subset\mathcal{N}$ (where $p_0\in\mathcal{N}$) such that

(i) *$\forall p \in B(p_0,R)$, the cut-locus of p does not intersect $B(p_0,R)$;*
(ii) *$R \leq 2\pi/\sqrt{\kappa}$, where $\kappa > 0$ is an upper bound of the Riemannian curvature on $B(p_0,R)$.*

Then Hildebrandt, Jäger and Widman proved *the existence of a solution to the Dirichlet problem with values in a HJW-convex ball* (see §5.2, f) and §6.2) and that **any weakly harmonic map** ϕ **with values in a HJW-convex ball is Hölder continuous** [120]. This result is optimal because of the following example: consider the map $u_\ominus \in W^{1,2}(B^m,\overline{S^m_+})$, where $\overline{S^m_+} := \{y\in S^m|\, y^{m+1}\geq 0\}$, defined by $u_\ominus(x) = (x/|x|,0)$, then, if $m\geq 3$ this maps has finite energy and is *weakly harmonic*. However $u_\ominus$ is clearly singular, but the hypothesis (i) of the above theorem is not satisfied.

With exactly the same hypothesis on the target, W. Jäger and H. Kaul in [126] found the following **uniqueness result**: assume that $\mathcal{M}$ is connected and $\partial\mathcal{M}\neq\emptyset$ and let ϕ_1,ϕ_2 :

$\mathcal{M} \longrightarrow \mathcal{B}$ be two smooth harmonic maps which agree on $\partial\mathcal{M}$; then, if $\mathcal{B}$ is a HKW-convex ball, $\phi_1 = \phi_2$. Again this result is optimal since, on the one hand, for $3 \leq m$, $W^{1,2}(B^m, \overline{S^m_+})$ contains the weakly harmonic map $u_\ominus$; on the other hand, for $3 \leq m \leq 6$, the minimum in $W^{1,2}_{u_\ominus}(B^m, \overline{S^m_+})$ is achieved by a smooth diffeomorphism onto $\overline{S^m_+}$, hence providing us with another harmonic map [127] (see [129] for improvements).

Influence of the topology of $\mathcal{N}$. Beyond more or less local assumptions on the curvature or the convexity of the target manifolds, many existence and regularity results are improved if one assumes that *there is no non-constant harmonic map from S^2 to $\mathcal{N}$*. This is related to the *bubbling phenomenon* which was discussed at length in §5.3 and 5.4.

7 Twistor theory and completely integrable systems

This is a rapid review of the development of the application of twistor theory and integrable systems to the study of harmonic maps. For further details, see, for example, [63, 94, 117, 73].

7.1 Twistor theory for harmonic maps

The genesis of the twistor theory for harmonic maps can be considered to be the following well-known result:[48] *Let $\phi : \mathcal{M}^2 \to \mathbb{R}^3$ be a conformal immersion from a Riemann surface $(\mathcal{M}^2, J^{\mathcal{M}})$. Then its Gauss map $\gamma : \mathcal{M}^2 \to S^2$ is antiholomorphic if and only if ϕ is harmonic (equivalently, minimal).*

The result was generalized to $\mathbb{R}^n$ by S.-S. Chern [44]. Indeed, let $\phi : \mathcal{M}^2 \to \mathbb{R}^n$ be a weakly conformal map. On identifying the Grassmannian $G_2^{\mathrm{or}}(\mathbb{R}^n)$ of oriented 2-planes in $\mathbb{R}^n$ with the complex quadric $Q_{n-2} = \{[z_1 : \cdots : z_n] \in \mathbb{C}P^n : z_1^2 + \ldots + z_n^2 = 0\}$, its Gauss map $\gamma : \mathcal{M}^2 \to G_2^{\mathrm{or}}(\mathbb{R}^n) = Q_{n-2}$ is given by the projective class of $\partial\phi/\partial\bar{z}$, where z is any local complex coordinate on $\mathcal{M}^2$. If ϕ is harmonic, γ is antiholomorphic by the harmonic equation, see (20). Note further that this antiholomorphicity implies that the Gauss map of a *weakly* conformal map extends smoothly across the set of branch points. Conversely, if γ is antiholomorphic, $\partial^2\phi/\partial z\partial\bar{z}$ is a multiple of the vector $\partial\phi/\partial\bar{z}$, which is tangential; but it is also a multiple of the mean curvature vector which is normal, thus it must vanish, hence ϕ is harmonic.

Now let $\mathcal{N} = \mathcal{N}^n$ be a general Riemannian manifold of dimension $n \geq 2$. Let $\pi : G_2^{\mathrm{or}}(\mathcal{N}) \to \mathcal{N}$ be the Grassmann bundle whose fibre at a point q of $\mathcal{N}^n$ is the Grassmannian of all oriented 2-dimensional subspaces of $T_q\mathcal{N}$. This is an associated bundle of the frame bundle $\mathrm{O}(\mathcal{N})$ of $\mathcal{N}$. Using the Levi-Civita connection, we may decompose the tangent bundle of $G_2^{\mathrm{or}}(\mathcal{N})$ into vertical and horizontal subbundles: $TG_2^{\mathrm{or}}(\mathcal{N}) = \mathcal{H} \oplus \mathcal{V}$; we denote the projections onto those subbundles by the same letters. Given any conformal immersion $\phi : \mathcal{M}^2 \to \mathcal{N}^n$, we define its *Gauss lift* $\gamma : \mathcal{M}^2 \to G_2^{\mathrm{or}}(\mathcal{N})$ by $\gamma(p) =$ the image of $d\phi_p$. Let $J^{\mathcal{V}}$ be the complex structure on the Grassmannian fibres of π. Say that γ is *vertically antiholomorphic* if

$$\mathcal{V} \circ d\gamma \circ J^{\mathcal{M}} = -J^{\mathcal{V}} \circ \mathcal{V} \circ d\gamma\,. \tag{63}$$

[48] This result is related to the Weierstrass–Enneper representation formula for a conformal parametrization $X : \Omega \subset \mathbb{C} \longrightarrow \mathbb{R}^3$ of a minimal surface in $\mathbb{R}^3$, which reads $X(z) = X(z_0) + \mathrm{Re}(\int_{z_0}^z (i(w^2-1),\, w^2 + 1,\, 2iw)(h/2)\,d\zeta)$, where w and h are respectively a meromorphic and a holomorphic function. Indeed, here w represents the Gauss map through an orientation reversing stereographic projection.

Then Chern's result extends to: *γ is vertically antiholomorphic if and only if ϕ is harmonic*. Further, the Gauss lift of a weakly conformal harmonic map extends smoothly over the branch points.

Maps into 4-dimensional manifolds. Suppose that $\mathcal{N} = \mathcal{N}^4$ is an oriented 4-dimensional Riemannian manifold. Then each $w \in G_2^{\rm or}(\mathcal{N}^4)$ defines an almost Hermitian structure J_w on $T_{\pi(w)}\mathcal{N}^4$. Further, if $\phi : \mathcal{M}^2 \to \mathcal{N}^4$ is a conformal immersion, then for any $p \in \mathcal{M}^2$, $d\phi_p$ intertwines $J_p^{\mathcal{M}}$ and $J_{\gamma(p)}$. Equivalently, lift J_w to an almost complex structure $J_w^{\mathcal{H}}$ on $\mathcal{H}_w$; then γ is *horizontally holomorphic* in the sense that

$$\mathcal{H} \circ d\gamma \circ J^{\mathcal{M}} = J^{\mathcal{H}} \circ \mathcal{H} \circ d\gamma\,. \tag{64}$$

We now define two almost complex structures J^1 and J^2 on the manifold $G_2^{\rm or}(\mathcal{N}^4)$ by setting J_w^1 (resp. J_w^2) equal to $J_w^{\mathcal{H}}$ on $\mathcal{H}_w$ and $J_w^{\mathcal{V}}$ (resp. $-J_w^{\mathcal{V}}$) on $\mathcal{V}_w$. Then the results above translate into: *the Gauss lift of a smooth immersion is holomorphic with respect to J^2 if and only if the map is conformal and harmonic.*

In fact, the projection of a J^2-holomorphic map into $G_2^{\rm or}(\mathcal{N})$ is always harmonic. More generally, let (Z, J^Z) be an almost complex manifold. A submersion $\pi : Z \to \mathcal{N}^4$ is called a *twistor fibration (for harmonic maps, with twistor space Z)* if the projection $\pi \circ f$ of any holomorphic map f from a Riemann surface to (Z, J^Z) is harmonic. The Grassmann bundle provides such a twistor fibration; we now find other twistor fibrations.

The Grassmann bundle $G_2^{\rm or}(\mathcal{N}^4)$ can be written as the product of two other bundles as follows. For any even-dimensional Riemannian manifold $\mathcal{N}^{2n}$, let $J(\mathcal{N}) \to \mathcal{N}$ be *the bundle of almost Hermitian structures on $\mathcal{N}$*. This is an associated bundle of $\mathrm{O}(\mathcal{N})$; indeed $J(\mathcal{N}) = \mathrm{O}(\mathcal{N}) \times_{\mathrm{O}(2n)} J(\mathbb{R}^{2n})$ where $J(\mathbb{R}^{2n}) = \mathrm{O}(2n)/\mathrm{U}(n)$ is the space of orthogonal complex structures on $\mathbb{R}^{2n}$. When $\mathcal{N}$ is oriented, $J(\mathcal{N})$ is the disjoint union of $J^+(\mathcal{N})$ and $J^-(\mathcal{N})$, *the bundles of positive* and *negative almost Hermitian structures on $\mathcal{N}^4$*, respectively. Give these bundles almost complex structures J^1 and J^2 in the same way as for $G_2^{\rm or}(\mathcal{N}^4)$. Then, when $\mathcal{N}$ is 4-dimensional, we have a bundle isomorphism $G_2^{\rm or}(\mathcal{N}^4) \to J^+(\mathcal{N}^4) \times J^-(\mathcal{N}^4)$ given by $w \mapsto (J_w^+, J_w^-)$ where J_w^+ (resp. J_w^-) is the unique almost Hermitian structure which is rotation by $+\pi/2$ on w. This isomorphism preserves J^1, J^2 and the horizontal spaces. The Gauss lift of an immersion $\phi : \mathcal{M}^2 \to \mathcal{N}^4$ thus decomposes into two *twistor lifts* $\psi_\pm : \mathcal{M}^2 \to J^\pm\mathcal{N}^4$. Both natural projections $J^\pm\mathcal{N}^4 \to \mathcal{N}$ are twistor fibrations; in fact we have the following result of J. Eells and S. Salamon [65]: **There is a bijective correspondence between non-constant weakly conformal harmonic maps $\phi : \mathcal{M}^2 \to \mathcal{N}^4$ and non-vertical J^2-holomorphic maps $\psi_\pm : \mathcal{M}^2 \to J^\pm\mathcal{N}^4$ given by setting $\psi_\pm$ equal to the twistor lift of ϕ.** For some related results in higher dimensions, see [180].

The problem with using this to find harmonic maps is that *J^2 is never integrable*. However, *J^1 is integrable if and only if the Riemannian manifold $\mathcal{N}^4$ is anti-selfdual*. Now a J^2-holomorphic map $\mathcal{M}^2 \to (Z, J^2)$ is also J^1-holomorphic if and only if it is *horizontal*, i.e., its differential has image in the horizontal subbundle $\mathcal{H}$, and horizontal holomorphic maps project to harmonic maps which are *real isotropic* in a sense that we now explain.

Real isotropic harmonic maps. A map $\phi : \mathcal{M}^2 \to \mathcal{N}^n$ from a Riemann surface to an arbitrary Riemannian manifold is called *real isotropic* if, for any complex coordinate z, all the derivatives $\nabla_Z^{\alpha}(\partial\phi/\partial z)$ lie in some isotropic subspace of $T^{\mathbb{C}}_{\phi(z)}\mathcal{N}$, i.e.

$$\eta_{\alpha,\beta} := \left\langle \nabla_Z^{\alpha}(\partial\phi/\partial z), \nabla_Z^{\beta}(\partial\phi/\partial z)\right\rangle = 0 \quad \text{for all } \alpha, \beta \in \{0, 1, 2, \ldots\}. \tag{65}$$

Here, $Z = \partial/\partial z$ and $\langle\ ,\ \rangle$ denotes the inner product on $T\mathcal{N}$ extended to $T^{\mathbb{C}}\mathcal{N}$ by complex bilinearity. For example, a holomorphic map to a Kähler manifold is real isotropic with the isotropic subspace being the $(1,0)$-tangent space. Now, in an extension to the argument showing that all harmonic maps from S^2 are weakly conformal (see §2.2), we show inductively on $k = \alpha + \beta$ that the inner products define holomorphic differentials $\eta_{\alpha,\beta}dz^k$ on S^2; since all holomorphic differentials on S^2 vanish for topological reasons, *all harmonic maps from S^2 to S^n are real isotropic*, and hence are obtained as projections of horizontal holomorphic maps into the twistor space. Such maps are easy to construct from 'totally isotropic' holomorphic maps into $\mathbb{C}P^n$ giving E. Calabi's theorem [36], as follows. Say that a map to a sphere or complex projective space is *full* if its image does not lie in a totally geodesic subsphere or projective subspace. Then **there is a** $2:1$ **correspondence between full harmonic maps** $\pm\phi : S^2 \to S^{2n}$ **and full totally isotropic holomorphic maps from** S^2 **to** $\mathbb{C}P^n$.

For an arbitrary oriented Riemannian manifold $\mathcal{N}$ of even dimension $2n$ greater than four, J^1 is integrable on $J^{\pm}(\mathcal{N})$ if and only if $\mathcal{N}$ is conformally flat. In order to apply twistor theory to more general manifolds, we need to find *reduced twistor spaces* on which J^1 is integrable. To do this, let $K \subset \mathrm{O}(2n)$ be the holonomy group of $\mathcal{N}$ and $\mathcal{P} \to \mathcal{N}$ the corresponding *holonomy bundle* given by reducing the structure group of $\mathrm{O}(\mathcal{N})$ to K. Then $J(\mathcal{N}) = \mathcal{P} \times_K J(\mathbb{R}^{2n})$. The holonomy group K acts on $J(\mathbb{R}^{2n})$ by conjugation, decomposing it into orbits O_i; it thus acts on $J(\mathcal{N})$, decomposing it into the union of subbundles associated to $\mathcal{P}$ and having fibre one of the orbits O_i. These subbundles are the candidates for our reduced twistor spaces.

For example, if $\mathcal{N}$ is a generic Kähler n-manifold, $K = U(n)$ and we find that the complex $U(n)$-orbits of $J(\mathcal{N})$ can be identified with the Grassmann bundles $G_r(T^{1,0}\mathcal{N}) \to \mathcal{N}$ $(r = 0, \ldots, n)$. These are thus twistor fibrations for harmonic maps. Note that, *for $0 < r < n$, J^1 is integrable on $G_r(T^{1,0}\mathcal{N})$ if and only if the Bochner tensor of $\mathcal{N}$ vanishes.*

Complex isotropic harmonic maps. Horizontal holomorphic maps into the Grassmann bundle project to harmonic maps which are *complex isotropic* in the sense that all the covariant derivatives $\nabla_Z^{\alpha}(\partial^{1,0}\phi/\partial z)$ are orthogonal in $T'_{\phi(z)}\mathcal{N}$ to all the covariant derivatives $\nabla_{\overline{Z}}^{\alpha}(\partial^{1,0}\phi/\partial\overline{z})$ with respect to the Hermitian inner product on $T'\mathcal{N}$. In particular, when $\mathcal{N} = \mathbb{C}P^n$, an argument again involving the holomorphicity of differentials constructed from the above inner products shows that all harmonic maps from $S^2 \to \mathbb{C}P^n$ are complex isotropic, and so given by such projections. In this case we can explicitly identify the Grassmann bundles and construct all holomorphic horizontal maps into it from holomorphic maps $S^2 \to \mathbb{C}P^n$ by considering their iterated derivatives. This leads to the result [68]: **There is a one-to-one correspondence between pairs** (f, r) **where** f **is a full holomorphic map from** S^2 **to** $\mathbb{C}P^n$ **and** $r \in \{0, 1, \ldots, n\}$ **and full harmonic maps from** S^2 **to** $\mathbb{C}P^n$.

Maps into symmetric spaces. Now let G be a compact Lie group and $\mathcal{N}^{2n} = G/K$ an irreducible Riemannian symmetric space. Then the natural projection $G \to G/K = \mathcal{N}$ is a reduction of the frame bundle with structure group K. As above, K acts on $J(\mathbb{R}^{2n})$

and thence on $J(\mathcal{N}) = G \times_K J(\mathbb{R}^{2n})$. Any orbit in $J(\mathbb{R}^{2n})$ is of the form K/H for some closed subgroup H; the corresponding orbit in $J(\mathcal{N})$ is the subbundle $\pi : G \times_K K/H \cong G/H \to G/K$ where π is the natural projection. This subbundle can alternatively be thought of an orbit of the action of G on $J(\mathcal{N})$. F. E. Burstall and J. H. Rawnsley [35] showed that such an orbit is almost complex manifold on which J^1 is integrable if and only if is contained in the zero set of the Nijenhuis tensor of J^1, They go on to prove that, if $\mathcal{N} = G/K$ is an *inner* symmetric space[49] of compact type, that zero set consists of finitely many orbits of G with each orbit G/H a *flag manifold* of G and that every flag manifold of G occurs for some inner symmetric space G/K. Further, any flag manifold G/H can be written alternatively as $G^{\mathbb{C}}/P$ for some suitable parabolic subgroup of the complexified group $G^{\mathbb{C}}$, and so has a natural complex structure J^1. On replacing J^1 by $-J^1$ on the fibres, we obtain a non-integrable almost complex structure J^2 and then *the natural projection* $(G/H, J^2) \to G/K = \mathcal{N}$ *is a twistor fibration for harmonic maps*. Further every harmonic map from S^2 to $\mathcal{N}$ is the projection of some J^2-holomorphic map into a suitable flag manifold. Moreover Burstall and Rawnsley exhibit holomorphic differentials;[50]if these vanish then the J^2-holomorphic curve is in fact holomorphic for the complex structure J^1. For the special case of *isotropic* harmonic maps, see below.

7.2 Loop group formulations

Again let G be a compact Lie group, and let ω be its (left) Maurer–Cartan form; this is a 1-form with values in the Lie algebra $\mathfrak{g}$ of G which satisfies the Maurer–Cartan equation $d\omega + \frac{1}{2}[\omega \wedge \omega] = 0$ where $[\omega \wedge \omega](X, Y) = 2[\omega(X), \omega(Y)]$ $(X, Y \in T_\gamma G, \gamma \in G)$. Note that ω gives an explicit trivialization $TG \cong G \times \mathfrak{g}$ of the tangent bundle; the Maurer–Cartan equation expresses the condition that the connection $d + \omega$ on this bundle is flat.

Maps into Lie groups. Now let $\phi : \mathcal{M}^m \to G$ be a smooth map from a Riemannian manifold to G. Let A be the $\mathfrak{g}$-valued 1-form given by the pull-back $\phi^*\omega$. Then A represents the differential $d\phi$; indeed, if G is a matrix group, $A = \phi^{-1}d\phi$. Pulling back the Maurer–Cartan equation shows that A satisfies

$$dA + \frac{1}{2}[A \wedge A] = 0. \tag{66}$$

This equation is an *integrability condition*: given a $\mathfrak{g}$-valued 1-form, we can find a smooth map $\phi : \mathcal{M} \to G$ with $A = \phi^{-1}d\phi$ if and only if (66) is satisfied. Further, it is easy to see that ϕ is harmonic if and only if

$$d^*A = 0\,. \tag{67}$$

Now let $\mathcal{M}^2$ be a simply connected Riemann surface and let (U, z) be a complex chart. Writing $A = A_z dz + A_{\bar z} d\bar z$ we may add and subtract the equations (66,67) to obtain the equivalent pair of equations:

$$\frac{\partial A_z}{\partial \bar z} + \frac{1}{2}[A_{\bar z}, A_z] = 0\,, \qquad \frac{\partial A_{\bar z}}{\partial z} + \frac{1}{2}[A_z, A_{\bar z}] = 0\,. \tag{68}$$

[49] An inner symmetric space is a Riemannian symmetric space whose involution is inner.

[50] These differentials vanish for harmonic maps from S^2 to S^{2n}, $\mathbb{C}P^n$ and $S^4 \simeq \mathcal{H}P^1$, so that one recovers the previous classification results for such maps [36, 28, 68].

We now introduce a parameter $\lambda \in S^1 := \{\lambda \in \mathbb{C}^* | \, |\lambda| = 1\}$ (called the *spectral parameter*), and consider the loop of 1-forms:

$$A_\lambda = \frac{1}{2}(1-\lambda^{-1})A_z dz + \frac{1}{2}(1-\lambda)A_{\bar z} d\bar z\,. \tag{69}$$

K. Uhlenbeck noticed[51] [218] that A satisfies the pair (66,67) if and only if

$$dA_\lambda + \frac{1}{2}[A_\lambda \wedge A_\lambda] = 0 \quad \text{for all } \lambda \in S^1; \tag{70}$$

this equation is a *zero curvature equation*: it says that for each λ, $d + A_\lambda$ is a flat connection on $\mathcal{M} \times \mathfrak{g}$. If satisfied, there is a loop of maps E_λ on $\mathcal{M}$ satisfying $E_\lambda^*(\omega) = A_\lambda$, since $\mathcal{M}$ is simply connected; equivalently, a map $\mathcal{E} : \mathcal{M} \to \Omega G$ into the *(based) loop group* of G: $\Omega G = \{\gamma : S^1 \to G \mid \gamma(1) = \text{identity of } G\}$ (where the loops γ satisfy some regularity assumption such as C^∞).

The map $\mathcal{E} : \mathcal{M} \to \Omega G$ is called the *extended solution* corresponding to ϕ. Now suppose that G is a matrix group, i.e., $G \subset GL(\mathbb{R}^N) \subset \mathbb{R}^{N\times N}$. It can be written as Fourier series

$$\mathcal{E}(z) : \lambda \longmapsto E_\lambda(z) = \sum_{i=-\infty}^{\infty} \lambda^i \widehat{E}_i(z) \qquad (z \in \mathcal{M})$$

for some maps $\widehat{E}_i : \mathcal{M} \to G$. If this is a finite series, we say that ϕ has *finite uniton number*. Uhlenbeck showed that *all harmonic maps from S^2 to the unitary group (and so to all compact groups) have finite uniton number*. She also gave a *Bäcklund-type transform* which gives new harmonic maps from old ones by multiplying their extended solution by a suitable linear factor called *uniton*, and showed that the extended solution of a harmonic map $\phi : S^2 \to \mathrm{U}(n)$ can be *factorized* as the product of unitons, so that ϕ can be obtained from a constant map by *adding a uniton* no more than n times. Another proof was given by G. Segal [196] using a Grassmannian model of $\mathrm{U}(n)$. An extension of the factorization theorem to maps into most other compact groups G was proved by Burstall and Rawnsley [35].

We can also consider the 'free' loop group $\Lambda G = \{\gamma : S^1 \to G\}$ and we may define loop groups $\Omega G^{\mathbb{C}}$ and $\Lambda G^{\mathbb{C}}$ for the complexified group $G^{\mathbb{C}}$ in the same way. Let $\Lambda^+ G^{\mathbb{C}}$ be the subgroup of loops which extend holomorphically to the disk $D^2 := \{\lambda \in \mathbb{C} | \, |\lambda| < 1\}$, i.e., have Fourier coefficients $\widehat{\gamma}_i$ zero for negative i. Then, we have an *Iwasawa decomposition* $\Lambda G^{\mathbb{C}} = \Omega G \cdot \Lambda^+ G^{\mathbb{C}}$ so that we can write ΩG as $\Lambda G^{\mathbb{C}}/\Lambda^+ G^{\mathbb{C}}$; this gives ΩG a *complex structure*. Now (69) tells us that the partial derivative $\mathcal{E}_{\bar z}$ lies in $\Lambda^+ \mathfrak{g}^{\mathbb{C}}$ which means that $\mathcal{E}$ is holomorphic. Further $\mathcal{E}_z$ lies in the subspace of $\Lambda \mathfrak{g}^{\mathbb{C}}$ where all Fourier coefficients other than A_{-1} and A_0 are zero; we say that $\mathcal{E}$ is *superhorizontal*. Thus we can interpret the fibration $\pi : \Omega G \longrightarrow G$ given by $\mathcal{E} \longmapsto \mathcal{E}|_{\lambda=-1}$ as a twistor fibration, since *any* harmonic map from $\mathcal{M}$ to G is the image by π of a holomorphic horizontal curve in ΩG.

Maps into Riemannian symmetric spaces. We can apply the above to harmonic maps into symmetric spaces G/K by including G/K by the totally geodesic Cartan embedding $\iota : G/K \longrightarrow G$ defined by $\iota(g \cdot K) = \tau(g)g^{-1}$, where $\tau : G \longrightarrow G$ is the Cartan involution[52] such that $(G^\tau)_0 \subset K \subset G^\tau$; here $G^\tau := \{g \in G | \, \tau(g) = g\}$ and $(G^\tau)_0$ is the

[51] Uhlenbeck's discovery was known previously to several physicists, see for example [172].

[52] $\tau(g) = s_o g s_o^{-1}$ where s_o is the point reflection in the base point of $\mathcal{N}$.

connected component of G^τ which contains the identity. However, there is an alternative more geometrical method which we now describe. For any map $\phi : \mathcal{M} \longrightarrow G/K$ choose a lift $f : \mathcal{M} \longrightarrow G$ of it and consider its Maurer–Cartan form $\alpha = f^*\omega \simeq f^{-1}df$. The Cartan involution τ induces a linear involution on the Lie algebra $\mathfrak{g} \simeq T_{\mathrm{Id}}G$ that we denote also by τ. The eigenvalues of τ are ± 1 and we have the eigenspace decomposition $\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}_1$, where, for $a = 0, 1$, $\mathfrak{g}_a$ is the $(-1)^a$-eigenspace. Note that $\mathfrak{g}_0 = \mathfrak{k}$ is the Lie algebra of K. Now we can split $\alpha = \alpha_0 + \alpha_1$ according to the eigenspace decomposition of $\mathfrak{g}$ and further split $\alpha_1 = \alpha_1' + \alpha_1''$, where $\alpha_1' := \alpha_1(\partial/\partial z)\, dz$ and $\alpha_1'' := \alpha_1(\partial/\partial \bar{z})\, d\bar{z}$. Then $\phi : \mathcal{M} \longrightarrow G/K$ *is harmonic if and only if, for all* $\lambda \in S^1$ *we have* $d\alpha_\lambda + (1/2)[\alpha_\lambda \wedge \alpha_\lambda] = 0$, where

$$\alpha_\lambda := \lambda^{-1}\alpha_1' + \alpha_0 + \lambda\alpha_1'' \quad \text{for all } \lambda \in S^1. \tag{71}$$

This relation allows us to construct a family of maps $f_\lambda : \mathcal{M} \longrightarrow G$ by integrating the relation $\alpha_\lambda = f_\lambda^*\omega \simeq f_\lambda^{-1}df_\lambda$. Each map f_λ lifts a *harmonic* map $\phi_\lambda : \mathcal{M} \longrightarrow G/K$ given by $\phi_\lambda(z) = f_\lambda(z)K$, hence $(\phi_\lambda)_{\lambda\in S^1}$ is an *associated family* of harmonic maps. Alternatively we can view the family $F = (f_\lambda)_{\lambda\in S^1}$ as a single map from $\mathcal{M}$ to the *twisted loop group* $\Lambda G_\tau := \{\gamma : S^1 \longrightarrow G |\ \gamma(-\lambda) = \tau(\gamma(\lambda))\}$ and the family $\Phi = (\phi_\lambda)_{\lambda\in S^1}$ as a map into $(\Lambda G_\tau)/K$. Given a harmonic map ϕ, the map Φ is unique if we assume for instance the extra condition $f_\lambda(z_0) = \mathrm{Id}$, for some $z_0 \in \mathcal{M}$. The representation of a harmonic map into G/K using twisted loop groups is related to the one using based loop groups through the relations $E_\lambda = f_\lambda f^{-1}$ and $\iota(\phi_\lambda) = \tau(f_\lambda)f_\lambda^{-1} = E_{-\lambda}E_\lambda^{-1}$.

A 'Weierstrass' representation. We denote the complexification of ΛG_τ by $\Lambda G_\tau^{\mathbb{C}}$. We also define $\Lambda^+ G_\tau^{\mathbb{C}}$ as the subgroup of loops $\gamma \in \Lambda G_\tau^{\mathbb{C}}$ which have a holomorphic extension (that we still denote by γ) in the disk D^2 and, if $\mathfrak{B} \subset G^{\mathbb{C}}$ is a solvable Borel subgroup such that the Iwasawa decomposition $G^{\mathbb{C}} = G \cdot \mathfrak{B}$ holds, we let $\Lambda^+_{\mathfrak{B}} G_\tau^{\mathbb{C}}$ be the subgroup of loops $\gamma \in \Lambda^+ G_\tau^{\mathbb{C}}$ such that $\gamma(0) \in \mathfrak{B}$. Now J. Dorfmeister, F. Pedit and H. Y. Wu [60] proved that an Iwasawa decomposition $\Lambda G_\tau^{\mathbb{C}} = \Lambda G_\tau \cdot \Lambda^+_{\mathfrak{B}} G_\tau^{\mathbb{C}}$ holds, so that we can define a natural fibration $\pi_\tau : \Lambda G_\tau^{\mathbb{C}} \longrightarrow \Lambda G_\tau^{\mathbb{C}}/\Lambda^+_{\mathfrak{B}} G_\tau^{\mathbb{C}} = \Lambda G_\tau$. They show also that if $H : \mathcal{M} \longrightarrow \Lambda G_\tau^{\mathbb{C}}$ is a *holomorphic curve* which satisfies the *superhorizontality condition* $\lambda H^*\omega \simeq \lambda H^{-1}dH \in \Lambda^+\mathfrak{g}^{\mathbb{C}}$, then $F = \pi_\tau \circ H$ (i.e., the unique map F into $\Lambda G_\tau^{\mathbb{C}}$ such that $H = FB$, for some map $B : \mathcal{M} \longrightarrow \Lambda^+_{\mathfrak{B}} G_\tau^{\mathbb{C}}$) lifts an *associated family of harmonic maps*. Conversely Dorfmeister, Pedit and Wu proved that any harmonic map from a simply connected surface to $\mathcal{N}$ arises that way. The superhorizontal holomorphic maps H which covers a given F are *not* unique. However we can use another *Birkhoff decomposition* $\Lambda G_\tau^{\mathbb{C}} \supset \mathcal{C} = \Lambda^-_* G_\tau^{\mathbb{C}} \cdot \Lambda^+ G_\tau^{\mathbb{C}}$, where $\Lambda^-_* G_\tau^{\mathbb{C}}$ is the subset of loops $\gamma \in \Lambda G_\tau^{\mathbb{C}}$ which have a holomorphic extension to $\mathbb{C}P^1 \setminus D^2 := \{\lambda \in \mathbb{C} \cup \{\infty\} |\ |\lambda| \geq 1\}$ and such that $\gamma(\infty) = \mathrm{Id}$. Here $\mathcal{C}$ is the *big cell*, a dense subset of the connected component of Id in $\Lambda G_\tau^{\mathbb{C}}$. Further Dorfmeister, Pedit and Wu showed that for any lift F of an associated family of harmonic maps into $\mathcal{N}$, there exist finitely many points $\{a_1, \cdots, a_k\}$ such that F takes values in $\mathcal{C}$ outside $\{a_1, \cdots, a_k\}$. We can hence decompose $F = F^- F^+$ on $\mathcal{M} \setminus \{a_1, \cdots, a_k\}$, where F^- (respectively F^+) takes values in $\Lambda^-_* G_\tau^{\mathbb{C}}$ (respectively $\Lambda^+ G_\tau^{\mathbb{C}}$), and then F^- extends to a meromorphic superhorizontal curve on $\mathcal{M}$ with poles at $a_1, \cdots, a_k$. Then the Maurer–Cartan form of F^-, $\mu = (F^-)^*\omega$, reads $\mu_\lambda = \lambda^{-1}\xi dz$, where $\xi : \mathcal{M} \longrightarrow \mathfrak{g}_1^{\mathbb{C}}$ is a meromorphic potential called the *meromorphic potential* of F. This provides *Weierstrass data* for the harmonic map and is known as the 'DPW' method [60].

Pluriharmonic maps. This can be extended to the more general case of 'pluriharmonic' maps: a smooth map from a complex manifold is called *pluriharmonic* if its restriction to every complex one-dimensional submanifold is harmonic. Let $\phi : (\mathcal{M}, J^{\mathcal{M}}) \to \mathcal{N}$ be a smooth map from a simply connected complex manifold to a Riemannian symmetric space $\mathcal{N} = G/K$. For $\lambda = e^{-i\theta} \in S^1$, define an endomorphism of $T\mathcal{M}$ by $R_\lambda = (\cos\theta)I + (\sin\theta)J$. Extending this by complex-linearity to the complexified tangent bundle $T^{\mathbb{C}}\mathcal{M}$, we have that $R_\lambda = \lambda^{-1}I$ on the $(1,0)$-tangent bundle $T'\mathcal{M}$ and $R_\lambda = \lambda I$ on the $(0,1)$-tangent bundle $T''\mathcal{M}$. Note that, if $\mathcal{M}$ is a Riemann surface, R_λ is rotation through θ. J. Dorfmeister and J.-H. Eschenburg [59] show that ϕ is pluriharmonic if and only if there is a parallel bundle isometry $\mathfrak{R}_\lambda : \phi^*T\mathcal{N} \to \phi_\lambda^*T\mathcal{N}$ preserving the curvature such that $\mathfrak{R}_\lambda \circ d\phi \circ R_\lambda = d\phi_\lambda$ for some smooth family of maps $\phi_\lambda \ (\lambda \in S^1)$, and that the maps ϕ_λ are all pluriharmonic; thus *pluriharmonic maps again come in associated S^1-families*. Then with similar definitions of superhorizontal and holomorphic to those above, we obtain the result: *there is a one-to-one correspondence between pluriharmonic maps $\phi : \mathcal{M} \to G/K$ and superhorizontal holomorphic maps $\Phi : \mathcal{M} \to \Lambda_\sigma G/K$ with $\phi = \pi \circ \Phi$.*

The twistor theory revisited. Twistor theory appears as a special case: a map is called *isotropic* if the associated family ϕ_λ is trivial, i.e. $\phi_\lambda = \phi$ up to congruence for all $\lambda \in S^1$. Then, for each $z \in \mathcal{M}$, the $\mathfrak{R}_\lambda(z)$ are automorphisms of $T_{\phi(z)}\mathcal{N}$, representing these by elements of G, they define a homomorphism $\mathfrak{R}(z) : S^1 \to G$, $\lambda \mapsto \mathfrak{R}_\lambda(z)$. By parallelity of the $\mathfrak{R}_\lambda(z)$ as z varies, these homomorphisms are all conjugate, so that the $\mathfrak{R}_\lambda$ define a map into the congugacy class of a circle subgroup $q : S^1 \to G$, $\lambda \mapsto q_\lambda$ with $q_{-1} = s_o$; this congugacy class is a flag manifold of the form G/C_q where C_q is the centralizer of q, and the $\mathfrak{R}_\lambda$ define a twistor lift into that manifold. Note that C_q is contained in K. Also a necessary condition for the existence of a circle subgroup q with $q_{-1} = s_o$ is that $\mathcal{N}$ be *inner*, i.e., s_o lies in the identity component of K. We thus obtain [69]: *Let $\phi : \mathcal{M} \to \mathcal{N}$ be a smooth map into an inner symmetric space $\mathcal{N} = G/K$ of compact type which is* full*, i.e., does not have image in a totally geodesic proper subspace of $\mathcal{N}$. Then ϕ is isotropic if and only if there is a flag manifold $Z = G/H$ with $H \subset K$ and a holomorphic superhorizontal map $\Phi : \mathcal{M} \to Z$ such that $\pi \circ \Phi = \phi$ where $\pi : G/H \to G/K$ is the natural projection.* In this setting, pluriharmonic maps into Lie groups G appear naturally by treating G as the symmetric space $G \times G/G$.

F. Burstall and M. A. Guest [34] take all this much further by showing that to every extended solution can be associated a homomorphism $q : \lambda \mapsto q_\lambda$ by flowing down the gradient lines of the energy of loops in G. The extended solution can be recovered from q by multiplication by a suitable holomorphic map into a loop group. The conditions (69) translate into conditions on the coefficients of the Fourier series of this map related to the eigenspace decomposition of $\mathrm{Ad}\, q_\lambda$. This leads to equations in the meromorphic parameters which can be solved by successive integrations leading to the theorem: *Every harmonic map $S^2 \to G$ arises from an extended solution which may be obtained explicitly by choosing a finite number of rational functions and then performing a finite number of algebraic operations and integrations.* They show how the work of Dorfmeister, Pedit and Wu [60] fits into this scheme, as well as Uhlenbeck's factorization.

Finite type solutions. An alternative way of finding harmonic maps into symmetric spaces, especially when the domain is a (2-)torus, is to integrate a pair of commuting Hamiltonian fields on the finite-dimensional subspace $\Omega^d\mathfrak{g} := \{\xi \in \Omega\mathfrak{g} |\ \xi_\lambda =$

$\sum_{k=-d}^{d}\widehat{\xi}_k(1-\lambda^k)\}$ of the based[53] loop algebra $\Omega\mathfrak{g}$, for some $d\in\mathbb{N}^*$. Indeed the vector fields X_1 and X_2 defined on $\Omega^d\mathfrak{g}$ by $X_1(\xi)-iX_2(\xi)=2[\xi,2i(1-\lambda)\widehat{\xi}_d]$ *are tangent to* $\Omega^d\mathfrak{g}$ *and commute*. Thus we can integrate the Lax type equation $d\xi=[\xi,2i(1-\lambda)\widehat{\xi}_d dz-2i(1-\lambda^{-1})\widehat{\xi}_{-d}d\overline{z}]$, where $\xi:\mathbb{R}^2\longrightarrow\Omega^d\mathfrak{g}$ (for a formulation of the harmonic map equations as a *Lax pair*, see the article by Wood in [73] or [53, 94]). Then, for any solution of this equation, the loop of 1-forms $A_\lambda:=2i(1-\lambda)\widehat{\xi}_d dz-2i(1-\lambda^{-1})\widehat{\xi}_{-d}d\overline{z}$ satisfies the relation (70) and hence provides an extended harmonic map by integrating the relation $E_\lambda^*\omega=A_\lambda$; the resulting harmonic maps are said to be of *finite type*.

A nontrivial result is that, for all $n\in\mathbb{N}^*$, **all non-isotropic harmonic maps from the torus to S^n or $\mathbb{C}P^n$ are of finite type.** This was proved by N. Hitchin [122] for tori in S^3, by U. Pinkall and I. Sterling [169] for constant mean curvature tori in $\mathbb{R}^3$ and by Burstall, D. Ferus, Pedit and Pinkall [33] for non-conformal tori in rank one symmetric spaces ([122] and [33] propose a different approach, see [123] for a comparison). The case of conformal but non-isotropic tori in S^n or $\mathbb{C}P^n$ requires the notion of *primitive* maps introduced by Burstall [32], i.e. maps with values in a k-symmetric space fibred over the target. See [160, 161] for further developments. To each finite type harmonic map of a torus can be associated a compact Riemann surface called its *spectral curve*, together with some data on it called *spectral data*. This leads to a representation using techniques from algebraic geometry, done by A. Bobenko [19] for constant mean curvature tori and by I. McIntosh [154] for harmonic tori in complex projective spaces.

Harmonic maps from a higher genus surface $\mathcal{M}$. They can, in principle, be found by the DPW method by investigating harmonic maps on the universal cover of $\mathcal{M}$, but this is hard to implement. Another possible approach investigated by Y. Ohnita and S. Udagawa [160] is to look for (finite type) pluriharmonic maps on the Jacobian variety $J(\mathcal{M})$ of $\mathcal{M}$ and compose them with the Abel map $\mathcal{M}\longrightarrow J(\mathcal{M})$.

Bibliography

[1] S. I. Al'ber: On n-dimensional problems in the calculus of variations in the large *Sov. Math. Dokl.* **5** (1964) 700–704

[2] S. I. Al'ber: Spaces of mappings into a manifold with negative curvature *Sov. Math. Dokl.* **9** (1967) 6–9

[3] F. Almgren, W. Browder and E. H. Lieb: Co-area, liquid crystals, and minimal surfaces *In: Partial differential equations* (Tianjin, 1986, Lecture Notes in Math., 1306, Springer, Berlin, 1988) 1–22

[4] F. Almgren and E. H. Lieb: Singularities of minimizing harmonic mappings from the ball to the sphere: examples, counterexamples and bounds *Ann. of Math., (2)* **128(3)** (1988) 483–530

[5] T. Aubin: Equations différentielles nonlinéaires et problème de Yamabe concernant la courbure scalaire *J. Maths. Pures et Appl.* **55** (1976) 269–296

[6] P. Baird and J. Eells: A conservation law for harmonic maps *In: Geometry* (Symposium Utrecht, 1980, Lecture Notes in Mathematics, vol. 894, Springer, Berlin, 1981) 1–25

[53] A similar formulation using twisted loops exists, see the paper by Burstall and Pedit in [73] or [117].

[7] P. Baird and J. C.Wood: *Harmonic Morphisms between Riemannian Manifolds* (London Math. Soc. Monogr. (N.S.), No. 29, Oxford Univ. Press, 2003)

[8] H. Bateman: The solution of partial differential equations by means of definite integrals *Proc. London Math. Soc.* **1(2)** (1904) 451–458

[9] V. Benci and J.-M. Coron: The Dirichlet problem for harmonic maps from the disk into the Euclidean n-sphere *Ann. Inst. H. Poincaré, Anal. Non Linéaire* **2(2)** (1985) 119–141

[10] F. Bethuel: A characterization of maps in $H^1(B^3, S^2)$ which can be approximated by smooth maps *Ann. Inst. H. Poincaré, Anal. Non Linéaire* **7(4)** (1990) 269–286

[11] F. Bethuel, J.-M. Coron, F. Demengel and F. Hélein: A cohomological criterion for density of smooth maps in Sobolev spaces between two manifolds *In: Nematics* (Orsay, 1990, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. vol. 332, Kluwer Acad. Publ., Dordrecht, 1991) 15–23

[12] F. Bethuel: The approximation problem for Sobolev maps between manifolds *Acta Math.* **167** (1991) 167–201

[13] F. Bethuel: On the singular set of stationary harmonic maps *Manuscripta Math.* **78(4)** (1993) 417–443

[14] F. Bethuel and H. Brezis: Regularity of minimizers of relaxed problems of harmonic maps *J. Funct. Anal.* **101** (1991) 145–161

[15] F. Bethuel, H. Brezis and J.-M. Coron: *Relaxed energies for harmonic maps* (Variational Methods, H. Berestycki, J.-M. Coron and I. Ekeland eds., Birhäuser, 1990)

[16] F. Bethuel and F. Demengel: Extensions for Sobolev mappings between manifolds *Calc. Var. Partial Differential Equations* **3(4)** (1995) 475–491

[17] F. Bethuel and X. Zheng: Density of smooth functions between two manifolds *J. Funct. Anal.* **80(1)** (1988) 60–75

[18] G. D. Birkhoff: Dynamical systems with two degrees of freedom *Trans. Amer. Math. Soc.* **18** (1917) 199–300

[19] A. I. Bobenko: All constant mean curvature tori in R^3, S^3, H^3 in terms of theta-functions *Math. Ann.* **290** (1991) 209–245

[20] J. Bourgain, H. Brezis and P. Mironescu: Lifting in Sobolev spaces *J. Anal. Math.* **80** (2000) 37–86

[21] H. Brezis: Analysis and topology in some nonlinear PDE problems *Bull. Amer. Math. Soc.* **40(2)** (2003) 179–201

[22] H. Brezis and J.-M. Coron: Large solutions for harmonic maps in two dimensions *Comm. Math. Phys.* **92** (1983) 203–215

[23] H. Brezis and J.-M. Coron: Multiple solutions of H-system and Rellich's conjecture *Comm. Pure Appl. Math.* **37** (1984) 149–187

[24] H. Brezis, J.-M. Coron and E. H. Lieb: Harmonic maps with defects *Comm. Math. Phys.* **107** (1986) 649–705

[25] H. Brezis and Y. Li: Topology and Sobolev spaces *J. Funct. Anal.* **183** (2001) 321–369

[26] H. Brezis, Y. Li, P. Mironescu and L. Nirenberg: Degree and Sobolev spaces *Topol. Methods Nonlinear Anal.* **13** (1999) 181–190

[27] H. Brezis and L. Nirenberg: Degree theory and BMO *Sel. Math., New Series* **1** (1995) 197–263

[28] R. L. Bryant: Conformal and minimal immersions of compact surfaces into the 4-sphere *J. Diff. Geom.* **17** (1982) 455–473

[29] R. L. Bryant: Harmonic morphisms with fibers of dimension one *Comm. Anal. Geom.* **8** (2000) 219–265

[30] F. E. Burstall: Harmonic maps of finite energy from non-compact manifolds *J. London Math. Soc.* **30** (1984) 361–370

[31] F. E. Burstall: *Nonlinear functional analysis and harmonic maps* (thesis, University of Warwick, 1984)

[32] F. E. Burstall: Harmonic tori in spheres and complex projective spaces *J. Reine Angew. Math.* **469** (1995) 146–177

[33] F. E. Burstall, D. Ferus, F. Pedit and U. Pinkall: Harmonic tori in symmetric spaces and commuting Hamiltonian systems on loop algebras *Annals of Math.* **138** (1993) 173–212

[34] F. E. Burstall and M. A. Guest: Harmonic two-spheres in compact symmetric spaces, revisited *Math. Ann.* **309(4)** (1997) 541–572

[35] F. E. Burstall and J. H. Rawnsley: *Twistor theory for Riemannian symmetric spaces. With applications to harmonic maps of Riemann surfaces* (Lecture Notes in Mathematics, 1424. Springer, Berlin, 1990)

[36] E. Calabi: Minimal immersions of surfaces in Euclidean spheres *J. Diff. Geom.* **1** (1967) 111–125

[37] K. C. Chang: Heat flow and boundary value problem for harmonic maps *Ann. Inst. Henri Poincaré Analyse Non Linéaire* **6(5)** (1989) 363–395

[38] K. C. Chang, W. Ding and R. Ye: Finite time blow up of the heat flow of harmonic maps from surfaces *J. Diff. Geom.* **36** (1992) 507–515

[39] S.-Y. A. Chang, L. Wang and P. C. Yang: Regularity of harmonic maps *Comm. Pure Appl. Math.* **52(9)** (1999) 1099–1111

[40] S. Chanillo: Sobolev inequalities involving divergence free maps *Comm. Partial Differential Equations* **16** (1991) 1969–1994

[41] Y. Chen: The weak solutions to the evolution problems of harmonic maps *Math. Z.* **201** (1989) 69–74

[42] Y. Chen and F. H. Lin: Remarks on approximate harmonic maps *Comment. Math. Helv.* **70(1)** (1995) 161–169

[43] Y. Chen and M. Struwe: Existence and partial regularity results for the heat flow for harmonic maps *Math. Z.* **201(1)** (1989) 83–103

[44] S.-S. Chern: Minimal surfaces in an Euclidean space of N dimensions *In: Differential and Combinatorial Topology* (A Symposium in Honor of Marston Morse, Princeton Univ. Press, Princeton, N.J., 1965) 187–198

[45] R. Coifman, P.-L. Lions, Y. Meyer and S. Semmes: Compensated compactness and Hardy spaces *J. Math. Pures Appl.* **72** (1993) 247–286

[46] T. H. Colding and W. P. Minicozzi II: An excursion into geometric analysis *In: Surveys in Differential Geometry* (S.-T. Yau, ed., vol. VIII, Int. Press, 2003) 75–107

[47] P. Collin and H. Rosenberg: Construction of harmonic diffeomorphisms and minimal graphs *Preprint* `arXiv:math/0701547` (2007)

[48] C. Constantinescu and A. Cornea: *Potential theory on harmonic spaces* (Grundlehren der mathematischen Wissenschaften, vol. 158, Springer, New York, 1972)

[49] J.-M. Coron: Nonuniqueness for the heat flow of harmonic maps *Ann. Inst. H. Poincaré Anal. Non Linéaire* **7(4)** (1990) 335–344

[50] J.-M. Coron and J.-M. Ghidaglia: Explosion en temps fini pour le flot des applications harmoniques *C. R. Acad. Sci. Paris Sér. I Math.* **308(12)** (1989) 339–344

[51] J.-M. Coron and R. Gulliver: Minimizing p-harmonic maps into spheres *J. Reine Angew. Math.* **401** (1989) 82–100

[52] J.-M. Coron and F. Hélein: Harmonic diffeomorphisms, minimizing maps and rotational symmetry *Compositio Math.* **69** (1989) 175–228

[53] Ĭ. Davidov and A. G. Sergeev: Twistor spaces and harmonic maps *Uspekhi Mat. Nauk* **48(3)(291)** (Russian), (1993) 3–96; *translation in Russian Math. Surveys* **48(3)** (1993) 1–91

[54] W. Y. Ding: Lusternik-Schnirelman theory for harmonic maps *Acta Math. Sinica* **2** (1986) 15–122

[55] W. Y. Ding: Symmetric harmonic maps between spheres *Comm. Math. Phys.* **118(4)** (1988) 641–649

[56] W. Y. Ding: Harmonic Hopf constructions between spheres *Internat. J. Math.* **5(4)** (1994) 641–649

[57] W. Y. Ding, H. Fan and J. Y. Li: Harmonic Hopf constructions between spheres. II *Calc. Var. Partial Differential Equations* **16(3)** (2003) 273–282

[58] W. Y. Ding and G. Tian: Energy identity for a class of approximate harmonic maps from surfaces *Comm. Anal. Geom.* **3** (1995) 543–554

[59] J. Dorfmeister and J.-H. Eschenburg: Pluriharmonic maps, loop groups and twistor theory *Ann. Global Anal. Geom.* **24(4)** (2003) 301–321

[60] J. Dorfmeister, F. Pedit and H. Wu: Weierstrass type representation of harmonic maps into symmetric spaces *Comm. Anal. Geom.* **6(4)** (1998) 633–668

[61] J. Eells and L. Lemaire: A report on harmonic maps *Bull. London Math. Soc.* **10** (1978) 1–68

[62] J. Eells and L. Lemaire: *Selected Topics in Harmonic Maps* (C.B.M.S. Regional Conf. Series, vol. 50 American Mathematical Society, Providence, RI, 1983)

[63] J. Eells and L. Lemaire: Another report on harmonic maps *Bull. London Math. Soc.* **20** (1988) 385–524

[64] J. Eells and A. Ratto: *Hamonic maps and minimal immersions with symmetries* (Annals of Math. Studies, Princeton University Press, 1993)

[65] J. Eells and S. Salamon: Twistorial construction of harmonic maps of surfaces into four-manifolds *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* **12(4)** (1985) 589–640

[66] J. Eells and J. H. Sampson: Harmonic mappings of Riemannian manifolds *Amer. J. Math.* **86** (1964) 109–160

[67] J. Eells and J. C. Wood: Restrictions on harmonic maps of surfaces *Topology* **15** (1976) 263–266

[68] J. Eells and J. C. Wood: Harmonic maps from surfaces to complex projective spaces *Adv. in Math.* **49** (1983) 217–263

[69] J.-H. Eschenburg R. Tribuzy: Associated families of pluriharmonic maps and isotropy *Manuscripta Math.* **95(3)** (1998) 295–310

[70] L. C. Evans: Partial regularity for stationary harmonic maps into spheres *Arch. Rational Mech. Anal* **116(2)** (1991) 101–113

[71] H. Federer: The singular sets of area minimizing rectifiable currents with codimension one and of area minimizing flat chains modulo two with arbitrary codimension *Bull. Amer. Math. Soc.* **76** (1970) 767–771

[72] C. Fefferman and E. M. Stein: H^p spaces of several variables *Acta Math.* **129(3-4)** (1972) 137–193

[73] A. P. Fordy and J. C. Wood, eds.: *Harmonic maps and integrable systems* (Vieweg, Wiesbaden; Available from `http://www.maths.leeds.ac.uk/pure/staff/wood/FordyWood/contents.html`, 1994)

[74] A. Freire: Uniqueness for the harmonic map flow in two dimensions *Calc. Var. Partial Differential Equations* **3** (1995) 95–105;
A. Freire: Uniqueness for the harmonic map flow from surfaces to general targets *Comment. Math. Helv.* **70(2)** (1995) 310–338; *correction in Comment. Math. Helv.* **71(2)** (1996) 330–337

[75] M. Fuchs: p-harmonic obstacle problems I. Partial regularity theory *Ann. Mat. Pura Appl.* **156(4)** (1990) 127–158

[76] M. Fuchs: p-harmonic obstacle problems II. Extensions of maps and applications *Manuscripta Math.* **63(4)** (1989) 381–419

[77] B. Fuglede: Harmonic morphisms between Riemannian manifolds *Ann. Inst. Fourier (Grenoble)* **28(2)** (1978) 107–144

[78] A. Gastel: Regularity theory for minimizing equivariant p-harmonic mappings *Calc. Var. Partial Differential Equations* **6(4)** (1998) 329–367

[79] A. Gastel: Construction of harmonic maps between spheres by joining three eigenmaps *Internat. J. Math.* **9(7)** (1998) 821–844

[80] A. Gastel: Torus equivariant harmonic maps between spheres *Topology* **41(2)** (2002) 231–227

[81] A. Gastel: On the harmonic Hopf construction *Proc. Amer. Math. Soc.* **132(2)** (2004) 607–615

[82] Y. Ge: Harmonic maps of degree 1 into the unit 2-sphere *Asian J. Math.* **4(3)** (2000) 621–632

[83] M. Giaquinta: *Multiple integrals in the calculus of variations and nonlinear elliptic systems* (Princeton University Press, 1983)

[84] M. Giaquinta and E. Giusti: On the regularity of the minima of variational integrals *Acta Math.* **148** (1982) 31–46

[85] M. Giaquinta and E. Giusti: The singular set of the minima of certain quadratic functionals *Ann. Scuola Norm. Sup. Pisa Cl. Sci., (4)* **11(1)** (1984) 45–55

[86] M. Giaquinta and S. Hildebrandt: A priori estimates for harmonic mappings *J. Reine Angew. Math.* **336** (1982) 124–164

[87] M. Giaquinta, G. Modica and J. Souček: The Dirichlet energy of mappings with values into the sphere *Manuscripta Math.* **65** (1989) 489–507

[88] M. Giaquinta, G. Modica and J. Souček: *Cartesian currents in the calculus of variations II* (Ergebnisse der Mathematik und ihrer Grenzgebiete, Vol. 38, Springer, Berlin, 1998)

[89] M. Giaquinta and J. Souček: Hamonic maps into a hemisphere *Ann. Scuola Norm. Sup. Pisa* **12** (1985) 81–90

[90] D. Gilbarg and N. S. Trudinger: *Elliptic partial differential equations of second order* (Springer, 1997)

[91] M. L. Gromov and V. A. Rohlin: Imbeddings and immersions in Riemannian geometry *Uspehi Mat. Nauk* **25(5)(155)** (Russian), (1970) 3–62; *English translation: Russian Math. Surveys* **25(5)** (1970) 1–57

[92] M. Grüter: Regularity of weak H-surfaces *J. Reine Angew. Math.* **329** (1981) 1–15

[93] S. Gudmundsson and M. Svensson: On the existence of harmonic morphisms from certain symmetric spaces *J. Geom. Phys.* **57** (2007) 353–366

[94] M. A. Guest: *Harmonic maps, loop groups, and integrable systems* (London Mathematical Society Student Texts, 38, Cambridge University Press, Cambridge, 1997)

[95] R. D. Gulliver, R. Osserman and H. L. Royden: A theory of branched immersions of surfaces *Amer. J. Math.* **95** (1973) 750–812

[96] P. Hajłasz: Approximation of Sobolev mappings *Nonlinear analysis, Theory, Methods and Appl.* **22(12)** (1994) 1579–1591

[97] P. Hajłasz: Sobolev mappings, co-area formula and related topics *In: Proc. on Analysis and Geometry, Novosibirsk* (Sobolev Institute Press, 2000) 227–254

[98] P. Hajłasz, T. Iwaniec, J. Malý and J. Onninen: *Weakly differentiable mappings between manifolds* (Memoirs of the American Math. Soc., to appear)

[99] R. Hamilton: *Harmonic maps of manifolds with boundary* (Springer Lect. Notes in Math. 471, 1975)

[100] F. B. Hang: Density problems for $W^{1,1}(M,N)$ *Commun. Pure Appl. Math.* **55** (2002) 937–947

[101] F. B. Hang: On the weak limits of smooth maps for the Dirichlet energy between manifolds *Comm. Anal. Geom.* **13(5)** (2005) 929–938

[102] F. B. Hang and F. H. Lin: Topology of Sobolev mappings, II *Acta Math.* **191** (2003) 55–107

[103] R. Hardt: Singularities of harmonic maps *Bull. Amer. Math. Soc.* **34(1)** (1997) 15–34

[104] R. Hardt, D. Kinderlehrer and F. H. Lin: Existence and partial regularity of static liquid crystal configurations *Comm. Math. Phys.* **105** (1986) 547–570

[105] R. Hardt, D. Kinderlehrer and F. H. Lin: Stable defects of minimizers of constrained variational problems *Ann. Inst. H. Poincaré Anal. Non Linéaire* **5** (1988) 297–322

[106] R. Hardt and F. H. Lin: A remark on H^1 mappings *Manuscripta Math.* **56** (1986) 1–10

[107] R. Hardt and F. H. Lin: Mappings minimizing the L^p-norm of the gradient *Commun. Pure Appl. Math.* **40** (1987) 556–588

[108] R. Hardt and F. H. Lin: The singular set of an energy minimizing map from B^4 to S^2 *Manuscripta Math.* **69(3)** (1990) 275–289

[109] R. Hardt, F. H. Lin and C. Poon: Axially symmetric harmonic maps minimizing a relaxed energy *Comm. Pure Appl. Math.* **45** (1992) 417–459

[110] R. Hardt and T. Rivière: Connecting topological Hopf singularities *Ann. Sci. Norm. Sup. Pisa Cl. Sci., (5)* **2(2)** (2003) 287–344

[111] P. Hartman: On homotopic harmonic maps *Canad. J. Math.* **19** (1967) 673–687

[112] F. Hélein: Difféomorphismes harmoniques entre un ouvert de $\mathbb{R}^3$ et une variété riemannienne *C. R. Acad. Sci. Paris* **308** (1989) 237–240

[113] F. Hélein: Régularité des applications harmoniques entre une surface et une sphère *C. R. Acad. Sci. Paris* **311** (1990) 519–524

[114] F. Hélein: Regularity of weakly harmonic maps form a surface into a manifold with symmetries *Manuscripta Math.* **70** (1991) 293–318

[115] F. Hélein: Harmonic diffeomorphisms with rotational symmetry *J. Reine Angew. Math.* **414** (1991) 45–49

[116] F. Hélein: Régularité des applications harmoniques entre une surface et une variété riemannienne *C. R. Acad. Sci. Paris* **312** (1991) 591–596

[117] F. Hélein: *Harmonic maps, constant mean curvature surfaces and integrable systems* (Lectures in Mathematics, ETH Zürich, Birkhäuser, 2001)

[118] F. Hélein: *Harmonic maps, conservation laws and moving frames* (Cambridge Tracts in Mathematics 150, Cambridge University Press, 2002)

[119] S. Hildebrandt, W. Jäger and K.-O. Widman: Harmonic mappings into Riemannian manifolds with non-positive sectional curvature *Math. Scand.* **37(2)** (1975) 257–263

[120] S. Hildebrandt, W. Jäger and K.-O. Widman: An existence theorem for harmonic mappings of Riemannian manifolds *Acta Math.* **138(1-2)** (1977) 1–16

[121] S. Hildebrandt and K.-O. Widman: On the Hölder continuity of weak solutions of quasilinear elliptic systems of second order *Ann. Scuola Norm. Sup. Pisa Cl. Sci., (4)* **4(1)** (1977) 145–178

[122] N. Hitchin: Harmonic maps from a 2-torus to a 3-sphere *J. Diff. Geom.* **31** (1990) 627–710

[123] N. Hitchin: Integrable systems in Riemannian geometry *In: Surveys in differential geometry: integral systems [integrable systems], 2181* (Surv. Differ. Geom., IV, Int. Press, Boston, MA, 1998) 21–81

[124] T. Ishihara: A mapping of Riemannian manifolds which preserves harmonic functions *J. Math. Kyoto Univ.* **19** (1979) 215–229

[125] C. G. J. Jacobi: Über eine Lösung der partiellen Differentialgleichung $\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0$ *J. Reine Angew. Math.* **36** (1848) 113–134

[126] W. Jäger and H. Kaul: Uniqueness and stability of harmonic maps and their Jacobi fields *Manuscripta Math.* **28** (1979) 269–291

[127] W. Jäger and H. Kaul: Rotationally symmetric harmonic maps from a ball into a sphere and the regularity problem for weak solutions of elliptic systems *J. Reine Angew. Math.* **343** (1983) 146–161

[128] F. John and L. Nirenberg: On functions of bounded mean oscillation *Comm. Pure and Appl. Math.* **14** (1961) 415–426

[129] J. Jost: A maximum principle for harmonic mappings which solves a Dirichlet problem *Manuscripta Math.* **38** (1982) 129–130

[130] J. Jost: Existence proof for harmonic mappings with the help of the maximum principle *Math. Z.* **184** (1983) 489–496

[131] J. Jost: The Dirichlet problem for harmonic maps from a surface with boundary onto a 2-sphere with nonconstant boundary values *J. Diff. Geom.* **19(2)** (1984) 393–401

[132] J. Jost: *Harmonic maps between surfaces* (Springer, Berlin-New York, 1984)

[133] J. Jost: *Two dimensional geometric variational problems* (Wiley-Interscience, Chichester, 1991)

[134] J. Jost: Unstable solutions of two-dimensional geometric variational problems *In: Differential geometry: partial differential equations on manifolds* (Proc. Sympos. Pure Math., 54, Part 1, Amer. Math. Soc., 1993) 205–244

[135] J. Jost: *Riemannian geometry and geometric analysis* (Universitext, Springer, 1998)

[136] J. Jost and H. Karcher: Geometrische Methoden zur Gewinnung von a priori-Schranken für harmonisch Abbildungen *Man. Math.* **10** (1982) 27–77;
J. Jost and H. Karcher: Almost linear functions and a priori estimates for harmonic maps *In: Global Riemmannian Geometry* (Durham, 1982, E. Horwood, Chichester, 1984) 148–155

[137] J. Jost and R. Schoen: On the existence of harmonic diffeomorphisms *Invent. Math.* **66(2)** (1982) 353–359

[138] J. Jost and M. Struwe: Morse-Conley theory for minimal surfaces of varying topological type *Invent. Math.* **102(3)** (1990) 465–499

[139] N. Kapouleas: Complete constant mean curvature surfaces in Euclidean three-space *Ann. of Math., (2)* **131(2)** (1990) 239-330;
N. Kapouleas: Compact constant mean curvature surfaces in Euclidean three-space *J. Diff. Geom.* **33(3)** (1991) 683–715

[140] N. Kapouleas: Constant mean curvature surfaces constructed by fusing Wente tori *Invent. Math.* **119(3)** (1995) 443–518

[141] H. Karcher and J. C. Wood: Non-existence results and growth properties for harmonic maps and forms *J. Reine Angew. Math.* **353** (1984) 165–180

[142] E. Kuwert: Minimizing the energy from a surface into a 2-sphere with prescribed degree and boundary values *Manuscripta Math.* **83** (1994) 31–38

[143] O. A. Ladyzhenskaya and N. N. Ural'tseva: *Linear and quasilinear elliptic equations* (New York, 1968)

[144] L. Lemaire: Applications harmoniques de surfaces riemanniennes *J. Diff. Geom.* **13** (1979) 51–78; *C. R. Acad. Sci. Paris* **279** (1974) 925–927; *C. R. Acad. Sci. Paris* **280** (1975) 897–899

[145] L. Lemaire and J.C. Wood: On the space of harmonic 2-spheres in $\mathbb{C}P^2$ *Internat. J. Math.* **7** (1996) 211–225

[146] F. H. Lin: A remark on the map $x/|x|$ *C. R. Acad. Sci. Paris* **305** (1987) 529–531

[147] F. H. Lin: Gradient estimates and blow-up analysis for stationary harmonic maps *Ann. of Math., (2)* **149(3)** (1999) 785–829

[148] F. H. Lin: Mapping problems, fundamental groups and defect measures *Acta Math. Sin.* **15(1)** (1999) 25–52

[149] F. H. Lin and T. Rivière: Energy quantization for harmonic maps *Duke Math. J.* **111(1)** (2002) 177–193

[150] P.-L. Lions: The concentration-compactness principle in the calculus of variations. The limit case I and II *Rev. Mat. Iberoamericana* **1(1)** (1985) 145–201; *Rev. Mat. Iberoamericana* **1(2)** (1985) 45–121

[151] S. Luckhaus: Partial Hölder continuity for minima of certain energies among maps into a Riemannian manifold *Indiana Univ. Math. J.* **37(2)** (1988) 349–367

[152] S. Luckhaus: Convergence of minimizers for the p-Dirichlet integral *Math. Z.* **213** (1993) 449–456

[153] R. Mazzeo and F. Pacard: Connected sums of constant mean curvature surfaces in Euclidean 3 space *J. Reine Angew. Math.* **536** (2001) 115–165

[154] I. McIntosh: Two remarks on the construction of harmonic tori in $\mathbb{C}P^n$ *Internat. J. Math.* **7(4)** (1996) 515–520

[155] P. Mironescu: On some properties of S^1-valued fractional Sobolev spaces *Contemporary Math.* **350** (2004) 201–207

[156] C. B. Morrey: The problem of Plateau on a Riemannian manifold *Annals of Math.* **49(4)** (1948) 807–851

[157] C. B. Morrey: *Multiple integrals in the calculus of variations* (Springer, 1966)
[158] L. Mou and P. Yang: Regularity of n-harmonic maps *J. Geom. Anal.* **6(1)** (1996) 91–112
[159] F. Murat: Compacité par compensation : condition nécessaire et suffisante de continuité faible sous une hypothèse de rang constant *Ann. Scuola Normale Sup., Pisa, Serie VIII* **1** (1981) 69–102
[160] Y. Ohnita and S. Udagawa: Harmonic maps of finite type into generalized flag manifolds, and twistor fibrations *In: Differential geometry and integrable systems* (Tokyo, 2000, Contemp. Math., 308, Amer. Math. Soc., Providence, RI, 2002) 245–270
[161] R. Pacheco: Twistor fibrations giving primitive harmonic maps of finite type *Int. J. Math. Math. Sci.* **20** (2005) 3199–3212
[162] M. R. Pakzad and T. Rivière: Weak density of smooth maps for the Dirichlet energy between manifolds *Geom. Funct. Anal.* **13(1)** (2003) 223–257
[163] R. Palais: The principle of symmetric criticality *Comm. Math. Phys.* **69(1)** (1979) 19–30
[164] T. H. Parker: Bubble tree convergence for harmonic maps *J. Diff. Geom.* **44(3)** (1996) 595–633
[165] R. Penrose: Solutions of the zero rest-mass equations *J. Math. Phys.* **10** (1969) 38–39
[166] O. Perron: Eine neue Behandlung der ersten Randwertaufgabe für $\Delta u = 0$ *Math. Z.* **18** (1923) 42–54
[167] V. Pettinati and A. Ratto: Existence and non-existence results for harmonic maps between spheres *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **17(2)** (1991) 273–282
[168] I. G. Petrowsky: Sur l'analyticité des solutions des systèmes d'équations différentielles *Rec. Math. (N. S.) [Mat. Sbornik], (5)* **47** (1939) 3–70
[169] U. Pinkall and I. Sterling: On the classification of constant mean curvature tori *Ann. Math.* **130** (1989) 407–451
[170] A. I. Pluzhnikov: Some properties of harmonic mappings in the case of spheres and Lie groups *Sov. Math. Dokl.* **27** (1983) 246–248
[171] A. I. Pluzhnikov: A topological criterion for the attainability of global minima of an energy function *Nov. Glob. Anal. Voronezh. Gos. Univ.* **177** (1986) 149–155
[172] K. Pohlmeyer: Integrable Hamiltonian systems and interactions through constraints *Comm. Math. Phys.* **46** (1976) 207–221
[173] H. Poincaré: Sur les équations aux dérivées partielles de la physique mathématique *Amer. J. Math.* **12** (1890) 211–294
[174] J. Qing: Remark on the Dirichlet problem for harmonic maps from the disc into the 2-sphere *Proc. Roy. Soc. Edinburgh* **122A** (1992) 63–67
[175] J. Qing: Multiple solutions of the Dirichlet problem for harmonic maps from discs into 2-spheres *Manuscripta Math.* **77(4)** (1992) 435–446
[176] J. Qing: The Dirichlet problem for harmonic maps from the disc into Kähler manifolds *J. London Math. Soc., (2)* **47(1)** (1993) 182–192
[177] J. Qing: On the singularities of the heat flow for harmonic maps from surfaces into spheres *Comm. Anal. Geom.* **3(1-2)** (1995) 297–315
[178] J. Qing and G. Tian: Bubbling of the heat flows for harmonic maps from surface *Comm. Pure Appl. Math.* **50** (1997) 295–310

[179] A. Ratto: Harmonic maps of spheres and the Hopf construction *Topology* **28(3)** (1989) 379–388

[180] J. H. Rawnsley: f-structures, f-twistor spaces and harmonic maps *In: Geometry Seminar L. Bianchi* (Lect. Notes Math. 1164, Springer, Berlin, Heidelberg, New York, 1985) 85–159

[181] T. Rivière: Harmonic maps from B^3 to S^2 having a line of singularities *C. R. Acad. Sci. Paris Série I Math.* **101** (1991) 583–587

[182] T. Rivière: Everywhere discontinuous harmonic maps into spheres *Acta Math.* **175** (1995) 197–226

[183] T. Rivière: Dense subsets of $H^{1/2}(S^2; S^1)$ *Ann. Global Anal. Geom.* **18(5)** (2000) 517–528

[184] T. Rivière: Conservation laws for conformally invariant variational problems *Invent. Math.* **168(1)** (2007) 1–22

[185] T. Rivière and M. Struwe: Partial regularity for harmonic maps, and related problems *Preprint* (2006)

[186] J. Rubinstein and P. Sternberg: Homotopy classification of minimizers of the Ginzburg-Landau energy and the existence of permanent currents *Comm. Math. Phys.* **179** (1996) 257–263

[187] E. A. Ruh and J. Vilms: The tension field of the Gauss map *Trans. Amer. Math. Soc.* **149** (1970) 569–573

[188] J. Sacks and K. Uhlenbeck: The existence of minimal immersions of two-spheres *Bull. Amer. Math. Soc.* **83** (1977) 1033–1036;
J. Sacks and K. Uhlenbeck: The existence of minimal immersions of 2-spheres *Ann. Math., (2)* **113** (1981) 1–24

[189] R. Schoen: Analytic aspects of the harmonic maps problem *In: Seminar on nonlinear PDE* (S. S. Chern (ed.), Springer, 1984) 321–358

[190] R. Schoen and K. Uhlenbeck: A regularity theory for harmonic maps *J. Diff. Geom.* **17** (1982) 307–335

[191] R. Schoen and K. Uhlenbeck: Boundary regularity and the Dirichlet problem for harmonic maps *J. Diff. Geom.* **18** (1983) 253–268

[192] R. Schoen and K. Uhlenbeck: Regularity of minimizing harmonic maps into the sphere *Invent. Math.* **78(1)** (1984) 89–100

[193] R. Schoen and S. T. Yau: Existence of incompressible minimal surfaces and the topology of three-dimensional manifolds with nonnegative scalar curvature *Ann. of Math.* **110** (1979) 127–142

[194] R. Schoen and S. T. Yau: *Lectures on harmonic maps* (International Press, 1994)

[195] L. Schwartz: *Théorie des Distributions* (Hermann, Paris Nouveau tirage, 1998); 1st edn., (1950/1951)

[196] G. Segal: Loop groups and harmonic maps *In: Advances in homotopy theory* (Cortona, 1988, London Math. Soc. Lecture Note Ser., 139, Cambridge Univ. Press, Cambridge, 1989) 153–164

[197] J. Shatah: Weak solutions and developments of singularities of the $SU(2)$ σ-model *Comm. Pure Appl. Math.* **41** (1988) 459–469

[198] L. Simon: Asymptotics for a class of nonlinear evolution equations, with applications to geometric problems *Ann. of Math., (2)* **118(3)** (1983) 525–571

[199] L. Simon: Theorems on the regularity and singularity of minimal surfaces and harmonic maps *In: Geometry and Global Analysis* (Tohoku Univ., Sendai, 1993) 111–145

[200] L. Simon: Rectifiability of the singular set of energy minimizing maps *Calc. Var. Partial Differential Equations* **3** (1995) 1–65

[201] L. Simon: *Theorems on regularity and singularity of energy minimizing maps* (Lectures in Mathematics, ETH Zürich, Birkhäuser, 1996)

[202] R. T. Smith: Harmonic mappings of spheres *Amer. J. Math.* **97** (1975) 364–385

[203] A. Soyeur: The Dirichlet problem for harmonic maps from the disc into the 2-sphere *Proc. Royal Soc. Edinburgh* **113A** (1989) 229–234

[204] K. Steffen: *An introduction to harmonic mappings* (SFB 256, Bonn Vorlesungsreihe, Vol. 18, Bonn, 1991)

[205] E. Stein: *Harmonic analysis* (Princeton University Press, Princeton, 1993)

[206] E. Stein and G. Weiss: On the theory of harmonic functions of several variables. I. The theory of H^p-spaces *Acta Math.* **103** (1960) 25–42

[207] M. Struwe: On the evolution of harmonic maps of Riemannian manifolds *Comment. Math. Helv.* **60** (1985) 558–581

[208] M. Struwe: On the evolution of harmonic maps in higher dimensions *J. Diff. Geom.* **28(2)** (1988) 485–502

[209] M. Struwe: *Variational methods. Applications to non linear partial differential equations and Hamiltonnian systems* (second edition, Springer, 1996)

[210] M. Struwe: Uniqueness of harmonic maps with small energy *Manuscripta Math.* **96(4)** (1998) 463–486

[211] L. Tartar: The compensated compactness method applied to systems of conservation laws *In: Systems of nonlinear partial differential equations* (J.M. Ball ed., NATO ASI Series C111, Reidel, New York, 1983) 263–285

[212] C. Taubes: The existence of a nonminimal solution to the $SU(2)$ Yang-Mills-Higgs equations on $\mathbb{R}^3$. I and II *Comm. Math. Phys.* **86(2)** (1982) 257–298;
C. Taubes: The existence of a nonminimal solution to the $SU(2)$ Yang-Mills-Higgs equations on $\mathbb{R}^3$. III *Comm. Math. Phys.* **86(3)** (1982) 299–320

[213] P. Topping: Rigidity of the harmonic map heat flow *J. Diff. Geom.* **45(3)** (1997) 593–610

[214] P. Topping: Reverse bubbling and nonuniqueness in the harmonic map flow *Int. Math. Res. Not.* **10** (2002) 505–520

[215] P. Topping: Repulsion and quantization in almost-harmonic maps, and asymptotics of the harmonic maps flow *Ann. of Math., (2)* **159(2)** (2004) 465–534

[216] K. Uhlenbeck: Harmonic maps: a direct method in the calculus of variations *Bull. Amer. Math. Soc.* **76** (1970) 1082–1087

[217] K. Uhlenbeck: Connections with L^p bounds on curvature *Comm. Math. Phys.* **83(1)** (1982) 31–42

[218] K. Uhlenbeck: Harmonic maps into Lie groups: classical solutions of the chiral model *J. Diff. Geom.* **30(1)** (1989) 1–50

[219] H. Urakawa: *Calculus of variations and harmonic maps* (Translations of Mathematical Monographs, vol. 132. American Mathematical Society, Providence, RI, 1993)

[220] G. F. Wang: Harmonic maps from a closed surface of higher genus to the unit 2-sphere *Internat. J. Math.* **4** (1993) 359–365

[221] H. Wente: An existence theorem for surfaces of constant mean curvature *J. Math. Anal. Appl.* **26** (1969) 318–344

[222] H. Wente: Counterexample to a conjecture of H. Hopf *Pacific J. Math.* **121(1)** (1986) 193–243

[223] H. Weyl: The method of orthogonal projection in potential theory *Duke Math. J.* **7** (1940) 411–440

[224] E. T. Whitakker: On the general solution of Laplace's equation and the equation of wave motions, and an undulatory explanation of gravity *Monthly Notices Roy. Astron. Soc.* **62** (1902) 617–620

[225] B. White: Homotopy classes in Sobolev spaces and energy minimizing maps *Bull. Amer. Math. Soc.* **13** (1986) 166–168

[226] B. White: Infima of energy functions in homotopy classes of mappings *J. Diff. Geom.* **23** (1986) 127–142

[227] B. White: Homotopy classes in Sobolev spaces and the existence of energy minimizing maps *Acta Math.* **160** (1988) 1–17

[228] B. White: Nonunique tangent maps at isolated singularities of harmonic maps *Bull. Amer. Math. Soc.* **26(1)** (1992) 125–129

[229] M. Wiegner: Ein optimaler Regularitätssatz für schwache Lösungen gewisser elliptischer Systeme *Math. Z.* **147** (1976) 21–28

[230] J. C. Wood: Non-existence of solutions to certain Dirichlet problems for harmonic maps *Preprint* University of Leeds (1981)

[231] Y. L Xin: *Geometry of harmonic maps* (Progress in Nonlinear Differential Equations and Their Applications, Birkhäuser, 1996)

FRÉDÉRIC HÉLEIN

Université Denis Diderot (Paris 7), Institut de Mathématiques de Jussieu — UMR 7586, Case 7012, 2 place Jussieu, 75251 Paris Cedex 05, France

E-mail: `helein@math.jussieu.fr`

JOHN C. WOOD

Department of Pure Mathematics, University of Leeds, Leeds LS2 9JT, Great Britain

E-mail: `J.C.Wood@leeds.ac.uk`

This page intentionally left blank

Handbook of Global Analysis
Demeter Krupka and David Saunders

Topology of differentiable mappings

Kevin Houston

Contents

1 Introduction

Given a smooth map $f : N \to P$ from one manifold to another, it is natural to ask for a description of the topology of $f(N)$ and $f^{-1}(p)$ for $p \in P$. One can see that for the topology of a map, generally, we have this option of studying *images* or *preimages*.

The theory for the latter is considerably more advanced than the former since Sard's Theorem and the Implicit Function Theorem tell us that, usually, the preimage of a smooth map is a manifold, while experience shows that, in general, the image of a map is singular. The selection of topics in this survey reveals that this is the case and attempts to redress the balance by presenting some lesser known results and, more importantly, techniques, in the study of the topology of images.

The dichotomy between images and preimages is presented in Section 2. Some fundamental examples such as the simple singularities of functions and the Whitney umbrella/cross-cap are presented.

The basic building block of the topology of differentiable maps is the Milnor fibre. This is presented in Section 3. This describes the local change, i.e., within a sufficiently small ball, of the topology of the singularity as one moves from the critical point to a nearby non-critical point. That is, from the singular fibre to a nearby non-singular one. Since the Milnor fibre is, in general, a non-trivial fibration over a circle we have monodromy. This is described in Section 4

In his 1978 obituary of Morse in [60], Smale said 'Morse theory is the single greatest contribution of American mathematics' (perhaps not surprisingly as the result formed the backbone of Smale's own Fields-medal-winning work on the h-cobordism theory and the higher dimensional Poincaré Conjecture. However, this obituary did cause discussion due to its critical nature). Morse Theory is well-known and covered in the highly readable [48] so it is the generalization to singular spaces, in fact, more properly, stratified spaces by Goresky and Macpherson that we describe Section 6. Their original intention had been to generalize the Lefschetz Hyperplane Theorem to the case of their (then) recently invented Intersection Cohomology. In pursuing this they invented Stratified Morse Theory. This subject is now fairly advanced, one can see by looking through [57] at the level of sophistication now possible. However, this sophistication is underused and there are many subjects to which it could be applied that have yet to be explored. The stratification of spaces is detailed in Section 5 and Stratified Morse Theory in Section 6.

The main result of Morse theory is that one can build up the topology of a manifold by placing on it a generic function that has non-degenerate singularities. The topology can then be described by attaching cells of dimension that depend on the index of the second differential of the the critical points of the map. In the stratified case we have to calculate the Morse index at a point but also have to take into account how the function behaves with respect to the space transverse to the stratum containing the point. The local intersection of the singular space and a manifold transverse to the stratum is called the normal slice. To apply Stratified Morse Theory we need to be able to describe the topology of the Morse function on this space. Unlike the usual Morse index, no very simple number exists to measure this topology. One method is to use Rectified Homotopical Depth, a concept, introduced by Grothendieck, which is analogous to the idea of depth from commutative algebra. In that theory regular rings and complete intersection rings have maximal depth, i.e., equal to the ring's dimension, we have that manifolds and local complete intersections have maximal Rectified Homotopical Depth, i.e., equal to the complex dimension of the space. Section 7 shows how this notion can be used with Stratified Morse Theory to describe the topology of certain complex analytic varieties in $\mathbb{CP}^n$.

Another interesting and greatly underutilized generalization of Stratified Morse Theory is given in Section 8. This is a relative version, i.e., for a stratified map $f : X \to Y$ between Whitney stratified sets a stratified Morse function on Y is used to describe the topology of X.

In the last sections we see a spectral sequence that allows us to deal with the topology of images. The potential applications of this sequence are quite large.

First, Section 9 gives examples of how images behave and discusses the multiple point spaces for a map. For a continuous map $f : X \to Y$ the kth multiple point space is

$$D^k(f) = \text{closure}\{(x_1, \dots, x_k) \in X^k \mid f(x_1) = \cdots = f(x_k), \text{ for } x_i \neq x_j,\, i \neq j\}.$$

The key here is that the image is very hard to describe, for example, the image of a real polynomial map may not be a real algebraic set. Now $D^k(f)$ can often be described as the zero-set of map and hence we can use the theory developed for level sets to produce a theory for images. Furthermore, $D^k(f)$ has a lot of symmetry since S_k, the group of permutations on k objects, acts on it, and this is exploited in describing the topology of the image. In fact, we need to focus on the the alternating homology of $D^k(f)$, that is, chains on $D^k(f)$ that are anti-invariant under the action of S_k. This alternating homology then forms the E^1 terms of a spectral sequence for a wide class of finite and proper maps. This sequence is called the Image Computing Spectral and is described, with examples, in Section 10.

2 Manifolds and singularities

Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is a smooth function with $n \geq 1$. A fundamental question in mathematics is 'What is the topology of the level sets $f^{-1}(c)$ for $c \in \mathbb{R}$?' Whitney showed the following.

Theorem 2.1 (Whitney [7]) *Let X be a closed set in $\mathbb{R}^n$. Then there exists a smooth function $f : \mathbb{R}^n \to \mathbb{R}$ such that $f^{-1}(0) = X$.*

Due to the wildness of closed sets – consider the pathological examples of the Cantor set and Hawaiian earrings – this, of course, means that finding a general structure theorem on the level set of an arbitrary smooth function f is essentially hopeless – and we have not even considered maps into higher dimensional spaces yet.

So we begin by specializing and look at a fundamental structure theorem for the level sets of certain smooth maps. First some definitions. Let $f : \mathbb{R}^n \to \mathbb{R}^p$ be a smooth map where n and p are arbitrary. If the differential at a point $x \in \mathbb{R}^n$ is surjective, then we say that f is a *submersion* at x and say that x is a *regular point* of f. If all points of $f^{-1}(c)$ are regular points, then we say that that c is a *regular value* of f. (This includes the case that c is not a value of f!)

Then we have the following.

Theorem 2.2 *Suppose that c is a regular value of f. Then $f^{-1}(c)$ is a $(n-p)$-dimensional submanifold of $\mathbb{R}^n$.*

The first remedy for the problem posed by Whitney's theorem is Sard's theorem which says that, in general, the level set is a manifold.

Theorem 2.3 (Sard's Theorem [56]) *The set of non-regular values of f has Lebesgue measure zero.*

Thus for mappings we have that, in general, a fibre is a manifold. For images the situation is not so good. However, we do have that if we have a map $f : \mathbb{R}^n \to \mathbb{R}^p$ with $n < p$ and maximal possible rank (i.e., n) at the point $x \in N$, then there is a neighbourhood U of x that maps into P such that $f(U)$ is a submanifold of P.

These last two theorems can be proved from the Inverse Function Theorem or Implicit Function Theorem. (These two theorems are in fact equivalent as each can be proved from the other.)

Theorem 2.4 (Implicit Function Theorem) *Suppose that $f : \mathbb{R}^{n+r} \to \mathbb{R}^r$ is a smooth map defined on a neighbourhood of $(x_0, y_0) \in \mathbb{R}^n \times \mathbb{R}^r$ with $f(x_0, y_0) = c$. If the $r \times r$ matrix*

$$\begin{pmatrix} \dfrac{\partial f_1}{\partial y_1} & \cdots & \dfrac{\partial f_1}{\partial y_r} \\ \vdots & & \vdots \\ \dfrac{\partial f_r}{\partial y_1} & \cdots & \dfrac{\partial f_r}{\partial y_r} \end{pmatrix}$$

is non-singular at (x_0, y_0), then there exists a neighbourhood of U of x_0 in $\mathbb{R}^n$ and V of

y_0 in $\mathbb{R}^q$ such that for all x in U there is unique point $g(x)$ in V with $f(x, g(x)) = c$. Furthermore, g is smooth.

The preceding can be generalized using the notion of transversality (which we shall use later in a different context).

Definition 2.5 Let U and V be submanifolds in $\mathbb{R}^n$. Then, U and V are *transverse at the point* $x \in U \cap V$ if

$$T_xU + T_xV = \mathbb{R}^n.$$

That is, the sum of the tangent spaces of U and V gives the tangent space to $\mathbb{R}^n$. We say that U and V are *transverse* if they are transverse for all points in $U \cap V$.

If U and V do not intersect, then automatically we say that the spaces are transverse.

The notion of transversality is very important as one would expect that two randomly chosen submanifolds would be transverse. In fact, if they were not, then using Sard's theorem one could perturb them slightly so that we had transversal intersection. Thus transversality is in some sense 'generic'.

We can produce a relative version of transversality.

Definition 2.6 Let $f : N \to P$ be a smooth map between the manifolds N and P. Suppose that C is a smooth submanifold of P. Then, f is said to be transverse to C at the point $x \in N$ if either $f(x) \notin C$ or the image of the tangent space to N under the differential d_xf is transverse to the tangent space of C at $f(x)$ in P. That is,

$$d_xf(T_xN) + T_{f(x)}C = T_{f(a)}P.$$

We say that f is *transverse to* C if it is transverse to C at all points $x \in N$.

Then we can generalize the structure theorem for level sets.

Theorem 2.7 *Suppose that $f : N \to P$ is a smooth between manifolds with C a submanifold of P. If f is transverse to C, then $f^{-1}(C)$ is a submanifold of N of codimension equal to the codimension of C in P.*

Singularities of spaces and mappings

Let us look at the case of $f : \mathbb{R}^n \to \mathbb{R}^p$, where $n > 0$ and $p > 0$. Loosely speaking, we know that if the differential of f has maximal rank at a point then we have a submersion or immersion; in the former the preimage of the value is a manifold and in the latter the image is a submanifold. Thus let us turn our attention to the case where the map does not have maximal rank.

Definition 2.8 We define a *singular point of f* to be a point x where d_xf has less than maximal rank.

Example 2.9 (i) Let $f : \mathbb{R}^n \to \mathbb{R}$ be given by $f(x) = \pm x_1^2 \pm x_2^2 \pm \cdots \pm x_n^2$. Then f is a *Morse singularity*. It is well-known that Morse proved that a function with a critical point such that the second differential is non-degenerate (equivalently, the square matrix of second derivatives is non-singular) is equivalent to such a Morse singularity (up to addition of a constant), see [48]. He also proved that such singularities are dense and stable – i.e., 'most' maps are Morse and they cannot be removed by perturbation.

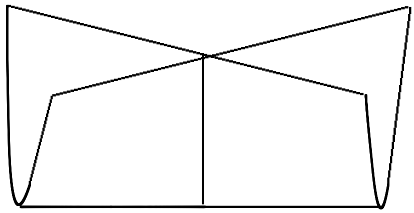


Figure 1: The Whitney Cross-Cap.

This classification has probably the profoundest effect in the theory of topology of manifolds as it leads to Morse Theory, of which more will be said later.

(ii) Let $f : \mathbb{R}^2 \to \mathbb{R}^3$ be given by $f(x,y) = (x, xy, y^2)$. This (and its image) is known as the *Whitney cross-cap* or *Whitney umbrella*. See Figure 1. This is also called the Whitney-Cayley cross-cap since it was known to Cayley. See [3] p217 for references and a discussion of his work in this area.

The singular set is the origin in $\mathbb{R}^2$. The image has singularities – points where the set is non-manifold – along the Z-axis in $\mathbb{R}^3$, and, apart from the origin, these are transverse crossing of manifolds.

(iii) The image of the Whitney cross cap is of codimension 1 in the codomain. If we attempt to find a polynomial that defines the image of f as a hypersurface we can try $h : \mathbb{R}^3 \to \mathbb{R}$ given by $h(X,Y,Z) = Y^2 - X^2 Z$. However, the image of f is actually a semi-algebraic set and so the zero-set of this h gives the image of f with a 'handle' that consists of the Z-axis. With this added to the image in Figure 1 one can see why the map is referred to as an umbrella.

If we work over the complex numbers rather than the reals, then the image and the zero-set descriptions coincide.

The last example shows the wide variety of behaviour that can occur for the image of a fairly simple algebraic map. One gets even stranger examples of if one considers images of smooth maps. For example, we can show that the corner of a cube can be produced from the image of an infinitely differentiable map.

Example 2.10 Let $\phi : \mathbb{R}^2 \to \mathbb{R}^+$ be the smooth map given by

$$\phi(x,y) = \begin{cases} e^{-1/x^2} & \text{for } x > 0 \\ 0 & \text{for } x \leq 0. \end{cases}$$

Let r_j be the map that rotates the plane about the origin through $\pi j/3$ radians. Let $\psi_1 = \phi \circ r_0 + \phi \circ r_1$, $\psi_2 = \phi \circ r_2 + \phi \circ r_3$ and $\psi_3 = \phi \circ r_4 + \phi \circ r_5$.

Then each ψ_i is a smooth function on the plane that is zero in a region bounded by two rays from the origin that are $2\pi/3$ radians apart. The interiors of these three regions do not overlap; the overlaps of the closures correspond to edge points of the corner of the cube.

The image of the map $h = (\psi_1, \psi_2, \psi_3)$ gives the corner of a cube.

Let us look at some more interesting examples, this time of less pathological singularities.

Example 2.11 (i) The simple singularities, A_k, D_k and E_k of maps $f : \mathbb{R}^2 \to \mathbb{R}$ are defined to be

$$\begin{array}{ll} A_k : & x^{k+1} + y^2, \text{ for } k \geq 1, \\ D_k : & x^2 y + y^{k-1}, \text{ for } k \geq 4, \\ E_6 : & x^3 + y^4, \\ E_7 : & x^3 + xy^3, \\ E_8 : & x^3 + y^5. \end{array}$$

The relation between the notation and that of simple Lie groups is not coincidental. See, for example, [2] page 99.

(ii) The Whitney cusp is given by $f : \mathbb{R}^2 \to \mathbb{R}^2$ is important in the study of maps between surfaces. It is given by $f(x, y) = (x, y^3 + xy)$. It is stable in the sense that if we perturb the map slightly, then there is some change of coordinates in the source and target that maps the perturbation back to the Whitney cusp.

The word ‘cusp’ is used in name of the last example because the *discriminant* of the map is a cusp.

Definition 2.12 Let $f : \mathbb{R}^n \to \mathbb{R}^p$ be a smooth with n and p arbitrary. Let C be the critical set of f, i.e., the points of $x \in \mathbb{R}^n$ such that rank $d_x f < p$.

Then the *discriminant* of f is $f(C)$, the image of the critical set.

Example 2.13 For the Whitney cusp the discriminant is diffeomorphic to the standard cusp, i.e., the image of $t \mapsto (t^2, t^3)$, and hence the name for the Whitney cusp.

The discriminant of the map is a very powerful invariant. It is a space which contains a lot of information concerning the map. In many cases given the discriminant it is possible to recover the map, see for example, [8] and [11].

3 Milnor fibre

We shall discuss later how classical Morse theory can be used to great effect in describing the topology of manifolds. For the moment we shall note that the essence of the theory is that the local behaviour of a function, in particular its singularities, is used to describe the global behaviour of the topology. In this section we consider the case of complex singularities, a field pioneered by the ancients (i.e., mathematicians pre-1900), but the approach is more recent and follows on from the seminal work of Milnor [49]. Whilst Milnor’s initial contribution cannot be underestimated it should be noted that many mathematicians have contributed to the theory – too many to do justice to in this paper. However, special mention should be made Lê Dũng Tráng, he has perhaps done more than any other to advance and popularize the theory of the Milnor fibre.

First, it should be noted that, in contrast to the case of real functions, for complex functions there is no local change in topology as one passes through a critical value. Instead one has to look at monodromy which is tackled in the next section. In this section we look at the local description of the fibres near to the singular fibre. The Milnor fibre is a fibre nearby to a singular fibre and is considered to be a local object, that is, we intersect with a neighbourhood.

The Milnor fibre of a complex function on a manifold

Let $f : \mathbb{C}^{n+1} \to \mathbb{C}$ be a non-constant complex analytic map such that $f(0) = 0$ and $n \geq 1$. We shall be interested in $f^{-1}(0)$, particularly at the origin in $\mathbb{C}^{n+1}$, and $f^{-1}(t)$ where $t > 0$ is small and $f^{-1}(t)$ is non-singular.

We restrict ourselves to local behaviour. Let us fix our notation for this. The sphere of radius ϵ centered at 0 in $\mathbb{C}^{n+1}$ is denoted S_ϵ; it bounds the closed ball B_ϵ; the open ball will be denoted B_ϵ°.

Our first result is the following:

Theorem 3.1 (Conic structure theorem) *There exists $\epsilon_0 > 0$ such that*

(i) $S_\epsilon \cap f^{-1}(0)$ *is homeomorphic to* $S_{\epsilon_0} \cap f^{-1}(0)$ *for all* $0 < \epsilon \leq \epsilon_0$*;*

(ii) $\mathrm{Cone}\left(S_{\epsilon_0} \cap f^{-1}(0)\right)$ *is homeomorphic to* $B_{\epsilon_0} \cap f^{-1}(0)$.

This Conic Structure Theorem holds for a far wider class of objects than just complex analytic sets, for example, Whitney stratified sets, a class we shall define later.

Definition 3.2 The space $L = S_\epsilon \cap f^{-1}(0)$ is called the *real link of f at* 0.

In his original ground-breaking text [49] Milnor showed this $(2n-1)$-dimensional space is $(n-2)$-connected. That is, $\pi_i(L) = 0$ for $0 \leq i \leq n-2$. (By convention, π_0 is trivial if and only if the space is path-connected). This was later improved by Hamm to complete intersections as discussed below.

We f has an *isolated singularity at* 0 if there exists an open neighbourhood U of 0 such that $U \cap f^{-1}(0)\backslash\{0\}$ is a manifold. Note that this includes the case that f is in fact non-singular at 0 – a standard, if perverse, use of terminology.

In the case that f has an isolated singularity, then K is a manifold. This has many interesting interpretations. For example, if $n = 1$, then the level set of f is a complex curve and so K is a knot (in fact a link, hence the name) in the 3-manifold S_ϵ (and hence in $\mathbb{R}^3$ as K does not fill the three manifold). Results relating these knots to analytic curves and vice versa were given in [49]. A short survey of more recent results can be found in [67].

If one goes to higher dimensions, then one can produce exotic spheres. Kervaire and Milnor showed in [34] that there exists manifolds homeomorphic to spheres which are not diffeomorphic to the standard differentiable structure on the sphere. These are called *exotic spheres*. Brieskorn gave the following example in [6].

Example 3.3 Let $f : \mathbb{C}^5 \to \mathbb{C}$ be given by

$$f(x, y, z, t, u) = x^2 + y^2 + z^2 + t^3 + u^{6k-1}.$$

Then, for $1 \leq k \leq 28$, the link of the origin of $f^{-1}(0)$ is a topological 7-sphere. Furthermore, these give the 28 different types of exotic 7-spheres.

The proof of this involves Smale's proof of the higher dimensional Poincaré conjecture and analysis of the monodromy of the singularity.

Thus, the link of a singularity is an interesting space in its own right. There is much to be investigated about it, particularly for surfaces and for non-isolated singularities.

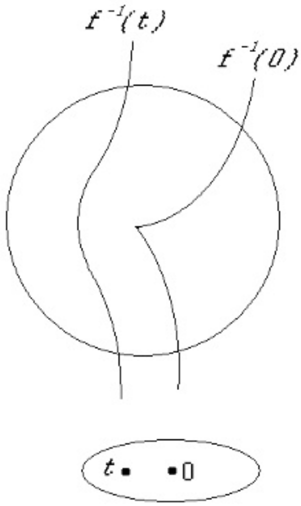


Figure 2: Schematic diagram of Milnor fibration

We now turn to the Milnor fibre of a singularity. One of the key results of [49] is the existence of a fibration connected with a neighbourhood of the singularity. Milnor originally defined this in a different fashion to the standard one about to be given, which is due to Lê in [38].

Theorem 3.4 (Milnor Fibration Theorem) *Let $f : \mathbb{C}^{n+1} \to \mathbb{C}$ be a complex analytic map and ϵ is taken small enough so that S_ϵ defines the real link of f. Let D^*_δ be the set $\{t \in \mathbb{C} : 0 < |t| < \delta\}$.*

Then, $f : B^\circ_\epsilon \cap f^{-1}(D^*_\delta) \to D^*_\delta$ is a smooth locally trivial fibration for $0 < \delta << \epsilon$.

In fact, $f : B_\epsilon \cap f^{-1}(D^*_\delta) \to D^*_\delta$ is locally trivial topological fibration with $f : S_\epsilon \cap f^{-1}(D^*_\delta) \to D^*_\delta$ a subfibration.

Furthermore, if f has an isolated singularity, then this subfibration extends over $D_\delta = \{t \in \mathbb{C} : |t| < \delta\}$.

Unsurprisingly, the proof of the first part of this involves the Ehresmann Fibration Theorem. The second can be proved using the First Thom–Mather Isotopy Lemma which will be discussed later.

Definition 3.5 The fibre of $f : B^\circ_\epsilon \cap f^{-1}(D^*_\delta) \to D^*_\delta$ is called the *Milnor fibre of f* and is denoted by F°_f, or F° if no confusion will result. The closed fibre is F_f or F. The boundary of F is denoted ∂F and, as can be seen from the theorem, this is also the fibre of a fibration.

A schematic picture of the fibrations is given in Figure 2. It should be noted that because we can collar ∂F in F that F° and F are homotopically equivalent and are effectively interchangeable in many theorems.

Example 3.6 (i) Recall that $f(x, y, z) = y^2 - x^2 z$ defines the Whitney Umbrella of Example 2.9(ii). The Milnor fibre is homotopically equivalent to a sphere, see Example 3.16.

(ii) Let $f(x_1, x_2, \dots, x_n, x_{n-1}) = x_1 x_2 \dots x_{n+1}$. Then, it is easy to calculate that the Milnor fibre of f is homotopically equivalent to $(S^1)^{n+1}$.

One of the reasons that the Milnor fibre is such a useful construction is it is a topological invariant in the following sense.

Definition 3.7 Two functions f and g from $\mathbb{C}^{n+1}$ to $\mathbb{C}$ with $f(0) = g(0) = 0$ are *topologically equivalent at* 0 if there exists a homeomorphism $h : U \to V$, where $U, V \subseteq \mathbb{C}^{n+1}$ are open sets containing the origin, such that $f = g \circ h$.

We have the following crucial theorem.

Theorem 3.8 ([36]) *Suppose that f and g are topologically equivalent. Then, their Milnor fibres are homotopically equivalent.*

Since the Milnor fibre is a Stein space, by Hamm [22] (see [17] for a simpler proof), it has the homotopy type of a CW-complex of real dimension equal to its complex dimension, i.e., n. This dimension is called *the middle dimension*. Thus we can place an upper bound on the non-vanishing of homology. Also, we can place a lower bound on the non-vanishing of reduced homology groups of the Milnor fibre, and in fact can do this for homotopy groups.

Proposition 3.9 (Kato-Matsumoto [33]) *If the singular set of f at 0 has dimension s, then F_f is $(n - s - 1)$-connected.*

From Example 3.6(ii) we see that this bound is in some sense sharp. However, the converse of the theorem is not true and so one would like a more accurate statement. This has proven difficult to find.

Also, we do not have many general theorems describing the homology groups between $n - s - 2$ and n, and so this is an area requiring more research. However, one example of such a theorem is given by Némethi [51], where some seriously heavy topological work is employed to the case of compositions of functions, that is to maps $f : \mathbb{C}^{n+1} \to \mathbb{C}^2$ defining complete intersections with an isolated singularity and curve singularities $g : \mathbb{C}^2 \to \mathbb{C}$. Let us define these.

Example 3.10 Let $f : \mathbb{C}^{n+1} \to \mathbb{C}^2$ be such that $f^{-1}(0)$ is of dimension $n - 1$ and has an isolated singularity at the origin in $\mathbb{C}^{n+1}$. Suppose that $g : \mathbb{C}^2 \to \mathbb{C}$ defines a reduced curve singularity. Then $h = g \circ f$ is called a *Generalized Zariski singularity*.

In this case, the singular set of h coincides with the fibre of f and so has codimension 1 in the fibre of h. Then, $H_*(F, \mathbb{Z}) = 0$ for $* \neq 0,\ 1,\ n$. See [51], [52] and, for more recent work, [25].

Isolated singularities and the unreasonable effectiveness of the Milnor number

An important corollary of Proposition 3.9 was first proved by Milnor.

Corollary 3.11 (Milnor, [49]) *If f has an isolated singularity at 0, then F_f has the homotopy type of wedge of spheres of dimension n.*

Recall that the wedge of two topological spaces is their one point union. Hence, a *wedge of spheres* (also known as a *bouquet of spheres*) is a collection of spheres, each member of which has a special point identified to the special point on the other members. An interesting feature of investigations of the local behaviour of Milnor fibres is the appearance of many results involving the wedge of spaces. This will be seen more clearly in Theorem 5.11.

Definition 3.12 The number of the spheres in the wedge is called the *Milnor number* of f at 0, and is denoted $\mu(f)$.

The Milnor number is a surprisingly effective topological invariant. Probably, one reason for this is that it can be calculated algebraically with ease: Suppose that coordinates on $\mathbb{C}^{n+1}$ are given by $x_1, x_2, \ldots, x_{n+1}$ and that $\mathbb{C}\{x_1, x_2, \ldots, x_{n+1}\}$ denotes the ring of convergent power series at 0. Then, Milnor [49] showed that

$$\mu(f) = \dim_{\mathbb{C}} \frac{\mathbb{C}\{x_1, x_2, \ldots, x_{n+1}\}}{\left\langle \dfrac{\partial f}{\partial x_1}, \ldots, \dfrac{\partial f}{\partial x_{n+1}} \right\rangle}.$$

The ideal in the denominator is called the *Jacobian ideal of* f and the quotient ring is called the *Milnor algebra of* f.

Example 3.13 Consider the case of a Morse singularity, i.e.,

$$f(x_1, \ldots, x_{n+1}) = x_1^2 + \cdots + x_{n+1}^2 + c.$$

In this case the Milnor number is easily seen to be equal to 1.

Since we can produce knots via the real link of a singularity this means that we can use μ as an invariant of knots constructed in this way.

Furthermore, $\mu(f)$ is related to the unfolding properties of f. On the space of complex analytic functions we can place a natural equivalence relation.

Definition 3.14 Two functions f and g are *right-equivalent* if there exists a biholomorphism $h : U \to V$ of open sets U and V in $\mathbb{C}^{n+1}$ such that $f = g \circ h$.

The orbit of a function can be referred to as a singularity type since all functions in the orbit have the same type of singularity.

The codimension of the orbit in the space of all complex analytic functions is $\mu(f)$. By taking a set of functions $\{\alpha_i(x)\}_{i=1}^{\mu(f)}$ that projects to a basis of the $\mathbb{C}$-vector space of the Milnor algebra of f we can produce an *unfolding* of f:

$$F(x, \lambda) = f(x) + \sum_{i=1}^{\mu(f)} \lambda_i \alpha_i(x).$$

The idea of this concept is that (up to isomorphism) this family of functions contains all functions near to f.

Thus, it can be seen that μ is linked closely to the topology of the Milnor fibre and that it measures how complicated the singularity is by measuring how 'deep' within the space of functions the singularity sits.

New Milnor fibres from old

As usual in mathematics one would like to construct new examples of objects from old examples in such a way that the properties of the new can be calculated from that of the old.

In this vein is the very useful following theorem first investigated by Sebastiani and Thom and proved in more generality by Sakamoto in [55]. First recall that the *join* of two topological spaces X and Y is denoted by $X * Y$ and is defined as $X \times [0,1] \times Y$ with the following identifications:

(i) $(x, 0, y) \sim (x', 0, y)$ for all $x, x' \in X$ and $y \in Y$,

(ii) $(x, 1, y) \sim (x, 1, y')$ for all $x \in X$ and $y, y' \in Y$.

Proposition 3.15 (Sebastiani-Thom Theorem, [58], [55]) *Let $f : \mathbb{C}^r \to \mathbb{C}$ and $g : \mathbb{C}^s \to \mathbb{C}$ be complex analytic maps with $f(0) = g(0) = 0$. Define*

$$f \oplus g : \mathbb{C}^{r+s} \to \mathbb{C} \text{ by } (f \oplus g)(x, y) = f(x) + g(y).$$

Then,

$$F_{f+g} \text{ is homotopically equivalent to } F_f * F_g.$$

Consequently, $\mu(f \oplus g) = \mu(f)\mu(g)$.

There are many generalizations of this theorem to different settings. For example, [45] shows how to generalise the underlying isomorphism to one in the derived category for very general singular functions. This paper also includes a number of references to other results in the area.

The concept of this result has an interesting application in Stratified Morse Theory when considering the splitting of local Morse data into local Normal and Tangential Morse Data, see Section 6.

Example 3.16 In the complex setting the Whitney umbrella is the image of $f : \mathbb{C}^2 \to \mathbb{C}^3$ be given by $f(x, y) = (x, xy, y^2)$. This set is also given as a zero-set using $h : \mathbb{C}^3 \to \mathbb{C}$ defined by $h(X, Y, Z) = Y^2 - X^2 Z$. From the Sebastiani-Thom result we can see that the Milnor fibre of h is the suspension of the Milnor fibre of $g(X, Z) = X^2 Z$. Since, g is homogeneous, its Milnor fibre is given by the solution set of $X^2 Z = 1$ and so can be given by

$$\left(X, \frac{1}{X^2} \right) \text{ for } X \neq 0.$$

The fibre is thus homeomorphic to $\mathbb{C} \backslash \{0\}$ which is in turn homotopically equivalent to a circle. Therefore, the Milnor fibre of the Whitney umbrella is homotopically equivalent to the suspension of a circle, i.e., a 2-sphere.

Complete intersections

Much of the preceding theory was generalized from functions to maps. Let $f : \mathbb{C}^{n+r} \to \mathbb{C}^r$ be a complex analytic map such that $f(0) = 0$ and, for simplicity, assume that $n, r \geq 1$. Again, we can define the link of $f^{-1}(0)$ as $S_\epsilon(0) \cap f^{-1}(0)$ for small enough ϵ. Hamm's theorem in [21] states that it is at least $(n-2)$-connected.

However, as we have no guarantee that a non-constant map will give a fibre of dimension n, we restrict to the case that $f^{-1}(0)$ has dimension n (the dimension we would expect in generic situations) and that it has an isolated singularity at the origin. In this case we say that f defines an *isolated complete intersection singularity* which is traditionally abbreviated to ICIS.

Again we find a Milnor fibration over the non-critical values in the target. As we assume that the singularity is isolated, we have a bouquet theorem: The Milnor fibre of an ICIS is a bouquet of n-spheres. It is quite common to find these bouquet theorems, we will see a reason for this in Theorem 5.11.

The number of spheres is again called the *Milnor number*. This is harder to calculate than the function case as it cannot be described as the quotient of some relatively simple ideal (or module); it can however be given be calculated quite effectively in low codimension by an alternating sum of numbers, see [37] or [41] page 76-77. The latter book is the standard reference for isolated complete intersection singularities.

For an ICIS one can define the *Tjurina number* as the dimension of the space

$$\tau(f) = \dim_{\mathbb{C}} \frac{(\mathbb{C}\{x_1, x_2, \dots, x_{n+r}\})^r}{\{\langle f_1, f_2, \dots, f_r\rangle e_i\}_{i=1}^r + \left\langle \dfrac{\partial f}{\partial x_1}, \dots, \dfrac{\partial f}{\partial x_{n+1}} \right\rangle}$$

where e_i is the standard basis, i.e., a column vector with a 1 in position i and 0 elsewhere. A basis of the space used in the definition of τ can be used to construct an unfolding of f that has all nearby functions up to $\mathcal{K}$-equivalence. See [66].

In the case of function $f : \mathbb{C}^{n+1} \to \mathbb{C}$ with an isolated singularity we see that

$$\tau(f) = \dim_{\mathbb{C}} \frac{\mathbb{C}\{x_1, x_2, \dots, x_{n+1}\}}{\left\langle f, \dfrac{\partial f}{\partial x_1}, \dots, \dfrac{\partial f}{\partial x_{n+1}} \right\rangle}.$$

It is easy to see that in this case that $\tau(f) \leq \mu(f)$ with equality if f is quasi-homogeneous.

Theorem 3.17 *For an ICIS we have $\tau \leq \mu$ with equality if f is quasi-homogeneous.*

The proof of this is considerably harder than the function version and relies on Hodge theory, see [42]. In one respect it is unsatisfactory in that it does not provide much insight into why the result is true, however, in 20 years no-one has improved upon the proof.

4 Monodromy

The Milnor fibration is a fibration over a punctured disc and so for a moment let us consider it as a fibration over the circle S^1. As this fibration is locally trivial we can consider what happens to a point $x \in f^{-1}(t)$ for any $t \in S^1$ as we go round a loop in S^1. Executing a complete loop gives a homeomorphism from F to F, which, in general, is not the identity.

Definition 4.1 The homeomorphism $h : F \to F$ is the *geometric monodromy of f*. The map induced on homology

$$h_* : H_*(F; \mathbb{Z}) \to H_*(F; \mathbb{Z})$$

is called the *(classical) monodromy operator of f*.

We can prove the existence of this map by taking a vector field on S^1 and using the Ehresmann Fibration Theorem to produce a corresponding vector field on the total fibre space. The geometric monodromy is then given by integrating this vector field.

Now consider the compact set F and its boundary ∂F. The monodromy can be chosen to be the identity on the boundary. So there exists a map from the relative pair $(F, \partial F)$ to F. Let c be a relative cycle, then, since h is the identity on ∂F, c and $h(c)$ have the same boundary and thus $h(c) - c$ is a cycle on F.

Definition 4.2 The *variation operator of f* is the map

$$\mathrm{var}_* : H_*(F, \partial F; \mathbb{Z}) \to H_*(F; \mathbb{Z}).$$

Let $i : (F, \emptyset) \to (F, \partial F)$ be the standard inclusion, then we have a commutative diagram

$$\begin{array}{ccc} H_*(F;\mathbb{Z}) & \stackrel{h_*-id}{\longrightarrow} & H_*(F;\mathbb{Z}) \\ \downarrow i_* & \stackrel{\mathrm{var}_*}{\nearrow} & \downarrow i_* \\ H_*(F,\partial F;\mathbb{Z}) & \stackrel{h_*-id}{\longrightarrow} & H_*(F,\partial F;\mathbb{Z}) \end{array}$$

This means that we can describe the monodromy through the variation operator:

$$h_* = id + \mathrm{var}_* \circ i_*$$

Suppose that f has an isolated singularity, then we know from Corollary 3.11 that the Milnor Fibre is a wedge of spheres of dimension n and so we can concentrate on $H_n(F; \mathbb{Z})$. A particular case is that of a quadratic singularity:

Example 4.3 Consider a Morse singularity, $f(x_1, \dots, x_{n+1}) = x_1^2 + \cdots + x_{n+1}^2$. We have seen that $\mu(f) = 1$ and so $H_n(F; \mathbb{Z}) \cong \mathbb{Z}$. In fact, one can calculate that the Milnor fibre is diffeomorphic to the tangent bundle of S^n. The construction is given explicitly in [2] and [41].

In this example we can consider the set $B_\epsilon(0) \cap f^{-1}(t)$ as $t \to 0$. We see that as $B_\epsilon(0) \cap f^{-1}(0)$ is a cone over its boundary, and hence is contractible, that the homology 'vanishes' as $t \to 0$.

Definition 4.4 The non-trivial homology class in $H_n(F; \mathbb{Z})$ of Example 4.3 is called *the vanishing cycle* of f.

Continuing this example, if we let ∇ denote the non-trivial class of $H_n(F, \partial F; \mathbb{Z})$ and Δ be vanishing cycle, then Picard and Lefschetz proved the following.

Theorem 4.5 ([40], [53]) *For a Morse singularity we have*

$$\mathrm{var}(\nabla) = (-1)^{n(n+1)/2} \Delta.$$

Since this theorem was first proved, the theory of vanishing cycles has been greatly developed and reframed in terms of a sheaf of vanishing cycles, see [9] for the original sheaf version and [10] and [57] for more modern exposition.

We can now generalize to more general functions. Suppose that f has an isolated singularity at $0 \in \mathbb{C}^{n+1}$. Then for a generic linear function $g : \mathbb{C}^{n+1} \to \mathbb{C}$ the function $f_\lambda : \mathbb{C}^{n+1} \to \mathbb{C}$ defined by $f_\lambda(x) = f(x) + \lambda g(x)$ will have only Morse singularities (with distinct critical values) for $0 \neq \lambda \in \mathbb{R}$.

Fix such a λ. Within the Milnor ball $B_\epsilon(0)$ we have a finite number of critical points of f_λ and this number is equal to $\mu(f)$, denote the points by $p_1, \dots, p_\mu$. Let $c_i = f_\lambda(p_i)$.

In $\mathbb{C}$ we have a disc $D_\delta = \{z \in \mathbb{C} : |z| \leq \delta\}$. Over $D' = D_\delta \backslash \{c_1, \ldots, c_\mu\}$ we in fact have a fibration which has its (open) fibre diffeomorphic to the Milnor fibre of f. (We use the Ehresmann fibration theorem again!) Basically, each critical point is Morse and so it will contribute one copy of $\mathbb{Z}$ to the homology of the Milnor fibre. Furthermore, the monodromy of each will determine the monodromy of f.

Let c_* be a point on the boundary of D_δ. Then, for each loop γ in D' based at c_* we have a monodromy $h_\gamma : H_n(F;\mathbb{Z}) \to H_n(F;\mathbb{Z})$. We get a homomorphism from $\pi_1(D')$ to $\operatorname{Aut} H_n(F;\mathbb{Z})$, the group of automorphisms of $H_n(F;\mathbb{Z})$.

Definition 4.6 The *monodromy group* of f is the image of the above homomorphism.

Let Δ_i be the vanishing cycle associated to the critical point c_i.

Proposition 4.7 *The cycles $\Delta_1, \ldots, \Delta_\mu$ form a basis of $H_n(F;\mathbb{Z})$.*

We can now define an intersection pairing on $H_n(F;\mathbb{Z})$. Suppose that N is an oriented compact $2n$-manifold with boundary ∂N such that the integer homology and cohomology of N has no torsion.

There exists the Poincaré duality map $\mu : H_n(N, \partial N) \to H^n(N)$ and we also have the standard inclusion map $i : (N, \emptyset) \to (N, \partial N)$ which induces a homomorphism from $H_n(N)$ to $H_n(N, \partial N)$.

By identifying $H^n(N)$ and the dual of $H_n(N)$ we have a pairing

$$\operatorname{ev}(\ ,\ ) : H^n(N) \times H_n(N) \to \mathbb{Z}$$

given by the usual evaluation for a space and its dual.

Thus, we can define the following.

Definition 4.8 The *intersection pairing/form* is the map $\langle\ ,\ \rangle : H_n(N) \times H_n(N) \to \mathbb{Z}$ given by

$$\langle x, y\rangle = \operatorname{ev}(\mu(i_*(x)), y).$$

Proposition 4.9 ([2]) *We have*

$$\langle \Delta_i, \Delta_i\rangle = \begin{cases} 0, & n \textit{ odd}, \\ (-1)^{n(n-1)/2}2, & n \textit{ even}. \end{cases}$$

Definition 4.10 The matrix $B = (\langle \Delta_i, \Delta_j\rangle)_{1\leq i,j\leq\mu}$ is called the *intersection matrix of f*.

From this matrix we can determine much about the monodromy of the singularity. We can read off results about the classical monodromy operator h_* and the variation operator var since we have $h_* = h_1 h_2 \cdots h_\mu$ where

$$h_i(x) = x - (-1)^{n(n-1)/2}\langle x, \Delta_i\rangle\Delta_i$$

for Δ_i a basis element. See [2] for details.

We can also describe the monodromy of direct sums of singularities:

Theorem 4.11 *For f and g with isolated singularities we can describe the monodromy and variation operators:*

$$\begin{aligned} h_{f\oplus g*} &= h_{f*} \otimes h_{g*}, \\ \operatorname{var}_{f\oplus g*} &= \operatorname{var}_{f*} \otimes \operatorname{var}_{g*} \end{aligned}$$

The theory in the case of non-isolated singularities is more complicated as one would expect. A survey can be found in [59]. However, much of the theory carries over to functions with isolated singularities on singular spaces. Details can be found in [64] and [65].

5 Stratifications of spaces

When dealing with singular spaces (and maps) one is usually faced with a choice of two methods. The first and perhaps oldest of these is to find a manifold associated to the singular space and some map between the two. A study of the manifold and the associated map will indirectly reveal information about the singular space. This is called the resolution method. The summit of achievement here is Hironaka's famous resolution theorem which relies heavily on algebraic constructions and would take us too far from our interests.

The second method in some sense gets us closer to the singularities but still relies on using the theory of manifolds. Basically, the space is partitioned into manifold subsets, (i.e., each manifold constitutes a subset). This is called the stratification method. Different conditions on how the manifolds meet one another give rise to different types of stratification, e.g., Whitney (the most common which we shall describe below), Bekka, A_f, and logarithmic. See [63] for how the various stratifications are related to each other.

Definition 5.1 Let X be a closed subset of a smooth manifold M and let X be decomposed into disjoint pieces S_i called strata. Then the decomposition is called a *Whitney Stratification* if the following conditions are met.

(i) Each stratum is a locally closed smooth submanifold of M.

(ii) $S_i \cap \mathrm{Closure}(S_j) \neq \emptyset$ if and only if $S_i \subseteq \mathrm{Closure}(S_j)$ for strata S_i, S_j $i \neq j$; this is called the *frontier condition* and we write $S_i < S_j$.

(iii) Whitney Condition (a): If $x_i \in S_a$ is a sequence of points converging to $y \in S_b$ and $T_{x_i}(S_a)$ converges to a plane τ (all this considered in the appropriate Grassmannian), then $T_y(S_b) \subseteq \tau$.

(iv) Whitney Condition (b): If $x_i \in S_a$ converges to $y \in S_b$, $y_i \in S_b$ also converges to y, l_i denotes the secant line between x_i and y_i, and l_i converges to l, then $l \subseteq \tau$.

Note that (b)⇒(a). The condition (a) is stated explicitly since it is a condition that many stratifications satisfy and is thus very useful. If X is a closed subanalytic subset of an analytic manifold, then X can be Whitney stratified, hence complex varieties, semi-analytic spaces, etc., can be Whitney stratified.

In Figure 3 we can see that we can stratify the space by taking the z-axis as a stratum. However, this is not a Whitney stratification. The singular point that is not a transverse crossing has to be a stratum. To see this consider a family of horizontal lines as in the figure. This will converge to a line perpendicular to the z-axis so violates the (b) condition.

A Whitney stratified space can be triangulated (see [16]) and is locally topologically trivial along the strata. Also, one has a conic structure theorem like Theorem 3.1: For any point x and a small enough sphere S_ϵ centred at x, the cone of $X \cap S_\epsilon$ is homeomorphic to $X \cap B_\epsilon$.

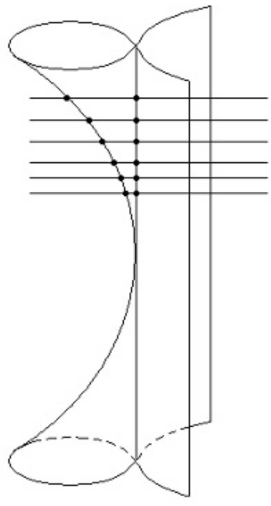


Figure 3: A space to Whitney stratify

The local topological triviality and the conic structure theorem both follow from the First Thom-Mather Isotopy Lemma. Perhaps, the most important lemmas in the applications of stratification theory are the Thom-Mather Isotopy Lemmas. The first concerns spaces and can be considered a direct generalization of the Ehresmann Fibration Theorem. Recall that the idea of the latter is that, at a non-critical value, a map on a manifold is in fact a fibration. The First Thom-Mather Isotopy Lemma essentially requires that the map is submersion on all the strata to produce a stratum preserving homeomorphism. Thus, it is very useful in proving that various spaces are homeomorphic.

The second lemma allows us to fibre certain mappings: it gives sufficient conditions for maps in a family of maps over $\mathbb{R}^p$ to be topologically right-left equivalent to one another. (Two maps $f : \mathbb{R}^n \to \mathbb{R}^p$ and $g : \mathbb{R}^n \to \mathbb{R}^p$ are *topologically right-left equivalent* if there exist homeomorphisms $h : \mathbb{R}^n \to \mathbb{R}^n$ and $k : \mathbb{R}^p \to \mathbb{R}^p$ such that $f \circ h = k \circ g$.)

Proof of both lemmas for Whitney spaces can be found in [15] and in the unpublished manuscript [46]. For Bekka (also known as (c)-regular) stratifications they are proved in [5]. We now outline the basic idea of proof. In differential topology we can find homeomorphisms by integrating vector fields on manifolds. For stratified spaces we find conditions so that we can define vector fields on the strata that when integrated give the required homeomorphisms. The surprising fact is that the vector fields do not have to be continuous when considered as a whole on the Whitney space. The method is very technical and the proof of the theorem occupies about a quarter of [15]. The original - unpublished - version is still readable, see [46]. More modern versions with some extra generalizations can be found in [54] and [63].

First Thom-Mather Isotopy Lemma

Suppose M and P are analytic manifolds. Let $X \subseteq M$ be a Whitney stratified subset.

Definition 5.2 A map $f : X \to P$ is a stratified submersion if $f|A$ is a submersion for all strata A in X.

Theorem 5.3 (Thom-Mather First Isotopy Lemma) *Let $f : X \to P$ be a proper stratified submersion. Then f is a locally trivial fibration such that the homeomorphisms are stratum preserving.*

As an application we can prove a Milnor fibration type theorem, first proved by Lê in [39]. We need a definition from there which clarifies what we mean by a function on a singular space having a singularity.

Definition 5.4 Suppose that $f : X \to \mathbb{C}$ is a complex analytic function with a Whitney stratification $\mathcal{S}$ and that X can be locally embedded in $\mathbb{C}^n$ at $x \in X$ for some n.

We say that f has an *isolated stratified singularity at x* if there exists $\epsilon > 0$ such that $f|B^\circ_\epsilon(x) \cap (A\backslash\{x\}) \to \mathbb{C}$ is a submersion for all $A \in \mathcal{S}$ where $B^\circ_\epsilon(x)$ is an open ball in $\mathbb{C}^n$ of radius ϵ centred at x.

We can now produce the general Milnor fibration set-up for singular spaces.

Proposition 5.5 ([39]) *Suppose X is a complex analytic space and $f : X \to \mathbb{C}$ has an isolated stratified singularity at $x \in f^{-1}(0)$. Then, there exists an embedding of X into $\mathbb{C}^n$ and real numbers ϵ and δ such that the map*

$$\phi : B_\epsilon(x) \cap X \cap f^{-1}(D^*_\delta) \to D^*_\delta$$

induced by f is a locally trivial fibration over the punctured disc $D^*_\delta = \{y \in \mathbb{C} : \ 0 < |y| < \delta\}\backslash\{0\}$. Here $B_\epsilon(x)$ is a small open or closed ball about x of radius ϵ in $\mathbb{C}^n$. In the closed case we produce a fibration on the boundary of $B_\epsilon(x) \cap X \cap f^{-1}(D^*_\delta)$ as well.

The fibration is sometimes referred to as the Milnor fibration. Just as in the case of a function on a complex manifold, often the distinction between the open and closed version of the fibres is blurred since the two spaces are homotopically equivalent.

The complex link

In the case where f is a general (complex) linear function, one gets the *complex link of x*. This is a very powerful space which contains a lot about the behaviour of the set X near to x.

Definition 5.6 Let x be a point of $X \subseteq \mathbb{C}^n$ a complex analytic set. Then, for a suitably generic linear function $L : \mathbb{C}^n \to \mathbb{C}$ the *complex link* of x is

$$\mathcal{L}_x = B_\epsilon(x) \cap X \cap L^{-1}(t)$$

where $t \neq 0$ and ϵ are sufficiently small.

Now take a manifold N transverse to the stratum A containing x such that $N \cap X = \{x\}$.

Definition 5.7 The complex link of x in the set $X \cap N$ is called the *complex link of the stratum A* and is denoted $\mathcal{L}_A$.

As one might imagine the topological type of $\mathcal{L}_A$ does not depend on the choice of $x \in A$ and the choices of ϵ and t when these are small.

For complex spaces, the complex link of strata turns out to be the fundamental building blocks of the space and analysis of them is vital to the use of Stratified Morse Theory in Section 6.

Few results are known that describe the complex link of strata in general settings. The following is very useful and quite general. This was proved by Lê following Hamm's work for the similar situation of real links.

Theorem 5.8 *Suppose that X is defined at x in $\mathbb{C}^n$ by r equations. Then, the complex link of the stratum containing x is $(n-r-2)$-connected.*

Corollary 5.9 *Suppose that X is a local complete intersection at x, i.e., the number of defining equations equals the codimension in $\mathbb{C}^n$ of X at x, then the complex link of the stratum A is homotopically equivalent to a wedge of spheres of dimension* $\dim X - \dim A - 1$.

There are further examples.

Examples 5.10 (i) Suppose that $f : N \to P$ is a complex analytic map between complex manifolds such that $\dim N < \dim P$ and that f is stable everywhere (see [66]) and the corank of the differential is less than or equal to 1 at all points in N. (In this case we say that f has corank 1 even though we include the case that the corank may be zero.)

We can stratify the image by stable type and this is the canonical Whitney stratification, see [14]. Then the complex link of a stratum A in the image is homotopically equivalent to a single sphere of dimension $k \dim N - (k-1)\dim P - \dim A - 1$ where k is determined precisely by the stable type of the germ at any point of A. See [29].

(ii) Let $F : \mathbb{C}^2 \to \mathbb{C}^4$ be defined by $F(x,t) = (x^2, x^3 + tx, tx^3, t)$. The image of this map has an isolated singularity at the origin of $\mathbb{C}^4$ and is defined in $\mathbb{C}^4$ by no fewer than four equations, so it is not a complete intersection. The link $\mathcal{L}$ of the origin is homeomorphic to the image of the map $f_t(x) = F(x,t)$ for any $t \neq 0$. Since, for $t \neq 0$, f_t is a proper injective immersion the image is homeomorphic to $\mathbb{C}^2$, which is contractible.

(iii) Suppose that X_n is the complex analytic set in the set of $2 \times n$ matrices given by the matrices of rank 1. Then X_n is $n+1$ dimensional, embedded in $\mathbb{C}^{2n}$ and has an isolated singularity: the matrix of rank zero. The set X_n can be seen to be an analytic set in $\mathbb{C}^{2n}$ via taking all the 2×2 minors of the following matrix, where coordinates on $\mathbb{C}^{2n}$ are given by z_i,

$$\begin{bmatrix} z_1 & z_2 & \cdots & z_n \\ z_{n+1} & z_{n+2} & \cdots & z_{2n} \end{bmatrix}.$$

Using this description it is possible to calculate that the complex link of the origin is homeomorphic to $\mathbb{CP}^n$.

We have seen that wedge-of-spheres theorems occur regularly for Milnor fibres and we will see in the next section that complex links are the building blocks of complex analytic spaces. We now show a general theorem that describes the Milnor fibre on a singular space in terms of a wedge of suspensions of complex links, see [61]. Hence in the case where the complex links are wedges of spheres, for example, X is a complete intersection, we can deduce the Milnor fibre is a wedge of spheres.

First recall that the *suspension* of a topological space Z, denoted ΣZ is the join of Z and two disjoint points. The repetition of this process is denoted $\Sigma^k Z$, where $\Sigma^1(Z) = \Sigma Z$.

Theorem 5.11 (General wedge theorem, [61]) *Let X be a complex analytic space with a Whitney stratification such that $f : X \to \mathbb{C}$ has an isolated stratified singularity.*

Then F_f is homotopically equivalent to

$$\bigvee_{A} \bigvee_{\mu_A(f)} \Sigma^{\dim_{\mathbb{C}} A} \mathcal{L}_A$$

where A runs all over all strata A such that $x \in \overline{A}$, the closure of A, and $\mu_A(f)$ is the number of copies to be taken (which depends on f).

This then encompasses many of the theorems we have met. For example, if $X = \mathbb{C}^{n+1}$ and f has an isolated singularity, then the complex links are empty and $\Sigma^{n+1}\emptyset = S^n$ (because $\Sigma\emptyset = S^0$). Thus, we recover the original Milnor fibre theorem, Corollary 3.11.

Second Thom-Mather Isotopy Lemma

The second isotopy lemma is a relative version of the first, rather than using it to fibre a singular space, we fibre a map between two singular spaces.

Suppose M, N and P are analytic manifolds. Let $X \subseteq M$ and $Y \subseteq N$ be Whitney stratified subsets.

Definition 5.12 Suppose X and Y are Whitney stratified spaces in the analytic manifolds M and N and $f : X \to Y$ is the restriction of a smooth map $F : M \to N$. Then, the map is called *stratified* if f is proper and if for any stratum $A \subseteq Y$ the preimage $f^{-1}(A)$ is a union of strata and f takes these strata submersively to A.

If X and Y are complex analytic and f is complex analytic, then it is possible to stratify X and Y into complex analytic strata so that f is stratified. See [17] I.1.7.

Just as it was necessary to impose further conditions on the strata of the space to get the First Isotopy Lemma we need to impose another important condition on the map for the Second. This is the Thom A_f condition:

Definition 5.13 Let $f : X \to Y$ be a stratified map. Then f is a *Thom A_f map* if for every pair of strata $B < A$ we have the following.

(i) $f|A$ and $f|B$ have constant rank.

(ii) A is Thom regular over B: If $a_i \in A$ is a sequence of points converging to $b \in B$ such that $\ker d_{a_i}(f|A)$ converges to a plane T then $\ker d_b(f|B) \subseteq T$.

This condition is important in its own right. The definition of (c)-regularity (or Bekka stratifications) involves the Thom condition, see [5]. It is generally thought that this type of stratification, which includes Whitney stratifications, is the 'correct' type of stratification for the study of maps between spaces.

Theorem 5.14 (Thom-Mather Second Isotopy Lemma) *Let $F : X \to Y$ be a proper Thom A_F map and let $f : Y \to P$ be a proper stratified submersion. Then $F : X \to Y$ is locally topologically trivial over P.*

That is, for every point $p \in P$ there exists a neighbourhood V of p, such that for every $q \in V$, $F : (f \circ F)^{-1}(q) \to f^{-1}(q)$ and $F : (f \circ F)^{-1}(p) \to f^{-1}(p)$ are topologically right-left equivalent by stratum preserving homeomorphisms.

The Thom-Mather Isotopy Lemmas are important ingredients in the study of the topological stability of maps, see [15], and for more recent progress and a highly detailed exposition see [12].

6 Stratified Morse theory

Classical Morse theory

Morse theory has a long history, going back before it even acquired that name. It is fairly obvious that given a topological space X and a continuous map $f : X \to \mathbb{R}$ then we can study the topology of X by seeing how it changes as we take the preimages under f of the set $(-\infty, a]$. Let $X_a = f^{-1}([-\infty, a])$; we can define a critical value v of f to be one such that $X_{v-\varepsilon}$ is not homeomorphic to $X_{v+\varepsilon}$, for any small ε. The idea is to find a suitable f so that, for instance, we have a finite number of critical values and can find out what happens as we pass them.

Of course such a situation is too general and we have to restrict to situations where X has some extra structure, for instance a nonsingular projective complex algebraic curve. According to Fulton in [13] Riemann had the following theorem: Suppose $f : X \to \mathbb{CP}^1$ is a meromorphic function on a smooth projective curve with n sheets and w simple branch points, then the genus g of X is given by $w = 2g + 2n - 2$.

Various generalizations were given until Morse eventually arrived at the following theorem:

Theorem 6.1 *Suppose $f : X \to \mathbb{R}$ is a 'sufficiently general' C^∞ function on the compact real manifold X. Then the topology of X_a changes only when we pass a critical value, which in this case is the image of a point where the differential of f is zero. Furthermore $X_{v+\varepsilon}$ is homotopically equivalent to $X_{v-\varepsilon}$ with a cell attached and the dimension of this cell is equal to the dimension of the space upon which the Hessian of f is negative definite (this is called the* index *of f).*

This is a truly great result, its effectiveness in the 20th Century can be seen in Smale's work on h-cobordism and the higher dimensional Poincaré conjecture, René Thom's work, Bott periodicity, etc. Classical Morse theory was explained very well in Milnor's classic book [48] and so I shall not explain it further here.

A stratified version

Now, we can easily ask if such a theory is possible for singular spaces rather than manifolds. Goresky and Macpherson answered this positively with their *Stratified Morse Theory*.

The idea here is that we stratify our singular space into manifolds and put a function on the space so that function is a Morse function on the manifolds and so that the func-

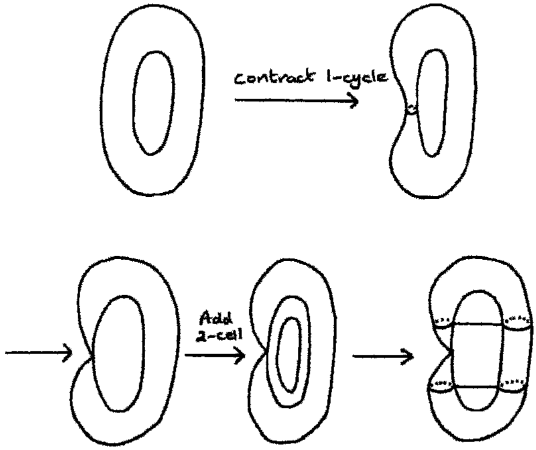


Figure 4: A stratified space from Goresky and MacPherson [17].

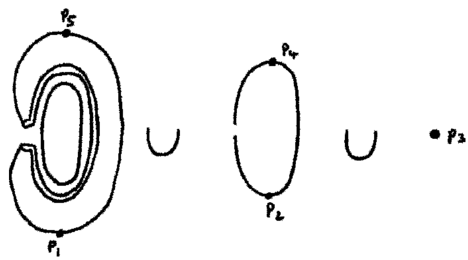


Figure 5: Stratification of torus with critical points.

tion behaves in a non-degenerate way with respect to how the manifolds meet each other. Then just as in Morse Theory we can describe how the topology of the singular space changes as we pass through critical points. The most developed method of stratification is Whitney stratification and, in fact, Goresky and Macpherson's original work was only for these stratifications but work of King, [35], and of Hamm, [23], allows us to use other stratifications, such as (c)-regular.

Let's see with an example how Stratified Morse Theory allows us to build up the topology of a space from simple building blocks. This example is taken from the introduction of [17]. The example used in illustrating Classical Morse theory is the torus with a height function. It is well known that part of the homology of the torus is generated by two circles. So to change the topology and make the space singular we can collapse one circle to a point and glue a 2-cell to the other as shown in Figure 4. The stratification is shown in Figure 5. The Morse function will be the height function h which gives the critical points $p_1, \dots, p_5$ labelled. Note that the stratum of dimension zero is needed otherwise the stratification would not satisfy the Whitney (b) condition.

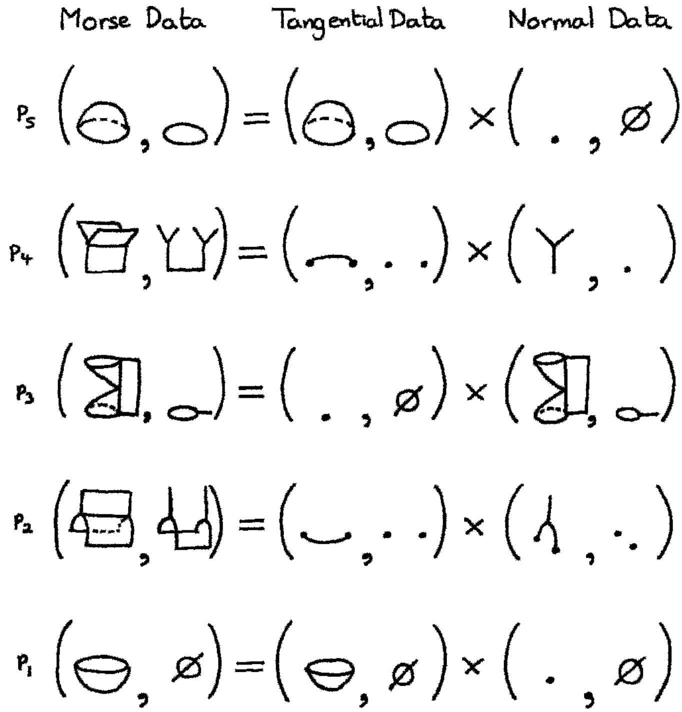


Figure 6: Product structure of the local Morse data.

As the height increases we can see that the topology of the space changes as we pass the critical values of h restricted to the various strata. Furthermore the change in topology depends only upon a small enough neighbourhood of the critical point. The space we attach to $X_{v-\varepsilon}$ to get $X_{v+\varepsilon}$ is called the *Morse data*.

The fact that is probably not obvious is that the Morse data is a product of two spaces. If we restrict our attention to the stratum containing the critical point then by classical Morse theory we add a cell to the manifold to get the new space. However, due to the Whitney conditions, along the stratum the neighbourhood is a product. Thus one can see that the Morse data should be a product of the classical Morse data and a space associated to a slice transverse to the stratum. Figure 6 shows the product structure of the Morse data for our example.

Morse functions

As in any Morse type theory we have to decide which functions are suitable and how common they are. Clearly the functions should satisfy the usual properties of Morse functions from classical theory: nondegenerate critical points with non-coincident critical values when restricted to the manifolds. The extra condition, (like most in stratification theory) relates to what is occurring in the normal direction to the manifold containing the critical point. First we need a definition:

Definition 6.2 Suppose X is a Whitney stratified space in the smooth manifold M and p is point in the stratum A of X. Then $Q \subseteq T_pM$ is a *generalized tangent space* at p to X if Q is the limit of tangent planes for a sequence of points converging to p. Hence, for example, T_pA is a generalized tangent space.

Definition 6.3 A *Morse function* f on the Whitney stratified set $X \subset M$ is a function that is the restriction of a smooth function F on M such that

(i) $f = F|X$ is proper;

(ii) for each stratum A of X the critical points of $f|A$ are nondegenerate and the critical values are distinct;

(iii) for every critical point $p \in A$ and for every generalized tangent plane Q at p, $dF(Q) \neq 0$ except if $Q = T_p(A)$.

Under mild restrictions there is a plentiful supply of Morse functions on a particular space.

Theorem 6.4 *Suppose X is a Whitney stratified subanalytic subset of the analytic manifold M. Then the functions $F : M \to \mathbb{R}$ that restrict to stratified Morse functions on X form an open and dense subset of the space of smooth functions. If $M = \mathbb{R}^n$, then the function given by restriction to X of the distance from a generic point in M is an example of a stratified Morse function.*

Morse data

Suppose X is a Whitney stratified space in the analytic manifold M and $f : X \to \mathbb{R}$ is a stratified Morse function with a critical $p \in A$, with critical value v, where A is a stratum of dimension n. Let N be a submanifold of M transverse to A with $N \cap A = \{p\}$ and let B be a small enough ball in M centred at p.

Define *Tangential Morse Data*, TMD, to be the space

$$\mathrm{TMD} = (D^\lambda \times D^{n-\lambda}, S^{\lambda-1} \times D^{n-\lambda})$$

where λ is the classical Morse index of $f|A$ at p, D^λ is a disc of dimension λ and $S^{\lambda-1}$ its boundary.

Define *Normal Morse Data*, NMD, to be the space

$$\mathrm{NMD} = (X \cap B \cap N) \cap (f^{-1}([v-\varepsilon, v+\varepsilon]), f^{-1}(v-\varepsilon)).$$

The space $X \cap B \cap N \cap f^{-1}(v-\varepsilon)$ is called *the halflink of f at p* and is denoted l^-.

Define the *Morse data* to be the space $\mathrm{MD} = \mathrm{TMD} \times \mathrm{NMD}$ where

$$(A, B) \times (C, D) = (A \times D, A \times D \cup B \times C).$$

Recall that $X_a = \{x \in X \,:\, x \in f^{-1}((-\infty, a]).$

Theorem 6.5 (Fundamental Theorem of Stratified Morse Theory) *With the above conditions*

$$X_{v+\varepsilon} \simeq X_{v-\varepsilon} \cup \mathrm{MD}$$

where the union of a space with a pair is taken to mean the attachment of the first space via the second.

Thus the Morse data for f is a product of the Tangential and Normal Morse Data.

The proof of the main theorem of Stratified Morse Theory, (the Morse Data of a function is a product of the Morse data for the stratum and the Morse data for the normal slice), occupies a large part of Goresky and MacPherson's book [17]. In effect they used repeated application of the First Thom-Mather Isotopy Lemma which has been disguised in a more useful (for their purposes) technique called Moving the Wall.

In subsequent years the proof has been simplified and generalized to stratified spaces other than Whitney, first by King in [35] and then by Hamm in [23] (which dealt with some of the weaknesses in [35]).

The method employed arises from the 'direct sum' idea in creating new singularities (see *New Milnor fibres from old* in Section 3). One can define the Morse data of an arbitrary continuous function as the pair $X \cap B \cap f^{-1}([v-\varepsilon, v+\varepsilon], v-\varepsilon)$ where B is a small ball around a critical point, v is the critical value and ε is a small number.

King's idea is that if we have a two functions $f_i : X_i \to \mathbb{R}$, $i = 1, 2$, where the functions and spaces are restricted to 'good' categories (Whitney stratified spaces and stratified Morse functions form a good category), then the Morse data for the function $f_1 \oplus f_2 : X_1 \times X_2 \to \mathbb{R}$ given by $(f_1 \oplus f_2)(x_1, x_2) = f_1(x_1) + f_2(x_2)$ is homeomorphic to the product of the Morse data for f_1 and f_2. The idea then is to show that for a stratified Morse function on the Whitney stratified space X we can 'split the function over' the product of the stratum and a normal slice.

This process can be exemplified using classical Morse theory. The main theorem arises from the Morse lemma: at a critical point of a Morse function f there is a choice of local coordinates such that

$$f(x_1, \dots, x_n) = \sum_i^k x_i^2 - \sum_{k+1}^n x_i^2.$$

The Morse data for $g(x_1, \dots, x_k) = \sum_i^k x_i^2$ and $h(x_{k+1}, \dots, x_n) = -\sum_{k+1}^n x_i^2$ are $(D^k, \emptyset)$ and (D^{n-k}, S^{n-k-1}) respectively. Since $f = g + h$ the Morse data for f is homeomorphic to

$$(D^k \times D^{n-k}, D^{n-k} \times \emptyset \cup D^k \times S^{n-k-1}).$$

This is equal to

$$(D^k \times D^{n-k}, D^k \times S^{n-k-1}).$$

Thus we recover the classical Morse Theory result.

The complex analytic setting

In obtaining the Morse data in the stratified case there are two objects to describe: the tangential data and the normal data. Finding the tangential data is merely the determination of a number: the number of negative eigenvalues of the Hessian. The normal data is more difficult to work out; it depends on the singularities of X and the function f.

Even if X is a (stratified) manifold this can prove problematic. However, for complex manifolds the Morse indices of the standard distance function are bounded above by the complex dimension rather than just the real dimension. In a similar way the normal data becomes simpler for complex analytic spaces. In fact, homotopically speaking, the normal Morse data does not depend on the function and hence to determine it we can investigate the normal data for a linear function on X. Better than that, the data is homeomorphic to a product of an interval and the intersection of X with a generic linear complex form. In other words, the complex link of the stratum! This allows us to prove theorems using induction as the complex link is a complex space of dimension one lower than X.

The main theorem for Stratified Morse Theory on a complex analytic space is the following.

Theorem 6.6 *Suppose that f is* any *stratified Morse function on X. Then the normal Morse data does not depend on the function. Furthermore,* $\mathrm{NMD} \simeq (\mathrm{Cone}(\mathcal{L}), \mathcal{L})$.

It should be noted that up to homotopy the above spaces do not depend on any choices involved, eg the choice of ε, metric on M, normal slice, etc.

7 Rectified homotopical depth

We now give a simple introduction to the notion of rectified homotopical depth (abbr. rhd). Rectified homotopical depth was introduced by Grothendieck in [20] to measure the failure of the Lefschetz hyperplane section theorem for singular spaces. The original theorem is the following.

Theorem 7.1 (Lefschetz Hyperplane theorem) *Suppose that $X \subseteq \mathbb{CP}^m$ is a non-singular projective algebraic variety and H a hyperplane. Then $\pi_i(X, X \cap H) = 0$ for $i < \dim(X)$.*

Lefschetz said this 'planted the harpoon of algebraic topology into the body of the whale of algebraic geometry', see [40] page 13.

Our interest in rhd arises from the fact that measuring the rhd for a complex analytic space tells us something about its Normal Morse data, (which as we have said, is independent of the Morse function). We shall see that if we replace the manifold X by a singular space, then we can replace $\dim(X)$ in the theorem with $\mathrm{rhd}(X)$.

Let X be a complex analytic space with stratification $\mathcal{S}$. For any stratum A in $\mathcal{S}$ let L_A denote the real link of A and let $\mathcal{L}_A$ denote the complex link. In our set up rhd keeps track of the vanishing of homotopy groups for these spaces.

We actually define the *rectified homotopical depth* of X, denoted $\mathrm{rhd}(X)$ using the following proposition.

Proposition 7.2 *The following are equivalent:*

(i) $\mathrm{rhd}(X) \geq n$,

(ii) $\pi_i(\mathrm{Cone}(L_A), L_A) = 0$ *for* $i < n - \dim_{\mathbb{C}} A$ *for all strata* $A \in \mathcal{S}$,

(iii) $\pi_i(\mathrm{Cone}(\mathcal{L}_A), \mathcal{L}_A) = 0$ *for* $i < n - \dim_{\mathbb{C}} A$ *for all strata* $A \in S$.

Thus the rhd of X is the largest number for which the second two statements hold. It is independent of the stratification chosen for the space, see [24].

The number $\mathrm{rhd}(X)$ is bounded above by the dimension of X because both types of link of the largest stratum are equal to the empty set and $\mathrm{Cone}(\emptyset)$ is defined to be a point. Thus $\pi_0(\mathrm{Cone}(\emptyset), \emptyset) = \pi_0(\text{point}, \emptyset) \neq 0$. In particular, if X is non-singular, then $\mathrm{rhd}(X) = \dim_{\mathbb{C}}(X)$.

Similarly we can define *rectified homological depth*, $\mathrm{rHd}(X;\mathbb{Z})$ by replacing the relative homotopy groups in the definition by relative homology groups. It is also possible to make a definition of rectified homological depth, $\mathrm{rHd}(X;G)$, for any coefficient group G.

We have the following lemma:

Lemma 7.3 *For a field F,*

$$\mathrm{rhd}(X) \leq \mathrm{rHd}(X;\mathbb{Z}) \leq \mathrm{rHd}(X;F) \leq \dim_{\mathbb{C}} X.$$

The first relation is proved by using Hurewicz's theorem. The second by the universal coefficient theorem. As above the last relation is simple.

Theorem 7.4 *Let $f : X \to \mathbb{R}$ be a stratified Morse function with a critical point at p in the stratum A with critical value v. Let $X_b = f^{-1}([-\infty, b])$. Let λ be the Morse index at p and ε a sufficiently small number, then*

(i) $\pi_i(X_{v+\varepsilon}, X_{v-\varepsilon}) = 0$ *for* $i < (\lambda - \dim_{\mathbb{C}} A) + \mathrm{rhd}(X)$,

(ii) $H_i(X_{v+\varepsilon}, X_{v-\varepsilon}) = 0$ *for* $i < (\lambda - \dim_{\mathbb{C}} A) + \mathrm{rHd}(X)$.

The second is easier to prove. Using excision we get

$$H_i(X_{v+\varepsilon}, X_{v+\varepsilon}) \cong H_i(\mathrm{NMD}, \partial\,\mathrm{NMD}).$$

But $(\mathrm{NMD}, \partial\,\mathrm{NMD})$ is equal to the product $(D^\lambda, \partial D^\lambda) \times (\mathrm{Cone}(\mathcal{L}_A), \mathcal{L}_A)$. The product theorem for homology gives the vanishing of homology that is required.

For (i) more advanced and less well known techniques need to be used but the spirit is the same. One can construct a proof using [17] II.4.

An important theorem, which is essentially a local Lefschetz type theorem, was proved by Hamm and Lê. See [24] 3.2.1.

Theorem 7.5 *Suppose that X is a complex analytic space and Y is a subspace defined set theoretically by no more than r equations, then*

$$\mathrm{rhd}(Y) \geq \mathrm{rhd}(X) - r.$$

The proof actually follows from reasoning similar to the proof of Theorem 5.8.

As a corollary of this we are able to find the rhd of a very large class of complex analytic spaces.

Corollary 7.6 *Suppose that X is a local complete intersection. Then* $\mathrm{rhd}(X) = \dim_{\mathbb{C}} X$.

Grothendieck's original intention for rectified homotopical depth was that is was analogous to the notion of depth from commutative algebra. For example, one can see that for regular rings and complete intersection rings that depth equals the dimension of the ring. For rectified homotopical depth we have that manifolds (equivalent to regular rings in the analogy) and local complete intersections have $\mathrm{rhd}(X) = \dim X$.

Rings with maximal depth are called Cohen–Macaulay and have many interesting properties. Similarly, spaces with maximal rectified homotopical depth has good properties - since, almost by definition, their complex links are wedges of spheres in middle dimension.

An interesting unexplored topic is the analogy with Gorenstein rings. To some extent this was tackled in Goresky and MacPherson's original work on Poincaré duality for singular spaces but it would be good to pursue the analogy further from the perspective of rectified homotopical depth.

Example 7.7 An example of a space that does not have $\mathrm{rhd}(X) = \dim_{\mathbb{C}} X$: Let $X = \mathbb{C}^2 \cup \mathbb{C}^2$ be two copies of $\mathbb{C}^2$ in $\mathbb{C}^4$ that intersect transversally at the origin. Stratify X by taking the origin as one stratum and the complement of the origin in X as the other. It is obvious that the complex link of the origin, $\mathcal{L}$, is a disjoint union of two discs. Thus $\pi_0(\mathrm{Cone}(\mathcal{L}), \mathcal{L}) = 0$ but $\pi_1(\mathrm{Cone}(\mathcal{L}), \mathcal{L}) \neq 0$. Therefore $\mathrm{rhd}(X) = 1 < \dim_{\mathbb{C}} X = 2$.

Another theorem of interest concerns the notion of perverse sheaves, see [4] for the definition. Let $\mathbb{C}^\bullet_X[\dim_{\mathbb{C}} X]$ denote the sheaf complex that has the constant sheaf of complex numbers in the $\dim_{\mathbb{C}} X$ position and zero in other degrees.

Theorem 7.8 ([24]) *We have:* $\mathrm{rHd}(X;\mathbb{Q}) = \dim_{\mathbb{C}} X$ *if and only if* $\mathbb{C}^\bullet_X[\dim_{\mathbb{C}} X]$ *is perverse in the sense of Bernstein-Beilensen-Deligne.*

The final theorem of this section shows how rhd can be used to give a Lefschetz theorem. The proof is a simple example of the type of technique used in proving Lefschetz type theorems.

Theorem 7.9 *Let X be a complex projective variety in $\mathbb{CP}^N$. Suppose H is a hyperplane in $\mathbb{CP}^N$ that does not contain all of X, then*

$$\pi_i(X, X \cap H) = 0 \text{ for } i < \mathrm{rhd}(X - H).$$

We give a proof here since almost all Lefschetz type theorems can be proved using the outline of this proof. **Proof** Identify $\mathbb{CP}^N - H$ with $\mathbb{C}^N$. For a generic point p in $\mathbb{C}^N$ the function $\phi(z) = ||z - p||^2$ is a stratified Morse function on $X \cap \mathbb{C}^N$. This function has a finite number of critical points and the Morse index for a critical point on the stratum A is not greater than $\dim_{\mathbb{C}} A$. For some large enough R all critical point of ϕ are contained in the set $\phi^{-1}([0, R])$. We consider the Morse function $-\phi$ on the set $X \cap \mathbb{C}^N$. The Morse indices are bounded below by $2\dim_{\mathbb{C}} A - \dim_{\mathbb{C}} A = \dim_{\mathbb{C}} A$. Let $X_a = (-\phi)^{-1}(-\infty, a)$. By passing through all critical points of $-\phi$ we build up X from X_{-R} using stratified Morse theory. The critical points of ϕ all lie outside $X \cap H$ and thus the connectivity of the complex link of the stratum depends only on $\mathrm{rhd}(X - H)$ and not $\mathrm{rhd}(X)$.

By Theorem 7.4 we get $\pi_i(X_{v+\varepsilon}, X_{v-\varepsilon}) = 0$ for $i < (\lambda - \dim_{\mathbb{C}} A + \mathrm{rhd}(X - H))$ for any critical value v, where λ is the Morse index of the critical point. As $\lambda \geq \dim_{\mathbb{C}} A$, the pair is therefore $(\mathrm{rhd}(X - H) - 1)$-connected. So (X, X_{-R}) is $(\mathrm{rhd}(X - H) - 1)$-connected.

X can be triangulated with $X \cap H$ being a subtriangulation. Thus there exists a neighbourhood U of $X \cap H$ that retracts onto $X \cap H$. For some $R' > R$ the space $X_{-R'}$ is a subset of U and since the interval $[-R', -R]$ contains no critical points for $-\phi$, by Thom's First Isotopy Lemma, X_{-R} retracts onto $X_{-R'}$. Thus in the chain of inclusions,

$$X \cap H \subseteq X_{-R'} \subseteq U \subseteq X_{-R}$$

the composition of any two inclusions induces an isomorphism on homotopy groups. This implies that the natural map $\pi_*(X \cap H) \to \pi_*(X_{-R})$ is an isomorphism. By examining the long exact sequence arising from the triple $(X, X_{-R}, X \cap H)$ we arrive at $\pi_i(X, X \cap H) = 0$ for $i < \mathrm{rhd}(X - H)$ as stated. □

This theorem also exemplifies the flavour of proofs using Stratified Morse Theory. Essentially the process involves taking a stratified Morse function and using it to build up one space from another. Something about the function allows us to bound the indices from above or below and the rhd hypothesis allows us to say something about the connectivity of the normal Morse data. Thus we get a theorem involving rhd which in effect tells us nothing practical until we 'attack' with an example where we have calculated the rhd. For example, we have the following.

Corollary 7.10 ([48] Theorem 7.4.) *Suppose that $X \cap H$ contains the singular locus of X. Then*

$$\pi_i(X, X \cap H) = 0 \textit{ for } i < \dim_{\mathbb{C}} X.$$

Since $X - H$ is nonsingular we have $\mathrm{rhd}(X - H) = \dim_{\mathbb{C}}(X - H) = \dim_{\mathbb{C}} X$ and we can apply the theorem. We also have the following.

Corollary 7.11 *Suppose that X is a local complete intersection. Then*

$$\pi_i(X, X \cap H) = 0 \textit{ for } i < \dim_{\mathbb{C}} X.$$

Since $X - H$ is a local complete intersection by Theorem 7.5 we have $\mathrm{rhd}(X - H) = \dim_{\mathbb{C}}(X - H) = \dim_{\mathbb{C}} X$. This of course includes the case that X is non-singular.

Once the overall method of proof has been grasped it is possible to generalize to other cases where the space is not in $\mathbb{CP}^n$. Here one also has to estimate the Morse index and this is usually done by using the concept of q-convexity. See [17], [22] and [57].

8 Relative stratified Morse theory

Suppose $X \subseteq M'$ and $Z \subseteq M$ are closed Whitney stratified subsets of smooth manifolds and that $\pi : X \to Z$ is a proper surjective stratified map, i.e., π is the restriction of $\pi' : M' \to M$, a smooth map such that π maps strata of X submersively to strata of Z. This is not an unnatural set up for if X and Z are complex analytic spaces and π is a proper complex analytic map, then there exist stratifications of X and Z such that π becomes a stratified map. (See [17] page 43.) Let $f : Z \to \mathbb{R}$ be a stratified Morse function with a critical point at p.

The idea of relative stratified Morse theory is to build up the space X using the maps f and π. The composition $f \circ \pi$ is not a Morse function and approximating $f \circ \pi$ by a Morse function may not viable since we could lose estimates on the Morse indices.

Nevertheless, this set up is amenable to study. For if the interval $[a, b]$ contains no critical value of f, then X_a is homeomorphic in a stratum preserving way to X_b. This is not too difficult to prove as it relies upon stratification techniques and the First Thom-Mather Isotopy Lemma. The important result is the following:

Theorem 8.1 *Let $l^- = \pi^{-1}(Z \cap N \cap B_\epsilon(p)) \cap f^{-1}(-\varepsilon)$. There exists a map $\phi : l^- \to \pi^{-1}(p)$, see [17] page 117. If λ is the Morse index of f at the critical point p, then X_b has the homotopy type of X_a with the attachment of the pair*

$$(D^\lambda, \partial D^\lambda) \times (\mathrm{cyl}(l^- \to \pi^{-1}(p)), \pi^{-1}(l^-))$$

where cyl *denotes the mapping cylinder of ϕ and D^λ is the standard disc of dimension λ.*

The proof is given in [17].

Thus we see that the tangential Morse data does not depend on π and our main difficulty with the set up is to calculate the 'relative' Morse data.

If Z is complex analytic, then l^- is less complicated because it is homotopically equivalent to the preimage of the complex link of the stratum containing p.

If X is also complex analytic and π is a finite complex analytic map then the structure of the mapping cylinder becomes simpler. It is the disjoint union of cones over the spaces involved.

There are many as yet unexplored situations where relative Stratified Morse Theory can be applied. In [17] various relative versions of the Lefschetz Hyperplane are proved. Another application is given in [26]. The set up there is the following for the case $G = S_k$, the group of permutations on k objects. Suppose G is a finite group acting upon the complex analytic space X. The map $\pi : X \to X/G$ is a finite surjective complex analytic map which we can stratify. This map is then a Thom A_π map. The relative Morse data will inherit a G action since we can lift a vector field giving a homeomorphism on X/G to a G-equivariant vector field on X. (We can lift the vector fields as π is a Thom A_π map and these fields are G-invariant). Thus X_b is G-homotopically equivalent to the union of X_a and the relative Morse data.

Furthermore, if M is a real manifold intersecting the strata of X/G transversally, then we can lift vector fields on $(X/G) \cap M$ to G-equivariant vector fields on $X \cap \pi^{-1}(M)$.

9 Topology of images and multiple point spaces

Through most of this article our concern has been with the topology of fibres. The Implicit Function Theorem, Sard's Lemma and the Morse Lemma effectively tell us that the topology of fibres is, in some sense, easy to deal with: Most fibres are non-singular and most singular functions have isolated quadratic singularities.

But what about the topology of images? This is considerably harder to deal with. Let us first consider one of the main difficulties. This is the realisation that singularities are common for images. Consider the Whitney cross-cap of Example 2.9(ii). This is a stable map - a small perturbation will produce an equivalent singularity under local diffeomorphisms of the source and target - and so this is a perfectly natural object. However, the image has non-isolated singularities. In the case of critical points of functions or maps we had the possibility of taking a nearby fibre which was non-singular. We do not have this possibility for the Whitney umbrella. There is no natural nearby non-singular image. Furthermore, to make matters worse, the umbrella is the image of an algebraic map, but the image itself is not algebraic in the real case - it is semi-algebraic.

To show that progress can be made in describing the topology of the image of a map consider the following. Suppose that we consider a smooth map $f : N \to P$, from a closed surface N, (i.e., non-singular, compact and without boundary), to a 3-manifold P, such that the only singularities of f are the stable. The singularities that occur are the Whitney cross cap or the transverse crossing of 2 or 3 sheets. In the latter case, these are called triple points and will be isolated just as the cross caps are isolated.

Theorem 9.1 (Izumiya-Marar [32]) *Suppose that $f : N \to P$ is a stable mapping from a*

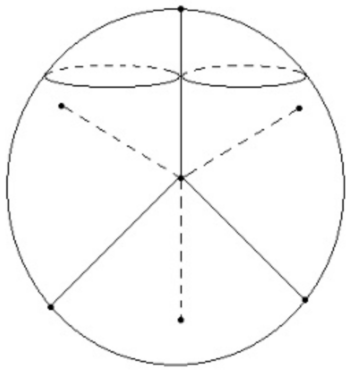


Figure 7: Steiner's Roman Surface.

closed surface to a 3-manifold. Then

$$\chi(f(N)) = \chi(N) + \frac{C(f)}{2} + T(f),$$

where $\chi(X)$ denotes the Euler characteristic of the space X, $C(f)$ is the number of cross caps and $T(f)$ is the number of triple points.

This has been generalized in a number of directions, see [28]. The point is that we look at significant features of the map.

Example 9.2 Steiner's Roman surface: Suppose that f is the map $f : \mathbb{RP}^2 \to \mathbb{R}^3$ given by $f([x:y:z]) = [xy : xz : yz : x^2+y^2+z^2]$, (since $x=y=z=0$ is impossible we have $[a:b:c:0] \notin f(\mathbb{RP}^2)$ so the target of the map is indeed $\mathbb{R}^3$). See [1] p40 for more information.

The map f is smooth and has stable singularities, with 6 cross caps and 1 triple point. See Figure 7. Note that three of the cross caps are hidden from view.

Here

$$\chi(R) = \chi(\mathbb{RP}^2) + C(f)/2 + T(f) = 1 + (6/2) + 1 = 5.$$

From the figure it is possible to see that the space is homotopically equivalent to a wedge of four 2-spheres.

This examples shows how we have to be careful. We have to look at the singularities of the *maps* rather than the singularities of the *image*. One can have a non-stable map such that the singularities of the image are locally homeomorphic to the singularities of the image of a stable map. For example, Steiner's Roman surface can also be given as the image of a map from S^2 to $\mathbb{R}^3$. The triple point is formed by four corners of a cube (see Example 2.10) coming together. In this case

$$\chi(S^2) + C(f)/2 + T(f) = 2 + (6/2) + 1 = 6$$

if we count the C and T by what the singularities of the image look like.

Multiple point spaces

We shall assume now that our maps are finite and proper, i.e., each point in the target has a finite number of preimages and the preimage of a compact set is compact); for the moment we shall not assume smoothness of the map, and hence will have a continuous map $f : X \to Y$.

There are many ways of defining multiple point spaces for a finite and proper map. For example, one can define the double point set as the set of points in X where f is not injective. That is, the closure of the set $x \in X$ such that there exists $y \neq x$ such that $f(x) = f(y)$. Alternatively, some authors define the double point set as the *image* of this set.

We take a third alternative which has a number of advantages. The double point space of a map f is the closure in $X^2 (= X \times X)$ of the set of pairs (x, y), with $x \neq y$, such that $f(x) = f(y)$. This first advantage of this is that often this space is, in some vague sense, less singular than that which the other definitions give. The second advantage, though this may not appear so useful at first sight, is that this space has more symmetry – the group of permutation on 2 objects acts on X^2 by permutation of copies.

We can generalize this so that the k^{th} *multiple point space of a map* is the closure of the set of k-tuples of pairwise distinct points having the same image:

Definition 9.3 Let $f : X \to Y$ be a finite map of topological spaces. Then, the k^{th} *multiple point space of* f, denoted $D^k(f)$, is defined to be

$$D^k(f) := \text{closure}\{(x_1, \dots, x_k) \in X^k \mid f(x_1) = \cdots = f(x_k) \,\text{for}\, x_i \neq x_j, i \neq j\}.$$

Just as in the case of the double point set these sets are considerably simpler than the sets in the target formed by counting the number of preimages, the former may be non-singular in contrast to the highly singular latter. In effect, the multiple point spaces act as a resolution of the image.

There exist maps $\varepsilon_{i,k}; D^k(f) \to D^{k-1}(f)$ induced from the natural maps $\tilde{\varepsilon}_{i,k} : X^k \to X^{k-1}$ given by dropping the i^{th} coordinate from X^k. There also exists maps $\varepsilon_k : D^k(f) \to Y$ given by $\varepsilon_k(x_1, \dots, x_k) = f(x_1)$.

We will now officially define the multiple point spaces in the image that we have mentioned above.

Definition 9.4 The k^{th} *image multiple point space*, denoted $M_k(f)$, is the space $\varepsilon_k(D^k(f))$.

As stated earlier, the spaces $M_k(f)$ can be highly singular compared to $D^k(f)$, in the sense that $D^k(f)$ could be non-singular but $M_k(f)$ could have non-isolated singularities.

Example 9.5 Let $f : \mathbb{R}^2 \to \mathbb{R}^3$ be the Whitney umbrella $f(x, y) = (x, xy, y^2)$. Then,

$$\begin{aligned} D^2(f) &= \text{closure}\{(x_1, y_1, x_2, y_2) \in \mathbb{R}^2 \times \mathbb{R}^2 \mid (x_1, x_1y_1, y_1^2) = (x_2, x_2y_2, y_2^2); \\ &\qquad\qquad (x_1, y_1) \neq (x_2, y_2)\} \\ &= \{(0, y_1, 0, -y_1) \in \mathbb{R}^4\}. \end{aligned}$$

From this we can see that $D^2(f)$ is a manifold – in fact, a line – and that $M_2(f)$ is not a manifold (although one could count it as a manifold with boundary).

Also, we can see clearly that S_2 acts on $D^2(f)$.

Morin described stable map germs $f : (\mathbb{R}^n, 0) \to (\mathbb{R}^p, 0)$ where $n < p$ and f has corank 1. He stated that f is equivalent to a map of the form,

$$(x_1, \ldots, x_s, u_1, \ldots, u_{l-2}, w_{1,1}, w_{1,2}, \ldots, w_{p-n+2,l}, y)$$
$$\mapsto (\underline{x}, \underline{u}, \underline{w}, y^l + \sum_{i=1}^{l-2} u_i y^i, \sum_{i=1}^{l-1} w_{1,i} y^i, \ldots, \sum_{i=1}^{l-1} w_{p-n+2,i} y^i),$$

where l is the multiplicity of the germ.

Mather proved in [47] that stable multi-germ maps are constructed from stable mono-germs with these mono-germs meeting transversally. Thus we have an explicit description of a map $f : N \to P$ if it is such that $\dim N < \dim P$ and is stable and corank 1. (Recall that a map is corank 1 if at each point the differential has corank at most 1. That is, the map could in theory, in this case, be an immersion at a point.)

From this explicit description we can produce a large number of examples via the following theorem.

Theorem 9.6 ([43]) *Suppose that $f : N \to P$ is such that $\dim N < \dim P$ and that the singularities of f are stable and corank 1. Then, $D^k(f)$ is non-singular.*

The proof of this is given in [43]. Furthermore, they give a method for calculating local defining equations for $D^k(f)$ using determinants. A simpler proof of the theorem (and one that holds in the smooth case as well) is to reduce to the problem of normal forms for the singularities considered. This is possible since, for a multi-germ f, if there is a local change of coordinates in source and target to produce f', then $D^k(f)$ and $D^k(f')$ are locally diffeomorphic.

Now, as discussed, corank 1 multi-germs have been classified by Morin, [50], for mono-germs, and Mather, [47], for multi-germs. Using Morin's description for mono-germs and the Marar-Mond description for defining equations it is straightforward to calculate that $D^k(f)$ is non-singular. In the case of multi-germs, we observe from Mather's classification that the multi-germ occurs as the trivial unfolding of a some stable mono-germ. Thus since we know the multiple point space is non-singular at the mono-germ it must be in some neighbourhood.

In fact, it is possible to prove a converse: If f is corank 1 and the kth multiple point spaces are non-singular of dimension $k \dim N - (k-1) \dim P - 1$ for all k, then f is stable.

For corank 2 stable maps, unfortunately, $D^2(f)$ can be singular (and probably is so in general).

Marar and Mond have a stronger result in [43] involving the restriction of $\tilde{D}^k(f)$ to fixed point sets of the action of S_k on $\mathbb{C}^{n-1} \times \mathbb{C}^k$. Also, they deal with the case of isolated instabilities. Here they show that the $D^k(f)$ can have isolated singularities, which furthermore are complete intersections. It is this which is key to later results: For isolated instabilities the multiple point spaces are isolated complete intersection singularities and hence we if we perturb the instability to produce a stable map, then the ICIS are perturbed to their Milnor fibres (as the multiple point spaces of a stable map are non-singular by the above).

It is easy to see that given a map we can associate lots of invariants to it by taking invariants of the multiple point spaces.

10 The image computing spectral sequence

Alternating homology of a complex

As stated earlier the multiple point spaces act as resolution of the image. However, since the triple point space is the double point space of the natural map $D^2(f)$ to X given by projection, it is obvious that if we use the homology of the multiple point spaces to give us the homology of the image that we get too much information. It turns out that we need to look at the alternating homology of the multiple point spaces. This is where we exploit the symmetry of the multiple point spaces. The group, S_k, of permutations on k objects acts naturally on $D^k(f)$ by permutation of copies of X^k. The alternating homology is the homology of the subcomplex of chains that alternate, i.e., are anti-symmetric with respect to S_k.

Let us put the details to this. Denote by sign the natural sign representation for S_k. The space $Z \subset X^k$ is called *S_k-cellular* if it is S_k-homotopy equivalent to a cellular complex. That is, there is a homotopy equivalence (respecting the action) to a complex of cells upon which S_k acts cellularly. (This latter means that cells go to cells and if a point of a cell is fixed by an element of S_k, then the whole cell is fixed by the element.) Whitney stratified spaces for which the strata are S_k-invariant can be triangulated to respect the action and hence have a cellular action.

Definition 10.1 Let

$$\mathrm{Alt}_{\mathbb{Z}} = \sum_{\sigma \in S_k} \mathrm{sign}(\sigma)\sigma.$$

We define alternating homology by applying this operator.

Definition 10.2 The *alternating chain complex of* Z, $C_*^{alt}(Z;\mathbb{Z})$ is defined to be the following subcomplex of the cellular chain complex, $C_n(Z;\mathbb{Z})$, of Z,

$$C_n^{alt}(Z;\mathbb{Z}) := \mathrm{Alt}_{\mathbb{Z}}\, C_n(Z;\mathbb{Z}).$$

The elements of $C_n^{alt}(Z;\mathbb{Z})$ are called *alternating* or *alternated chains*.
There is an alternative way to define or a useful way to calculate $C_n^{alt}(Z;\mathbb{Z})$:

$$C_n^{alt}(Z;\mathbb{Z}) \cong \{c \in C_n(Z;\mathbb{Z}) \mid \sigma c = \mathrm{sign}(\sigma)c \text{ for all } \sigma \in S_k\}.$$

Now we just need to apply the homology functor to this subcomplex to get alternating homology.

Definition 10.3 The *alternating homology of* Z, denoted $H_*^{alt}(Z;\mathbb{Z})$, is defined to be the homology of $C_*^{alt}(Z;\mathbb{Z})$.

Note that in [18] $H_i^{alt}(Z;\mathbb{Z})$ denotes the alternating part of integral homology. However, our notation is more in keeping with traditional notation in homology.

If we wish to define alternating homology over general coefficients then we may do so in the usual way by tensoring $C_*^{alt}(Z;\mathbb{Z})$ by the coefficient group.

Example 10.4 Suppose $T = S^1 \times S^1$ denotes the standard torus. Then S_2 acts on T by permutation of the copies of S^1. Let Z be the points $(z, z+\pi) \in T$, then Z is just a circle with antipodal action. We can give Z a cellular structure by choosing two antipodal points p_1 and p_2 as 0-cells and then the complement of these points will form two 1-cells,

e_1 and e_2, upon which S_2 acts by permutation and whose orientation we induce from an orientation of the circle. Then $\sigma(e_1) = e_2$ and $\sigma(e_2) = e_1$, where σ is the non-trivial element of S_2.

The group $C_0^{alt}(Z;\mathbb{Z})$ is generated by $c_0 = p_1 - p_2$ and $C_0^{alt}(Z;\mathbb{Z})$ is generated by $c_1 = e_1 - e_2$. The boundary of c_1 is $-p_1 + p_2 - p_1 + p_2 = -2(p_1 - p_2)$. Therefore c_1 is not a cycle and $2(p_1 - p_2)$ is a boundary, hence

$$\begin{aligned} H_0^{alt}(Z;\mathbb{Z}) &= \mathbb{Z}_2, \\ H_1^{alt}(Z;\mathbb{Z}) &= 0. \end{aligned}$$

An alternating homology group is not a subgroup of ordinary homology as the example shows: $H_0^{alt}(D^2(f);\mathbb{Z}) = \mathbb{Z}_2$ is not a subgroup of $H_0(D^2(f);\mathbb{Z}) = \mathbb{Z}$.

Our fundamental example for alternating homology is $D^k(f) \subset X^k$. Suppose that, for $k > 1$, the S_k-action on $D^k(f)$ is cellular. Then we can define alternating homology for $D^k(f)$.

Example 10.5 Let $f : B^2 \to \mathbb{RP}^2$ be the quotient map that maps the unit disc B^2 to real projective space by antipodally identifying points on the boundary of the disc. Then $D^2(f) \subset B^2 \times B^2$ is just the circle in Example 10.4 and so $H_0^{alt}(D^2(f);\mathbb{Z})$ has the alternating homology of that example. The set $D^3(f)$ is empty.

Let $D^k(f)^g$ denotes the fixed point set in $D^k(f)$ of the element $g \in S_k$, and let $\chi^{alt}(D^k(f)) = \sum_i (-1)^i \dim_{\mathbb{Q}} \mathrm{Alt}\, H_i(X;\mathbb{Q})$ denote the alternating Euler characteristic.

Lemma 10.6 ([31]) *Assume that S_k induces a cellular action on $D^k(f)$. Then,*

$$\chi^{alt}(D^k(f)) = \frac{1}{k!} \sum_{g \in S_k} \mathrm{sign}(g) \chi(D^k(f)^g),$$

provided that each $\chi(D^k(f)^g)$ is defined.

A similar theorem is stated in section 3 of [19] in the case that each $D^k(f)^g$ is the Milnor fibre of an isolated complete intersection singularity, however, their result as stated is false.

This result is very useful for low dimensional cases for calculating the homology of an image since, as we shall see in a moment, that the Euler characteristic of the image is the alternating sum of the alternating Euler characteristics of the multiple point spaces.

Example 10.7 Let us see in a simple example how the alternating homology of multiple point spaces arises in the computation of the homology of an image.

Consider a map $f : M \to P$ such that $D^3(f) = \emptyset$. Then, there is a homeomorphism of $D^2(f)$ onto its image under the natural projection map $D^2(f)$ to M. We can see that there is a surjective map $C_*(M)$ to $C_*(f(M))$. The kernel of this is the complex of alternating chains.

Figure 8 shows an example where two hemispheres are glued along their edges to produce a sphere. One can see that in this case there is an inclusion of the double point set into M. Here the double point set is two circles and these can given an orientation arising from the equator in the circle. At the level of chains we can see that the symmetric chain maps to the chain given by the equator. Hence the alternating chains must be in the kernel of the map $C_*(M)$ to $C_*(f(M))$.

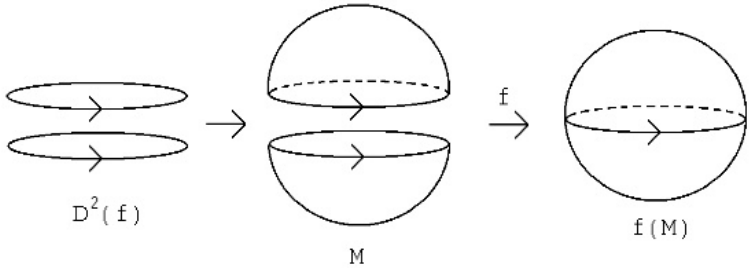


Figure 8: Example of spaces used in the short exact sequence

Thus, in general, we get a short exact sequence of complexes:

$$0 \to C_*^{alt}(D^2(f)) \to C_*(M) \to C_*(f(M)) \to 0,$$

and this naturally leads to a long exact sequence:

$$\cdots \to H_i^{alt}(D^2(f)) \to H_i(M) \to H_i(f(M)) \to H_{i-1}^{alt}(D^2(f)) \to \dots$$

So we have a long exact sequence that relates the homology of the image to the source and its double point space.

The spectral sequence

We saw that for maps with (at worst) double points we had a long exact sequence relating the homology of the image to the homology of the domain and the alternating homology of the double point space. It should therefore come as no surprise that to generalize the relation arising from this long exact sequence we need a spectral sequence.

In this section we describe such a spectral sequence that relates the homology of the image to the alternating homology of the multiple point spaces.

Theorem 10.8 *Suppose $f : X \to Y$ is a finite and proper continuous map, such that $D^k(f)$ has the S_k-homotopy type of an S_k-cellular complex for all $k > 1$ and $M_k(f)$ has the homotopy type of a cellular complex for all $k > 1$. Then there exists a spectral sequence*

$$E^1_{p,q} = H_q^{alt}(D^{p+1}(f); \mathbb{Z}) \implies H_{p+q+1}(f(X); \mathbb{Z}).$$

The differential is the naturally induced map

$$\varepsilon_{1,k_*} : H_i^{alt}(D^k(f); \mathbb{Z}) \to H_i^{alt}(D^{k-1}(f); \mathbb{Z}).$$

The proof of this is given in [30].

Example 10.9 We have already calculated the alternating homology of the multiple point spaces for the map $f : B^2 \to \mathbb{RP}^2$ of Example 10.5. The image computing spectral sequence for f is given below.

H_2^{alt}	0	0	0
H_1^{alt}	0	0	0
H_0^{alt}	$\mathbb{Z}$	$\mathbb{Z}_2$	0
	B^2	$D^2(f)$	$D^3(f)$

All the differentials of this sequence must be trivial and so the sequence collapses at E^1 and since there are no extension difficulties we can read off the homology of the image of f, i.e., the real projective plane.

Example 10.10 We can use the sequence to prove Theorem 9.1 of Izumiya and Marar

We can triangulate $f(N)$ with the cross caps and triples among the vertices and so that the image of D^2 is a subcomplex. Since f is proper and finite we can pull back the triangulation of $f(N)$ to give one for N. Thus we have a finite, proper, surjective map of CW–complexes and so we can apply the image computing spectral sequence.

The alternating homology of D^1 is just the ordinary homology of N and as there are no quadruple points the alternating homology of D^k is trivial for $k > 3$.

The triple point set D^3 is just six copies of the triple points of the source and we can see that these form alternating zero chains, one for each triple point. So $H_i^{alt}(D^3) = \mathbb{Z}^{T(f)}$ for $i = 0$ and zero otherwise.

Each normal crossing of two sheets has a cross cap at each end or is a circle and so we can pair the cross caps. The two sheeted normal crossings meet at triple points. This means that D^2 has $C(f)$ points in the diagonal from which come two 1-cells which permute under the action of S_2 on D^2. So any particular 1-cell will join a pair of cross caps and if any two 1-cells cross then they cross at a triple point. The only other parts of D^2 are μ pairs of circles.

This allows us to calculate the alternating homology of D^2. We have

$$H_i^{alt}(D^2) = \begin{cases} \mathbb{Z}^{C(f)/2} \oplus \mathbb{Z}^{\mu}, & \text{for } i = 1, \\ \mathbb{Z}^{\mu}, & \text{for } i = 0, \\ 0, & \text{otherwise.} \end{cases}$$

We can work out the Euler characteristic of the limit of a spectral sequence from the Euler characteristic of any level, (provided of course that the terms are finitely generated). So,

$$\begin{aligned} \chi(f(N)) &= \chi(E_\infty^{p,q}) \\ &= \chi(E_1^{p,q}) \\ &= \chi^{alt}(D^1) - \chi^{alt}(D^2) + \chi^{alt}(D^3) \\ &= \chi^{alt}(D^1) - [\mu - (\mu + C(f)/2)] + T(f) \\ &= \chi(N) + C(f)/2 + T(f), \end{aligned}$$

where χ^{alt} means the Euler characteristic of the alternating homology.

It should be remarked that if we can work out how the spectral sequence collapses and there are no extension problems then it may give more information than just the Euler characteristic.

Final remarks

The image computing spectral sequence has not been sufficiently applied and there are many unexplored areas in which it could be used. Consider the case of the quotient space given by the action of a finite group G on a set X. The quotient map $\pi : X \to X/G$ is finite and in many important case will be continuous. The sequence has not been investigated in this case, not even to see the relation with the classic theorems on quotients of spaces.

Bibliography

[1] F. Apéry: *Models of the real projective plane* (Braunschweig; Weisbaden: Vieweg, 1987)

[2] V. I. Arnol'd, S. M. Guseĭn-Zade, A. N. Varchenko: *Singularities of differentiable maps. Vol. II. Monodromy and asymptotics of integrals* (Monographs in Mathematics, 83. Birkhäuser Boston, Inc., Boston, MA, 1988)

[3] V. I. Arnol'd, V. S. Afrajmovich, Yu. S. Il'yashenko, L. P. Shil'nikov: *Bifurcation theory and catastrophe theory, Dynamical systems V* (Encyclopaedia Math. Sci., 5, Springer, Berlin, 1994)

[4] A. Beilinson, J. Bernstein and P. Deligne: Faisceaux Pervers *Astérisque* **100** (Soc. Math. de France), (1983) 172

[5] K. Bekka: C-regularité et trivialité topologique *In: Singularity theory and its applications, Part I* (Coventry, 1988/1989, Lecture Notes in Math., 1462, Springer, Berlin, 1991) 42–62

[6] E. Brieskorn: Beispiele zur Differentialtopologie von Singularitäten *Invent. Math.* **2** (1966) 1–14

[7] Th. Bröcker: *Differentiable germs and catastrophes* (London Mathematical Society Lecture Note Series, No. 17. Cambridge University Press, Cambridge-New York-Melbourne, 1975)

[8] J. W. Bruce, A. A. du Plessis, L. C. Wilson: Discriminants and liftable vector fields *J. Algebraic Geom.* **3(4)** (1994) 725–753

[9] P. Deligne: *Groupes de monodromie en géométrie algébrique II* (Séminaire de Géométrie Algébrique du Bois-Marie 1967–1969 (SGA 7 II), Lecture Notes in Mathematics, Vol. 340. Springer-Verlag, Berlin-New York, 1973)

[10] A. Dimca: *Sheaves in topology* (Universitext. Springer-Verlag, Berlin, 2004)

[11] A. du Plessis, T. Gaffney, L. C. Wilson: Map-germs determined by their discriminants *In: Stratifications, singularities and differential equations, I* (Marseille, 1990; Honolulu, HI, 1990, Travaux en Cours, 54, Hermann, Paris, 1997) 1–40

[12] A. du Plessis, C. T. C. Wall: *The geometry of topological stability* (London Mathematical Society Monographs. New Series, 9. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1995)

[13] W. Fulton: On the Topology of Algebraic Varieties *In: Proc. Symp. Pure Math.* (46, AMS Providence R.I., 1987) 15–46

[14] T. Gaffney: Polar multiplicities and equisingularity of map germs *Topology* **32(1)** (1993) 185–223

[15] C. G. Gibson, K. Wirthmüller, A. A. du Plessis and E. J. N. Looijenga: *Topological Stability of Smooth Mappings* (SLNM 552, Springer Verlag, 1976)

[16] M. Goresky: Triangulation of stratified objects *Proc. Amer. Math. Soc.* **72** (1978) 193–200

[17] M. Goresky, R. MacPherson: *Stratified Morse theory* (Ergebnisse der Mathematik und ihrer Grenzgebiete (3), 14. Springer-Verlag, Berlin, 1988)

[18] V. V. Goryunov: Semi-simplicial resolutions and homology of images and discriminants of mappings *Proc. London Math. Soc.* **70** (1995) 363–385

[19] V. V. Goryunov and D. Mond: Vanishing cohomology of singularities of mappings *Compositio Math.* **89** (1993) 45–80

[20] A. Grothendieck: *Cohomologie locale des faisceaux cohérents et théorèmes de Lefschetz locaux et globaux (SGA 2). Séminaire de Géométrie Algébrique du Bois-Marie 1962* (Advanced Studies in Pure Mathematics, Vol. 2. North-Holland Publishing Co., Amsterdam, 1968)

[21] H. A. Hamm: Topology of isolated singularities of complex spaces *In: Proceedings of Liverpool Singularities Symposium II* (1969/1970, Lecture Notes in Math., Vol. 209, Springer, Berlin, 1971) 213–217

[22] H. A. Hamm: Zum Homotopietyp q-vollständiger Räume *J. Reine Angew. Math.* **364** (1986) 1–9

[23] H. A. Hamm: On stratified Morse theory *Topology* **38(2)** (1999) 427–438

[24] H. A. Hamm, D. T. Lê: Rectified homotopical depth and Grothendieck conjectures *In: The Grothendieck Festschrift* (Vol. II, Progr. Math., 87, Birkhäuser Boston, Boston, MA, 1990) 311–351

[25] A. Horváth and A. Némethi: On the Milnor fiber of non-isolated singularities *Studia Sci. Math. Hungar.* **43(1)** (2006) 131–136

[26] K. Houston: Local Topology of Images of Finite Complex Analytic Maps *Topology* **36** (1997) 1077–1121

[27] K. Houston: Global Topology of Images of Finite Complex Analytic Maps *Mathematical Proceedings of the Cambridge Philosophical Society* **122** (1997) 491–501

[28] K. Houston: Images of finite maps with one-dimensional double point set *Topology Appl.* **91(3)** (1999) 197–219

[29] K. Houston: On the topology of augmentations and concatenations of singularities *Manuscripta Math.* **117(4)** (2005) 383–405

[30] K. Houston: General Image Computing Spectral Sequence *Preprint* June 2006, To appear in the Proceedings of the Singularities Conference in Luminy in Feb 2005 (2006)

[31] K. Houston and N. P. Kirk: On the Classification and Geometry of Corank 1 Map-Germs from Three-Space to Four-Space *In: Singularity Theory* (ed. J.W. Bruce and D.M.Q. Mond, LMS Lecture note series 263, CUP, 1999) 325–351

[32] S. Izumiya and W. L. Marar: On Topologically Stable Singular Surfaces in a 3-Manifold *Journal of Geometry* **52** (Springer Verlag, Basel), (1995) 108–119

[33] M. Kato, Y. Matsumoto: On the connectivity of the Milnor fiber of a holomorphic function at a critical point *In: Manifolds* (Tokyo 1973 (Proc. Internat. Conf., Tokyo, 1973), Univ. Tokyo Press, Tokyo, 1975) 131–136

[34] M. A. Kervaire, J. Milnor: Groups of homotopy spheres. I. *Ann. of Math.* **77(2)** (1963) 504–537

[35] H. King: Topology of isolated critical points of functions on singular spaces *In: Stratifications, Singularities and Differential Equations* (Vol. II: Stratifications and Topology of Singular Spaces, ed. D. Trotman and L. Wilson, Hermann, Paris, 1997) 63–72

[36] D. T. Lê: Calcul du nombre de cycles évanouissants d'une hypersurface complexe *Ann. Inst. Fourier (Grenoble)* **23(4)** (1973) 261–270

[37] D. T. Lê: Computation of the Milnor number of an isolated singularity of a complete intersection *Functional Analysis and its Applications* **8** (1974) 127–131

[38] D. T. Lê: Some remarks on relative monodromy *In: Real and complex singularities* (Proc. Ninth Nordic Summer School/NAVF Sympos. Math., Oslo, 1976, Sijthoff and Noordhoff, Alphen aan den Rijn, 1977) 397–403

[39] D. T. Lê: Le concept de singularité isolee de fonction analytique *Advanced Studies in Pure Math.* **8** (1986) 215–227

[40] S. Lefschetz: *L'analysis situs et la géométrie algébrique* (Gauthier-Villars, Paris, 1950)

[41] E. J. N. Looijenga: *Isolated singular points on complete intersections* (London Mathematical Society Lecture Note Series, 77. Cambridge University Press, Cambridge, 1984)

[42] E. Looijenga, J. H. M. Steenbrink: Milnor number and Tjurina number of complete intersections *Math. Ann.* **271(1)** (1985) 121–124

[43] W. L. Marar and D. M. Q. Mond: Multiple point schemes for corank 1 maps *J. London Math. Soc.* **39(2)** (1989) 553–567

[44] D. B. Massey: *Lê cycles and hypersurface singularities* (Lecture Notes in Mathematics, 1615. Springer-Verlag, Berlin, 1995)

[45] D. B. Massey: The Sebastiani-Thom isomorphism in the derived category *Compositio Math.* **125(3)** (2001) 353–362

[46] J. N. Mather: Notes on Topological Stability *Preprint* available from `http://www.math.princeton.edu/directory/`, July (1970)

[47] J. N. Mather: Stability of C^∞ mappings. III. Finitely determined mapgerms *Inst. Hautes Études Sci. Publ. Math.* **35** (1968) 279–308

[48] J. Milnor: *Morse theory* (Annals of Mathematics Studies, No. 51 Princeton University Press, Princeton, N.J., 1963)

[49] J. Milnor: *Singular points of complex hypersurfaces* (Annals of Mathematics Studies, No. 61 Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1968)

[50] B. Morin: Formes canoniques des singularités d'une application différentiable *C. R. Acad. Sci. Paris* **260** (1965) 5662–5665

[51] A. Némethi: The Milnor fiber and the zeta function of the singularities of type $f = P(h, g)$ *Compositio Math.* **79(1)** (1991) 63–97

[52] A. Némethi: Hypersurface singularities with 2-dimensional critical locus *J. London Math. Soc. (2)* **59(3)** (1999) 922–938

[53] E. Picard, G. Simart: *Théorie des fonctions algébriques de deux variables indépendantes* (Paris: Gauthier-Villars et Fils. VI + 246, 1897)

[54] M. J. Pflaum: *Analytic and Geometric Study of Stratified Spaces* (Springer Lecture Notes in Mathematics, Vol. 1768, Springer-Verlag, Berlin, 2001)

[55] K. Sakamoto: Milnor fiberings and their characteristic maps *In: Manifolds* (Tokyo 1973 (Proc. Internat. Conf., Tokyo, 1973), Univ. Tokyo Press, Tokyo, 1975) 145–150

[56] A. Sard: The measure of the critical values of differentiable maps *Bull. Amer. Math. Soc.* **48** (1942) 883–890

[57] J. Schürmann: *Topology of singular spaces and constructible sheaves* (Mathematics Institute of the Polish Academy of Sciences. Mathematical Monographs (New Series), 63. Birkhäuser Verlag, Basel, 2003)

[58] M. Sebastiani and R. Thom: Un résultat sur la monodromie *Invent. Math.* **13** (1971) 90–96

[59] D. Siersma: The vanishing topology of non isolated singularities *In: New developments in singularity theory* (Cambridge, 2000, NATO Sci. Ser. II Math. Phys. Chem., 21, Kluwer Acad. Publ., Dordrecht, 2001) 447–472

[60] S. Smale: Marston Morse (1892–1977) *Math. Intelligencer* **1(1)** (1978/79) 33–34

[61] M. Tibăr: Bouquet decomposition of the Milnor fibre *Topology* **35(1)** (1996) 227–241

[63] D. Trotman: Comparing Regularity Conditions on Stratifications *In: Singularities* (Proc. Symp Pure Math. 40 (Part 2) AMS Providence R.I., 1983) 575–586

[64] V. A. Vassiliev: *Ramified integrals, singularities and lacunas* (Mathematics and its Applications, 315, Kluwer Academic Publishers Group, Dordrecht, 1995)

[65] V. A. Vassiliev: *Applied Picard-Lefschetz theory* (Mathematical Surveys and Monographs, 97, American Mathematical Society, Providence, RI, 2002)

[63] A. Verona: *Stratified mappings – structure and triangulability* (Lecture Notes in Mathematics, 1102, Springer-Verlag, Berlin, 1984)

[66] C. T. C. Wall: Finite determinacy of smooth map-germs *Bull. Lond. Math. Soc.* **13** (1981) 481–539

[67] C. T. C. Wall: *Singular points of plane curves* (London Mathematical Society Student Texts, 63, Cambridge University Press, Cambridge, 2004)

KEVIN HOUSTON

School of Mathematics, University of Leeds, Leeds, LS2 9JT, U.K.

E-mail: `k.houston@leeds.ac.uk`

Handbook of Global Analysis
Demeter Krupka and David Saunders

Group actions and Hilbert's fifth problem

Sören Illman

Contents

Introduction

In the first section we consider Hilbert's fifth problem concerning Lie's theory of transformation groups. In his fifth problem Hilbert asks the following. Given a continuous action of a locally euclidean group G on a locally euclidean space M, can one choose coordinates in G and M so that the action is real analytic? We discuss the affirmative solutions given in Theorems 1.1 and 1.2, and also present known counterexamples to the general question posed by Hilbert. Theorem 1.1 is the celebrated result from 1952, due to Gleason, Montgomery and Zippin, which says that every locally euclidean group is a Lie group. Theorem 1.2 is a more recent result, due to the author, which says that every Cartan (thus in particular, every proper) C^s differentiable action, $1 \leq s \leq \infty$, of a Lie group G is equivalent to a real analytic action.

In the remaining part of the article, Sections 2–18, we give a complete, and to a very large extent selfcontained, proof of Theorem 1.2. This tour brings us into many different topics within the theory of transformation groups, as the above list of contents shows.

1 Hilbert's fifth problem

At the Second International Congress of Mathematicians in Paris 1900, Hilbert posed twenty-three mathematical problems, which have had a great impact on mathematical research ever since then up to the present day, see [14]. Of these problems the fifth problem is concerned with Lie's theory of transformation groups, and in a second part of the problem with what Hilbert calls "infinite groups," which are not groups in the modern use of the term. The questions in this second part of the fifth problem concern functional equations and difference equations, and have for example connections with the work of N. H. Abel. These questions lie completely outside the theory of transformation groups, and we shall not discuss them here any further.

Recall that a topological transformation group consists of a topological group G and a topological space X, together with a continuous action of the group G on X, i.e., a continuous map

$$\Phi\colon G \times X \to X, \ \ (g, x) \mapsto gx, \tag{1}$$

with the following two properties,

i) $ex = x$, for all $x \in X$, where e is the identity element in G, and

ii) $g(g'x) = (gg')x$, for all $g, g' \in G$, and all $x \in X$.

We shall in this article automatically assume that all given topological spaces are Hausdorff spaces. By an *m-dimensional topological manifold* we mean a topological space M such that each point in M has an open neighborhood which is homeomorphic to an open subset of $\mathbb{R}^m$.

A *Lie group* G is a topological group G, which at the same time is a real analytic manifold, and the multiplication $\mu\colon G \times G \to G,\ (g, g') \mapsto gg'$, and $\iota\colon G \to G,\ g \mapsto g^{-1}$, are real analytic maps.

In the case when G is a Lie group and M is a C^t differentiable manifold, $1 \leq t \leq \omega$, and we are given an action $\Phi\colon G \times M \to M$ of G on M that is a C^t differentiable map, we speak of a C^t differentiable transformation group. We also say in this case that M is a C^t differentiable G-manifold, $1 \leq t \leq \omega$. (Here $1 \leq t \leq \omega$, means that t is a positive integer $1 \leq t < \infty$, or that $t = \infty$ or $t = \omega$, where ω stands for real analytic.) When $t = \infty$ we also speak of a smooth transformation group, or we say that M is a smooth G-manifold. In the case when $t = \omega$, i.e., when M is a real analytic manifold and the action $\Phi\colon G \times M \to M$ of G on M is real analytic, we speak of a real analytic transformation group, or we say that M is a real analytic G-manifold. In complete analogy with the definition of a Lie group, one should call a real analytic transformation group a *Lie transformation group.*

In his fifth problem Hilbert asks the following. Let G be a locally euclidean topological group, and let M be a locally euclidean topological space (i.e., G is a topological group which at the same time is a topological manifold, and M is a topological manifold) and suppose that we are given a continuous action

$$\Phi\colon G \times M \to M \tag{2}$$

of G on M. Is it then always possible to choose the local coordinates in G and M in such a way that the action Φ becomes real analytic? In other words, is it possible to give the topological manifolds G and M real analytic structures so that Φ is real analytic?

In his discussion of the fifth problem Hilbert also expresses the possibility that some assumption of differentiability is actually unavoidable for a positive answer to the question in (2). Hilbert mentions the theorem, announced by Lie [31] but first proved by F. Schur [51], which says that any transitive C^2 differentiable transformation group can be made real analytic by means of suitable coordinate changes. This result is often considered to be the origin of Hilbert's fifth problem, see for example [50], p. 177-178.

Let us first discuss the special case of Hilbert's question where $M = G$. In this case the question is whether we can give G a real analytic structure so that the multiplication

$$\mu\colon G \times G \to G \tag{3}$$

is real analytic.

In this special case the answer to Hilbert's question is always *yes*. This affirmative answer is obtained by combining the result in Gleason [11] with the result in Montgomery-Zippin [38], and we can express the combined result as follows.

Theorem 1.1 *Every locally euclidean group is a Lie group.*

We say that a topological group G has no small subgroups if there exists a neighborhood of the identity element which contains no other subgroup than the trivial subgroup $\{e\}$. It is easy to see from the structure of one-parameter subgroups of a Lie group that a Lie group has no small subgroups, see [8], p. 193. Gleason proves in [11] that every finite-dimensional, locally compact topological group G that has no small subgroups is a Lie group. In [38] Montgomery and Zippin prove, by inductively using the above result of Gleason, that a locally connected, finite-dimensional, locally compact topological group has no small subgroups. Since a locally euclidean topological group is clearly both locally connected, locally compact, and finite-dimensional, we see that [11] and [38] together prove Theorem 1.1. The work by Gleason was further extended by Yamabe [56], [57], who proved, without any assumption of finite-dimensionality, that every locally compact topological group which has no small subgroups is a Lie group.

The formulation "is a Lie group" in Theorem 1.1 above is due to the fact that if a locally euclidean topological group G can be given a real analytic structure such that G is a Lie group, then the real analytic structure on G is uniquely determined. This is a consequence of the well-known result, which says that every continuous homomorphism between two Lie groups is real analytic, see Theorem 2.3 and Corollary 2.4.

Theorem 1.1 is sometimes regarded as the solution of Hilbert's fifth problem, but as we have noted Hilbert's question is more general and is concerned with transformation groups. Compare also with the remark by Montgomery in [37], p. 185. We refer to Montgomery [36], p. 442-443 for some interesting speculation, made in 1950, concerning the possible answers to Hilbert's general question in (2). An authoritative and very good discussion of the state in 1955 of Hilbert's fifth problem is given by Montgomery and Zippin in [40], Section 2.15.

Before the result in Theorem 1.1 was proved by Gleason, Montgomery and Zippin, the result had been known in some special cases. It follows by von Neumann [43] that Theorem 1.1 holds when G is compact. For commutative groups Theorem 1.1 was proved by Pontryagin [49], Theorem 44, and for solvable groups by Chevalley [7] and Mal'cev [32]. Iwasawa in his famous paper [24] also proved Theorem 1.1 for solvable groups, and in addition established many other far reaching results.

Let us here also mention that it is proved in Pontryagin [49], Chapter IX, that each C^r differentiable group, $r \geq 3$, is a Lie group, and that G. Birkhoff proved in [3] that each C^1 differentiable group is a Lie group.

Let us now return to Hilbert's general question whether it is possible to give G and M real analytic structures such that the group action in (2) becomes real analytic. We have already seen in Theorem 1.1 that a locally euclidean group G can always be given a real analytic structure so that it becomes a Lie group, and that such a real analytic structure on G is uniquely determined. Hence we can now assume that G is a Lie group, and that M is a topological manifold on which G acts by a continuous action Φ as in (2), and we are asking if M can be given a real analytic structure such that Φ becomes real analytic.

In [1] Bing constructs a continuous action of $\mathbb{Z}_2$ on $\mathbb{R}^3$ that cannot be C^r differentiable for any $r \geq 1$, and hence in particular it cannot be real analytic. If one in Bing's example instead considers the action to be on the one-point compactification S^3 of $\mathbb{R}^3$, one obtains a continuous action of $\mathbb{Z}_2$ on S^3, with the property that the fixed point set is an Alexander horned sphere in S^3. (An Alexander horned sphere is an imbedded, not locally flat, 2-sphere Σ^2 in S^3.) Montgomery and Zippin [39] modified the example of Bing to give an example of a continuous action of the circle S^1 on R^4 that cannot be C^r differentiable for any $r \geq 1$, and hence in particular it cannot be real analytic.

In both these examples, in [1] and [39], the action is *not locally smooth,* in the sense of [5], Section IV.1. (Some authors use the term *locally linear action*, instead of locally smooth action.) But there also exist continuous locally smooth actions that cannot be made smooth, and hence not either real analytic. For example, there exists a 12-dimensional, compact smoothable manifold M, which admits a locally smooth effective action of S^1, but which does not admit a non-trivial smooth action of S^1 in any of the smooth structures on M, see [5], Corollary VI.9.6. Thus we see that the answer to the general question in (2) is *no* in the case of continuous actions. One may in fact point out that the answer to the general question in (2) is no even for the trivial group $G = \{e\}$, since there exist topological manifolds that do not have any smooth structure, and hence also no real analytic structure. The first example of such a manifold was given by Kervaire [27].

In [40], p. 70, the following easy example of a smooth, i.e., a C^∞ differentiable, action that cannot be real analytic is given. The group is the group of reals $\mathbb{R}$, and it acts on the plane by the map

$$\Phi\colon \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2,$$

where $\Phi(t, re^{i\varphi}) = e^{i\alpha(r)t} \cdot re^{i\varphi}$, for all $t \in \mathbb{R}$ and all $re^{i\varphi} \in \mathbb{R}^2$, $r \geq 0$ and $\varphi \in \mathbb{R}$. Here

$$\alpha \colon \mathbb{R} \to \mathbb{R}$$

is a C^∞ differentiable function such that

$$\alpha(x) = 0, \text{ for all } x \leq 1,$$
$$\alpha(x) = 1, \text{ for all } x \geq 2.$$

Clearly Φ is a C^∞ differentiable map, and Φ is an action of the group of reals $\mathbb{R}$ on $\mathbb{R}^2$. Note that for $0 \leq r \leq 1$ we have that

$$\Phi(t, re^{i\varphi}) = re^{i\varphi}, \quad \text{for all } \; t \in \mathbb{R}.$$

For $r \geq 2$ we have that

$$\Phi(t, re^{i\varphi}) = e^{it} \cdot re^{i\varphi}, \quad \text{for all } \; t \in \mathbb{R}.$$

Thus the action is the trivial action in the unit disk, and outside the open disk with radius 2 the action of $\mathbb{R}$ is by standard rotation in the plane.

The above action of $\mathbb{R}$ on $\mathbb{R}^2$ cannot be real analytic, in any real analytic structure on $\mathbb{R}^2$. This is because the action is the trivial action in the open unit disk $\mathring{D}^2$, and thus, if the action was real analytic, it would have to be the trivial action in the whole plane, which is not the case.

If G is a compact Lie group the above kind of phenomena cannot occur. It is proved in Matumoto and Shiota [34], Theorem 1.3, that if G is a compact Lie group then every C^s differentiable action, $1 \leq s \leq \infty$, of G on a second countable C^s differentiable manifold can be made into a real analytic action. The technique of the proof in [34] is the same as in Palais [46], which in turn is an equivariant version of Whitney's proof [55] of the fact that every second countable C^s smooth manifold can be given a real analytic structure, compatible with the given C^s differentiable structure, $1 \leq s \leq \infty$.

How about the general case with actions of an arbitrary Lie group G? We saw in the elementary example above that there exist smooth actions, that is C^∞ differentiable actions, that cannot be real analytic. However we were able to prove the following, see [16].

Theorem 1.2 *Let G be a Lie group which acts on a C^s differentiable manifold M by a C^s differentiable Cartan action, $1 \leq s \leq \infty$. Then there exists a real analytic structure β on M, compatible with the given C^s differentiable structure on M, such that the action of G on M_β is real analytic.*

In Theorem 1.2 the manifold M *is not assumed to be paracompact*. But we wish to emphasize that this great generality *is in no way* a main point of the result. The chief interest of Theorem 1.2 is of course when M is paracompact or second countable.

An action of a Lie group G, on any locally compact Hausdorff space X, is said to be *Cartan* if each point x in X has a compact neighborhood A such that

$$G[A] = \{g \in G \mid gA \cap A \neq \emptyset\}$$

is a compact subset of G. We also say in this case that X is a Cartan G-space. This terminology was introduced by Palais in [45].

The notion of a Cartan action is a generalization of the more widely known notion of a *proper action*. We recall that an action of G on a locally compact space X is proper if $G[A]$ is a compact subset of G for any compact subset A of X. There exist smooth actions of Lie groups that are Cartan but not proper, and such actions have non-Hausdorff orbit spaces, see Example 4.10 and cf. Corollary 4.16.

When G is a discrete group the notion of a proper action coincides with the classical notion of a *properly discontinuous action*. Thus Theorem 1.2 in particular holds in the case of properly discontinuous actions of any discrete group G. (The reader should note that in the literature the meaning of the term "properly discontinuous" varies. For example, some authors use the term "properly discontinuous action" to mean a Cartan action of a discrete group G. But as far as the result in Theorem 1.2 is concerned this difference in the use of terminology does not cause any problems here, since Theorem 1.2 in any case holds for Cartan actions.)

Theorem 1.2 answers Hilbert's question concerning which group actions can be made real analytic. Furthermore the answer is best possible since, as we have seen above, there exist differentiable, in fact C^∞ differentiable, non-Cartan actions of Lie groups that cannot be made real analytic.

Concerning the proof of Theorem 1.2 we note the following. In Theorem 1.2 the group G is assumed to be an arbitrary Lie group, that is we are *not* restricting our attention to actions of linear Lie groups. (See Birkhoff [2] for the first example of a connected Lie group which is not a linear Lie group.) Hence one cannot in general, even if M is assumed to be second countable, imbed the G-manifold M as a G-invariant subset of some finite-dimensional linear representation space for G. Therefore it is *not possible* to prove Theorem 1.2, even if we assume that M is second countable, by using some equivariant version of Whitney's method [55], as was done in the case when G is *compact* in [34]. Instead we use a maximality argument, involving the use of Zorn's lemma, for the global part of the proof of Theorem 1.2. This argument is analogous to the one used by W. Koch and D. Puppe in [28], in a non-equivariant situation. However the main work goes into the part of the proof which deals with questions of a more local nature. An important role in the technical part of the proof of Theorem 1.2 is played by a result concerning approximations of C^s slices, $1 \leq s \leq \infty$, see Lemma 17.1.

In [16] we gave Theorem 1.2 in the smooth case, i.e., in the C^∞ case. We should also here remark that Lemma 2.3 in [16], given for the strong C^∞ topology, is not correct as stated. This mistake was pointed out to me by Sarah Packman (a graduate student at Berkeley) [44]. The best way to correct this mistake is to simply replace Lemma 2.3 in [16] by its valid very-strong C^∞ topology version, see Lemma 8.1 in [21] and Lemma 12.1 in the present article. As a consequence of this change one also needs to make another change in [16]. Theorem 1.2 in [34], cf. Theorem 2.1 in [16], cannot anymore be used, one needs a corresponding result for the very-strong C^∞ topology. Such a result is proved in Theorem 7.2 in [21], see Theorem 16.6 in the present article. After these two replacements the proof of the C^∞ case of Theorem 1.2 given in [16] requires no further changes, and is correct as it stands. We gave this correction to [16] the first time in [18], see also [20] and [21].

2 Lie groups and manifolds

By an *m-dimensional topological manifold M* we mean a topological Hausdorff space M such that each point $x \in M$ has an open neighborhood U which is homeomorphic to an open subset of $\mathbb{R}^m$, $m \geq 0$. Thus we do not automatically make any assumption concerning paracompactness or second countability of M, we will always mention such assumptions separately. Every second countable topological manifold is paracompact, but the converge does not hold in general. For example, an uncountable disjoint topological union of real lines is a paracompact 1-dimensional manifold, but it is not second countable. The relation between second countable and paracompact manifolds is given by the following well known result.

Proposition 2.1 *A topological manifold M, with or without boundary, is paracompact if and only if each connected component of M is second countable.*

Proof Follows by Théorème 3 and Corollaire in [9], and Theorem 5 in Chapter I, §9, no. 10 in [4]. □

By a C^t *differentiable manifold M*, where $1 \leq t \leq \omega$, we mean a topological manifold M together with a C^t *differentiable structure α on M*. A C^t differentiable structure α on M is, by definition, a *maximal* C^t *differentiable atlas α on M*, $1 \leq t \leq \omega$. In this terminology "C^ω differentiable" means "real analytic". Thus a real analytic manifold M is a topological manifold M together with a *real analytic structure α on M*, and a real analytic structure on M is by definition a *maximal real analytic atlas on M*.

We use the exact analogue of the above definition to define the notion of a C^t *differentiable manifold M with boundary,* $1 \leq t \leq \omega$. In this way we have now defined the notion of a C^t *differentiable manifold M, with or without boundary,* $1 \leq t \leq \omega$. In order to shorten the terminology we will in the following drop the word "differentiable", and for example speak about a C^t manifold M, with or without boundary, $1 \leq t \leq \omega$. Suppose that $M_\beta = (M, \beta)$ is a C^t manifold, where $1 \leq t \leq \omega$, and let $1 \leq s < t$. The C^t structure β is, by definition, a maximal C^t atlas on the topological manifold M. Since $1 \leq s < t$ we have in particular that β is a C^s atlas on M. Therefore β uniquely determines a maximal C^s atlas α on M. Thus $M_\alpha = (M, \alpha)$ is a C^s manifold, and we call M_α the C^s manifold determined by the C^t manifold M_β, $1 \leq s < t \leq \omega$. We have that $\beta \subset \alpha$ and we also say that the C^t structure β is *compatible* with the C^s structure α.

We say that a C^s manifold M_α, where $1 \leq s \leq \infty$, is *equivalent* to a real analytic manifold, if there is a real analytic manifold N_δ and a C^s diffeomorphism $f\colon M_\alpha \to N_\delta$, i.e., f is a C^s diffeomorphism between M_α and the C^s manifold N_γ determined by the real analytic manifold N_δ. Now the C^s structure induced from N_γ through f equals α, i.e., $f^*(\gamma) = \alpha$. Let us denote $f^*(\delta) = \beta$. Then β is a real analytic structure on M, and since $\delta \subset \gamma$ we have that $f^*(\delta) \subset f^*(\gamma)$, that is $\beta \subset \alpha$. This shows that the real analytic structure β on M is compatible with the given C^s structure α on M.

Thus we have seen that a C^s manifold $M_\alpha = (M, \alpha)$, $1 \leq s \leq \infty$, is equivalent to a real analytic manifold, if and only if, there exists a real analytic structure β on M that is compatible with the given C^s structure α on M.

A topological group G is a *Lie group,* if G is a real analytic manifold and both the multiplication $\mu : G \times G \to G$, $(g, g') \mapsto gg'$, and the inverse map $\iota : G \to G$, $g \mapsto g^{-1}$, are real analytic maps. Here below we record two fundamental results about Lie groups.

Theorem 2.2 *Suppose H is a closed subgroup of a Lie group G. Then H is a Lie group.*

Proof See e.g. Theorem II.2.3 in [13]. □

We will use Theorem 2.2 freely, without referring to it.

Theorem 2.3 *Let $\gamma : G \to G'$ be a continuous homomorphism between Lie groups. Then γ is real analytic.*

Proof See e.g. Theorem II.2.6 in [13]. □

We already referred to Theorem 2.3 in Section 1, more precisely we used:

Corollary 2.4 *Let G be a topological group, and suppose that there exist real analytic structures α_1 and α_2 on G such that $G_1 = (G_1, \alpha_1)$ and $G_2 = (G, \alpha_2)$ are Lie groups. Then $\alpha_1 = \alpha_2$, and thus $G_1 = G_2$.*

Proof By Theorem 2.3 the maps $\mathrm{id} : G_1 \to G_2, g \mapsto g$, and $\mathrm{id} : G_2 \to G_1, g \mapsto g$, are real analytic. Hence the claim follows. □

Proposition 2.5 *Let G be a Lie group. Then G is paracompact.*

Proof See, e.g. Remark on page 88 in [13]. □

3 Group actions

Definition 3.1 Let G be a topological group and let X be a topological space. An *action of G on X* is a (continuous!) map

$$\Phi : G \times X \to X \ , \ (g, x) \mapsto gx,$$

such that:

(a) $ex = x$, for all $x \in X$, where e is the identity element in G, and

(b) $g(g'x) = (gg')x$, for all $g, g' \in G$ and all $x \in X$

By a G-space X we mean a topological space X together with an action of G on X. It follows directly by Definition 3.1 that if X is a G-space, then the map $g\colon X \to X \ , \ x \mapsto gx$, is a homeomorphism of X, for each $g \in G$.

Thus we may think of a G-space X as a topological space X together with a group of homeomorphisms of X, but one should keep in mind that the precise definition of a G-space is based upon Definition 3.1. We say that the action of G on X is *effective* if the only element $g \in G$, which acts as the identity homeomorphism $\mathrm{id}\colon X \to X$, is the identity element $e \in G$, i.e., if $gx = x$, for all $x \in X$, then $g = e$.

Let $x \in X$, where X is a G-space. The set

$$Gx = \{gx \in X \mid g \in G\}$$

is called the G–orbit of x in X, or simply the orbit of x. Observe that the set of all G–orbits in X is a partition of X, and we denote the set of G–orbits by X/G. Furthermore we let $\pi : X \to X/G \ , \ x \mapsto Gx$, be the natural projection onto X/G. The topology on X/G is the quotient topology from $\pi : X \to X/G$, i.e., a subset $V \subset X/G$ is defined to be open in X/G if and only if $\pi^{-1}(V)$ is open in X. Suppose U is an open subset of X. Then $\pi^{-1}(\pi(U)) = \bigcup_{g \in G} gU$ is open in X, as a union of open subsets, and hence we have, by the definition of the quotient topology, that $\pi(U)$ is open in X/G. Thus we see that the natural projection $\pi : X \to X/G$ is an open (and continuous!) map. If $J \subset G$ and $A \subset X$, we denote

$$JA = \{ga \mid g \in J \text{ and } a \in A\},$$

and we say that A is G–invariant if $GA = A$.

Suppose X is a G-space, and let $x \in X$. We denote

$$G_x = \{g \in G \mid gx = x\}.$$

Then G_x is a closed subgroup of G, and it is called the *isotropy subgroup of* G *at* $x \in X$. If $g \in G$ then

$$G_{gx} = gG_xg^{-1}.$$

Thus we see that the isotropy subgroups at points on the some G-orbit are conjugate, in G, to each other. By a *G-isotropy type* we mean a conjugacy class [K] of a compact subgroup K of G. If X is a G-space, we say that the G-isotropy type [K] *occurs in* X, if there exists $x \in X$ such that $[G_x] = [K]$.

We say that a map $f\colon X \to Y$ is *G-equivariant*, if $f(gx) = gf(x)$, for each $g \in G$ and every $x \in X$. Note that if $f\colon X \to Y$ is a G-equivariant map between G-spaces, and $x \in X$, then

$$G_x \subset G_{f(x)}.$$

Proposition 3.2 *Let X be a K-space, where K is a compact group. Then the action $\Phi : K \times X \to X$, $(k, x) \mapsto kx$, is a closed map. In particular KA is closed in X, for every closed subset A of X.*

Proof See e.g. Theorem I.1.2 in [5]. □

Corollary 3.3 *Let X be a G-space, and let $x \in X$ be such that G_x is compact. Suppose W is any open neighborhood of x in X. Then there exists a G_x-invariant open neighborhood V of x in X, such that $V \subset W$.*

Proof *Proof.* The set $A = X - W$ is closed in X and $K = G_x$ is a compact subgroup of G. It follows by Proposition 3.2 that KA is closed in X. Now $x \notin KA = G_x(X - W)$, and hence $V = X - KA$ is a K-invariant open neighborhood of x in X, and $V \subset W$. □

Proposition 3.4 *Let X be a K-space, where K is a compact group. Then the orbit space X/K is Hausdorff, and the natural projection $\pi\colon X \to X/K$ is both a closed map and a proper map.*

Proof See e.g. [5], Theorem I.3.1. □

Now suppose that G is a Lie group and let M be a C^t manifold, where $1 \le t \le \omega$. A C^t *differentiable action*, or in short, a C^t *action,* of G on M, is an action, as in Definition 3.1, where the action map $\Phi\colon G \times M \to M$ is a C^t map, $1 \le t \le \omega$. Note that since G is a real analytic manifold, i.e., a C^ω manifold, and M is a C^t manifold, $1 \le t \le \omega$, it follows that the product $G \times M$ is a C^t manifold, $1 \le t \le \omega$. A C^t *G-manifold* M, $1 \le t \le \omega$, consists of a C^t manifold M together with a C^t action of G on M, $1 \le t \le \omega$.

Let M_β be a C^t G-manifold, where $1 \le t \le \omega$. Suppose that $1 \le s < t$, and let M_α be the C^s manifold determined by M_β, see Section 2. Then M_α is a C^s G-manifold. This is clear because β is a C^s atlas, although not a maximal one, on M, and since $\beta \subset \alpha$ it follows that $\Phi\colon G \times M_\alpha \to M_\alpha$, $(g, x) \mapsto gx$, is a C^s map.

The problem we are faced with, and which we solve in Theorem 1.2, concerns the opposite direction to the above, and $t = \omega$. That is given a C^s G-manifold M_α, where $1 \leq s < \omega$ and $\Phi\colon G \times M_\alpha \to M_\alpha$, $(g, x) \mapsto gx$, denotes the given C^s action, does there exist a real analytic structure β on M such that:

(a) The real analytic structure β is compatible with the given C^s structure α, i.e., $\beta \subset \alpha$.

(b) The action $\Phi\colon G \times M_\beta \to M_\beta$, $(g, x) \mapsto gx$, is a real analytic map.

We say that a C^s G-manifold $M_\alpha = (M, \alpha)$, where $1 \leq s \leq \infty$, *is equivalent to a real analytic G-manifold*, or that the C^s action (of G on M_α) is equivalent to a real analytic action, if there exist a real analytic G-manifold N_δ and a G-equivariant C^s diffeomorphism $f\colon M_\alpha \to N_\delta$. Let N_γ denote the C^s manifold determined by the real analytic manifold N_δ, see Section 2. Then $f^*(\gamma) = \alpha$. We define $\beta = f^*(\delta)$. Then β is a real analytic structure on M, and $\beta \subset \alpha$ since $\delta \subset \gamma$ and $f^*(\gamma) = \alpha$. Thus the real analytic structure β on M is compatible with the given C^s structure α on M. Let $\Psi\colon G \times N_\delta \to N_\delta$ denote the real analytic action of G on N_δ, and let $\Phi\colon G \times M_\beta \to M_\beta$, $(g, x) \mapsto gx$, be the action of G on M_β. Since $f\colon M_\alpha \to N_\delta$ is a G-equivariant map it follows that $\Psi \circ (\mathrm{id} \times f) = f \circ \Phi$. Since $f\colon M_\beta \to N_\delta$ is a real analytic isomorphism and Ψ is real analytic it now follows that $\Phi\colon G \times M_\beta \to M_\beta$ is real analytic.

Thus we have seen that a C^s G-manifold M_α, where $1 \leq s \leq \infty$, is equivalent to a real analytic G-manifold, if and only if, there exists a real analytic structure β on M, such that conditions (a) and (b) above hold.

Let $\mathrm{Gl}(n, \mathbb{R})$ denote the general linear group of all non-singular $(n \times n)$-matrices with entries in $\mathbb{R}$. (Alternatively the reader may wish to think of $\mathrm{Gl}(n, \mathbb{R})$ as the group of all linear automorphisms of the vector space $\mathbb{R}^n$.) Then $\mathrm{Gl}(n, \mathbb{R})$ is a Lie group, and the natural action of $\mathrm{Gl}(n, \mathbb{R})$ on $\mathbb{R}^n$,

$$\Phi : \mathrm{Gl}(n, \mathbb{R}^n) \times \mathbb{R}^n \to \mathbb{R}^n\ ,\ (T, x) \to Tx,$$

is real analytic. Here Tx denotes matrix multiplication (or, if the reader prefers the second interpretation of $\mathrm{Gl}(n, \mathbb{R})$, the value of the linear automorphism T at x.)

Now let G be a Lie group, and let $\theta : G \to \mathrm{Gl}(n, \mathbb{R})$ be a linear representation of G, that is θ is a continuous homomorphism from G into $\mathrm{Gl}(n, \mathbb{R})$. Since both G and $\mathrm{Gl}(n, \mathbb{R})$ are Lie groups it now follows that $\theta : G \to \mathrm{Gl}(n, \mathbb{R})$ is a real analytic homomorphism. We denote the representation space for G corresponding to the linear representation $\theta : G \to \mathrm{Gl}(n, \mathbb{R})$ by $\mathbb{R}^n(\theta)$. That is $\mathbb{R}^n(\theta)$ is a notation for the vector space $\mathbb{R}^n$ on which G acts by

$$\Theta : G \times \mathbb{R}^n \to \mathbb{R}^n\ ,\ (g, x) \mapsto gx = \theta(g)x.$$

Since the homomorphism $\theta : G \to \mathrm{Gl}(n, \mathbb{R})$ is real analytic, it follows that the action Θ is real analytic. For each $g \in G$, the map $g : \mathbb{R}^n \to \mathbb{R}^n$, $x \mapsto gx = \theta(g)x$, is a linear automorphism of $\mathbb{R}^n$.

Let $\theta, \psi\colon G \to \mathrm{Gl}(n; \mathbb{R})$ be two linear representations of G. We say that θ and ψ are equivalent if there exists $T \in \mathrm{Gl}(n; \mathbb{R})$ such that

$$\psi(g) = T\theta(g)T^{-1},\ \text{for all } g \in G.$$

In this case the linear isomorphism $T\colon \mathbb{R}^n(\theta) \to \mathbb{R}^n(\psi),\ \ x \mapsto T(x)$, is a G-equivariant map, since $T(gx) = T(\theta(g)x) = (T\theta(g)T^{-1}T)(x) = (\psi(g)T)(x) = \psi(g)(T(x)) = gT(x)$. Thus the linear representation spaces $\mathbb{R}^n(\theta)$ and $\mathbb{R}^n(\psi)$ are G-equivariantly, linearly, isomorphic.

Theorem 3.5 *Each linear representation $\theta\colon K \to \mathrm{Gl}(n, \mathbb{R})$ of a compact group K is equivalent to an orthogonal representation $\psi\colon K \to \mathrm{O}(n) \hookrightarrow \mathrm{Gl}(n, \mathbb{R})$.*

Proof See e.g. [5], Theorem 0.3.5. □

Now suppose that K is a compact Lie group, and let $\psi : K \to \mathrm{O}(n)$ be an orthogonal representation of K. Then the corresponding action $\Psi : K \times \mathbb{R}^n(\psi) \to \mathbb{R}^n(\psi)$ restrics to give an induced real analytic action of K on the unit sphere S^{n-1} and on the unit disk D^n, and we denote the corresponding K-spaces by $\mathrm{S}^{n-1}(\psi)$ and $\mathrm{D}^n(\psi)$. Then $\mathrm{S}^{n-1}(\psi)$ is a real analytic K-manifold, and $\mathrm{D}^n(\psi)$ is a real analytic K-manifold with boundary.

Theorem 3.6 *Let K be a compact Lie group and let $\mathbb{R}^n(\theta)$ be a linear representation space for K. Then the number of K-isotropy types occurring in $\mathbb{R}^n(\theta)$ is finite.*

Proof See e.g. [26], Corollary 4.25. □

4 Cartan and proper actions of Lie groups

By a G-space X we in this section mean a locally compact space X together with an action of G on X, where G is a Lie group, or equally well any locally compact group. Given a G-space X and any two subsets A and B of X we define

$$G[B, A] = \{g \in G \bigm| gA \cap B \neq \emptyset\}. \tag{i}$$

Furthermore we denote

$$G[A] = G[A, A]. \tag{ii}$$

We have that

$$G[A, B] = G[B, A]^{-1}. \tag{iii}$$

If $A_0 \subset A$ and $B_0 \subset B$, then

$$G[B_0, A_0] \subset G[B, A]. \tag{iv}$$

If $\{A_i\}_{i\in\Lambda}$ and $\{B_i\}_{i\in\Lambda}$ are families of subsets of X, then

$$G[B, \bigcup_{i\in\Lambda} A_i] = \bigcup_{i\in\Lambda} G[B, A_i] \ \text{ and } \ G[\bigcup_{i\in\Lambda} B_i, A] = \bigcup_{i\in\Lambda} G[B_i, A]. \tag{v}$$

Let $g_0 \in G$, then we obtain by direct verification that

$$G[g_0 B, A] = g_0 G[B, A] \ \text{ and } \ G[B, g_0 A] = G[B, A] g_0^{-1}. \tag{vi}$$

More generally, if J is a subset of G, then

$$G[JB, A] = JG[B, A], \ \text{ and } \ G[B, JA] = G[B, A] J^{-1}. \tag{vii}$$

Lemma 4.1 *Let X and Y be G-spaces and $f : X \to Y$ a surjective G-map. If $B \subset Y$ we have that*

$$G[B] = G[f^{-1}(B)].$$

Proof Direct verification. □

Lemma 4.2 *Let X be a G-space, and let A be a compact and B a closed subset of X. Then $G[B, A]$ is closed in G. Therefore also $G[A, B] = G[B, A]^{-1}$ is closed in G.*

Proof We show that $G - G[B, A]$ is open in G. Let $g \in G - G[B, A]$. Then $gA \cap B = \emptyset$, i.e., $gA \subset X - B$. Since A is compact, and the action map $G \times X \to X$, $(g, x) \mapsto gx$, is continuous, and $X - B$ is open in X, there exists an open neighborhood U of g in G, such that

$$UA \subset X - B. \tag{1}$$

Now $U \subset G - G[B, A]$, since if $h \in U$ then $hA \subset X - B$ by (1), i.e., $hA \cap B = \emptyset$, and hence $h \in G - G[B, A]$. Thus $G - G[B, A]$ is open in G. □

Definition 4.3 Let X be a G-space. We say that the action of G on X is *Cartan*, or that X is a *Cartan G-space*, if each point $x \in X$ has a compact neighborhood V such that $G[V]$ is compact.

Note that if X is a Cartan G-space, then the isotropy subgroup G_x at $x \in X$ is a compact subgroup of G, for every $x \in X$.

Now consider G itself as a G-space, where the group G acts on G by multiplication from the left. If $A \subset G$, then one obtains directly that

$$G[A] = AA^{-1}. \tag{viii}$$

Let H be a closed subgroup of G, where we let G be any locally compact group. Then the homogeneous space G/H is Hausdorff, see e.g. [40], Theorem on p. 27. Moreover G/H is locally compact, since G is locally compact and the natural projection $\pi \colon G \to G/H$, $g \mapsto gH$, is an open map. Thus the homogeneous space G/H, with the natural action of G given by $G \times G/H \to G/H$, $(g', gH) \mapsto g'gH$, is a G-space. If $B \subset G/H$ we have that

$$G[B] = \pi^{-1}(B)(\pi^{-1}(B))^{-1}, \tag{ix}$$

which is easy to see directly, or alternatively (ix) follows by Lemma 4.1 and (viii).

Note that the G-space G/H is Cartan if and only if H is a compact subgroup of G. The isotropy subgroup of G at eH equals H, so in order for G/H to be a Cartan G-space, it is necessary that H is compact. Now assume that H is compact and let B be a compact subset of G/H. Then $\pi^{-1}(B)$ is compact, see Proposition 3.4, and hence it follows by (ix) that $G[B]$ is compact.

We give the proof of Lemma 4.4 below under the assumption that G is any locally compact group and X is any locally compact space. If one in addition assumes that both G and X are first countable, for example if G is a Lie group and X is a topological manifold, one can use ordinary sequences instead of the more general notion of a net.

Lemma 4.4 *Suppose X is a Cartan G-space, and let $x \in X$. Then $\varphi_x : G \to X$, $g \mapsto gx$, is a closed map.*

Proof Let J be a closed subset of G. Our claim is that Jx is closed in X. Suppose $y \in \overline{Jx}$, and let $g_\alpha x \in Jx$, where $g_\alpha \in J$, $\alpha \in \Lambda$, be a net in Jx, such that $\lim_\alpha g_\alpha x = y$. Choose a compact neighborhood V of y, such that $G[V]$ is compact. We can assume that $g_\alpha x \in V$, for all $\alpha \in \Lambda$. Let $\alpha_0 \in \Lambda$ be fixed. Then the identity $(g_\alpha g_{\alpha_0}^{-1})(g_{\alpha_0} x) = g_\alpha x$ shows that $g_\alpha g_{\alpha_0}^{-1} \in G[V]$, for all $\alpha \in \Lambda$. Since $G[V]$ is compact a subnet of the net $g_\alpha g_{\alpha_0}^{-1}$, $\alpha \in \Lambda$, converges to a point in G, say to $\overline{g}$. By abuse of notation we may assume that the net $g_\alpha g_{\alpha_0}^{-1}$, $\alpha \in \Lambda$, converges to $\overline{g}$, i.e., that $\lim_\alpha g_\alpha g_{\alpha_0}^{-1} = \overline{g}$. Then $\lim_\alpha g_\alpha = \overline{g} g_{\alpha_0} = g$. Since $g_\alpha \in J$, for all $\alpha \in \Lambda$, and J is closed, we have that $g \in J$. Now $y = \lim_\alpha g_\alpha x = gx \in Jx$, which proves our claim. □

Corollary 4.5 *Suppose X is a Cartan G-space, and let $x \in X$. Then the orbit Gx is closed in X, and $\overline{\varphi}_x : G/G_x \to Gx$, $gG_x \mapsto gx$, is a G-equivariant homeomorphism.*

Proof Clearly $\overline{\varphi}_x$ is a G-equivariant bijection. The natural projection $\pi : G \to G/G_x$, $g \mapsto gG_x$, is an open map, and hence a quotient map. Since $\overline{\varphi}_x \circ \pi = \varphi_x$, and φ_x is continuous, it follows that $\overline{\varphi}_x$ is continuous. Since φ_x is a closed map, by Lemma 4.4, it follows that also $\overline{\varphi}_x$ is a closed map. Thus $\overline{\varphi}_x$ is a homeomorphism. □

It follows by Corollary 4.5 that if X is a Cartan G-space, then each point in the orbit space X/G is closed in X/G, i.e., X/G is a T_1 space.

Lemma 4.6 *Let X be a Cartan G-space, and let $x \in X$. Suppose U is an open neighborhood of G_x in G. Then there exists an open neighborhood V of x in X, such that $G[V] \subset U$. Furthermore V may be chosen to be G_x-invariant.*

Proof We have that $x \notin (G-U)x$, and by Lemma 4.4 the set $(G-U)x$ is closed in X. Let A be a compact neighborhood of x in X, such that $G[A]$ is compact and $A \cap (G-U)x = \emptyset$. Then

$$Q = G[A] \cap (G - U)$$

is a compact subset of G. Furthermore $Qx \subset (G - U)x \subset X - A$. Since Q is compact, and the action map $\Phi : G \times X \to X$, $(g, x) \mapsto gx$, is continuous, and $X - A$ is open in X, there exists an open neighborhood V of x in X, such that

$$QV \subset X - A. \tag{1}$$

Moreover we can choose V so that $x \in V \subset A$.

We claim that $G[V] \subset U$. Suppose $G[V] \not\subset U$, and let $g \in G[V] \cap (G - U)$. Now $G[V] \subset G[A]$, since $V \subset A$, and hence $g \in G[A] \cap (G - U) = Q$. Thus (1) gives us $gV \subset QV \subset X - A \subset X - V$. Therefore $gV \cap V = \emptyset$, which is a contradiction since $g \in G[V]$. Hence our claim that $G[V] \subset U$ holds. Since the isotropy subgroup G_x is compact, the last claim in Lemma 4.6 follows by Corollary 3.3. □

It follows directly by Definitions 4.7 (a), (b) and (c) below, that if X is a proper G-space in sense of Koszul, Palais or Borel, then X is in particular a Cartan G-space. By Corollary 4.9 these three notions of a proper G-space are equivalent. *Hence we can use anyone of the three definitions below as the definition of a proper G-space.* (This author prefers to take a proper G-space to mean a Borel proper G-space.)

Definition 4.7 (a) Let X be a G-space. We say that the action of G on X is *Koszul proper,* or that X is a *Koszul proper G-space*, if for any two points x and y in X, there exist compact neighborhoods V and W, of x and y, respectively, such that $G[W, V]$ is compact.

Definition 4.7 (b) Let X be a G-space. We say that the action of G on X is *Palais proper*, or that X is a *Palais proper G-space*, if for each point x in X there exists a compact neighborhood V_0 of x in X, such that every $y \in X$ has a compact neighborhood W for which $G[W, V_0]$ is compact.

Definition 4.7 (c) Let X be a G-space. We say that the action of G on X is *Borel proper*, or that X is a *Borel proper G-space*, if for each compact subset A of X the set $G[A]$ is compact.

Lemma 4.8 *Every Koszul proper G-space X is a Borel proper G-space.*

Proof Let X be a Koszul proper G-space, and let A and B be compact subsets of X. We shall show that $G[B, A]$ is compact. Let $x \in A$ and $y \in B$. We choose compact neighborhoods $V(x; y)$ and $W(y; x)$, of x and y, respectively such that $G[W(y; x), V(x; y)]$ is compact. The interior of the sets $W(y; x),\ y \in B$, form an open covering of the compact set B, and hence a finite number of the sets $W(y; x),\ y \in B$, say $W(y_j; x),\ 1 \leq j \leq q$, cover B. Thus $B \subset \bigcup_{j=1}^{q} W(y_j; x)$.

Let $V(x; y_j)$ denote the compact neighborhood of x, which corresponds to the compact neighborhood $W(y_j; x)$ of y_j, $1 \leq j \leq q$. Then $V(x; B) = \bigcap_{j=1}^{q} V(x; y_j)$ is a compact neighborhood of x in X. Now $G[B, V(x; B)]$ is compact since the set $G[B, V(x; B)]$ is closed in G, by Lemma 4.2, and

$$G[B, V(x; B)] \subset G[\bigcup_{j=1}^{q} W(y_j; x), V(x; B)]$$

$$= \bigcup_{j=1}^{q} G[W(y_j; x), V(x, B)] \subset \bigcup_{j=1}^{q} G[W(y_j; x), V(x; y_j)].$$

Here the last union is a compact set, since it is a finite union of compact sets.

Thus we have shown that each point $x \in A$ has a compact neighborhood $V(x; B)$, in X, such that $G[B, V(x; B)]$ is compact. Now the interior of the sets $V(x; B),\ x \in A$, form an open covering of A, and since A is compact it follows that a finite number of the sets $V(x; B),\ x \in A$, say $V(x_i; B)\ , 1 \leq i \leq p$, cover A. Thus $A \subset \bigcup_{i=1}^{p} V(x_i; B)$. Since $G[B, A]$ is closed in G, by Lemma 4.2, and

$$G[B, A] \subset G[B, \bigcup_{i=1}^{p} V(x_i; B) \subset \bigcup_{i=1}^{p} G[B, V(x_i; B)],$$

it follows that $G[B, A]$ is compact. In particular $G[A] = G[A, A]$ is compact, for any compact subset A of X, and thus X is a Borel proper G-space. □

Corollary 4.9 *Let X be a G-space. Then the following are equivalent:*

(a) *X is a Koszul proper G-space.*

(b) *X is a Palais proper G-space.*

(c) *X is a Borel proper G-space.*

Proof The fact that (a) implies (c) is given by Lemma 4.8. Suppose that (c) holds, and let $x \in X$. Choose a compact neighborhood V_0 of x in X. If $y \in X$, we let W be any compact neighborhood of y in X. Then $G[W, V_0] \subset G[V_0 \cup W, V_0 \cup W] = G[V_0 \cup W]$. Now $G[W, V_0]$ is closed in G, by Lemma 4.2, and $G[V_0 \cup W]$ is compact, since $V_0 \cup W$ is a compact subset of X and X is a Borel proper G-space. Hence $G[W, V_0]$ is compact. This shows that X is Palais proper, and thus (c) implies (b). The fact that (b) implies (a) is immediate from the definitions. □

A good example of a Cartan G-space X, which is not a proper G-space is the following one, see [30], Section I.1.

Example 4.10 Define an action of $\mathbb{R}$ on $\mathbb{R}^2 - \{0\}$ by, $(t, (x, y)) \mapsto (e^t x, e^{-t} y)$, for all $t \in \mathbb{R}$ and $(x, y) \in \mathbb{R}^2 - \{0\}$. This action is a C^∞ action, and it is Cartan but not proper. The orbit space of this action is not Hausdorff, for example the orbits of the points $(1, 0)$ and $(0, 1)$ are two different points in the orbit space which do not have disjoint neighborhoods. If one restricts this action to the integers $\mathbb{Z}$ one obtains an action of $\mathbb{Z}$ on $\mathbb{R}^2 - \{0\}$, which is Cartan but not proper.

Lemma 4.11 *Let X be a Cartan G-space and let $x \in X$. Then there exists a G-invariant open neighborhood V^* of x in X, such that V^* is a proper G-space.*

Proof Let V be an open neighborhood of x in X, such that $\overline{V} = A$ is compact and $G[A]$ is compact. Then $V^* = GV$ is an open subset of X, and hence V^* is locally compact. Thus V^* is a G-space, and we claim that V^* is a proper G-space. By Corollary 4.9 it is enough to prove that V^* is a Koszul proper G-space. Let $x, y \in V^* = GV$. Then $x = ga$ and $y = g'b$, where $g, g' \in G$ and $a, b \in V$. Now gV and $g'V$ are open neighborhoods of x and y, respectively, in V^*, and hence also in X. Since X is locally compact there exist compact neighborhoods W and W' of x and y, respectively, in V^*, such that $x \in W \subset gV$, and $y \in W' \subset g'V$. Now

$$G[W', W] \subset G[g'V, gV] = g'G[V, V]g^{-1} = g'G[V]g^{-1} \subset g'G[A]g^{-1}. \tag{1}$$

Since $G[W', W]$ is closed in G, by Lemma 4.2, and $g'G[A]g^{-1}$ is compact, it follows by (1) that $G[W', W]$ is compact. □

Lemma 4.12 *Suppose X is a Cartan G-space such that X/G is Hausdorff. Then X is a proper G-space.*

Proof By Corollary 4.9 it is enough to verify that X is Koszul proper. Let $x, y \in X$, and assume first that x and y belong to the same orbit, i.e., $y = gx$, for some $g \in G$. Let V be a compact neighborhood of x in X such that $G[V]$ is compact. Then gV is a compact neighborhood of y, and $G[gV, V] = gG[V, V] = gG[V]$ is compact.

Now assume that $\pi(x) \neq \pi(y) \in X/G$, and let V^* and W^* be disjoint open neighborhoods of $\pi(x)$ and $\pi(y)$, respectively, in X/G. Then $\pi^{-1}(V^*)$ and $\pi^{-1}(W^*)$ are disjoint G-invariant open neighborhoods of x and y, respectively. Let V and W be compact neighborhoods of x and y, respectively, such that $V \subset \pi^{-1}(V^*)$ and $W \subset \pi^{-1}(W^*)$. Then $G[W, V] \subset G[\pi^{-1}(W^*), \pi^{-1}(V^*)] = \emptyset$. □

Proposition 4.13 *Let X be a G-space. Then the following are equivalent:*

(a) $\Phi : G \times X \to X$, $(g, x) \mapsto gx$, *is a Borel proper action.*

(b) $\Phi^* : G \times X \to X \times X$, $(g, x) \mapsto (gx, x)$, *is a proper map.*

(c) *The restriction $\Phi| : G \times A \to X$ is a proper map, for each compact subset A of X.*

Proof Let $A \subset X$. A direct verification gives us the equality $G[A] = p(\Phi^{*-1}(A \times A))$, where $p : G \times X \to G$ denotes the projection. Using this equality we see that (a) and (b) are equivalent. Similarly we have for any $A \subset X$ and $B \subset X$ that $(\Phi|(G \times A))^{-1}(B) = \Phi^{*-1}(B \times A)$, and using this we see that (b) and (c) are equivalent. □

Corollary 4.14 *Let X be a proper G-space. Suppose that J is a closed subset of G and A is a compact subset of X. Then $JA = \{gx \mid g \in J \ , \ x \in A\}$ is a closed subset of X. In particular GA is closed in X.*

Proof By Corollary 4.9 we may use "a proper G-space" to mean "a Borel proper G-space". Hence it follows by Proposition 4.13 that the restriction $\Phi| : G \times A \to X$ is a proper map, and therefore also a closed map, since X is locally compact. (It is easy to see that if $f : Y \to X$ is a proper map and X is compactly generated, then f is a closed map, cf. e.g. Corollary in [47]. Furthermore every locally compact space X is compactly generated, see [52], 2.2.) Thus $JA = \Phi(J \times A)$ is closed in X. □

Corollary 4.15 *Let X be a proper G-space. Then X/G is Hausdorff.*

Proof Let $\overline{x}, \overline{y} \in X/G$, where $\overline{x} \neq \overline{y}$. Choose $x, y \in X$ such that $\pi(x) = \overline{x}$ and $\pi(y) = \overline{y}$, where $\pi : X \to X/G$ is the natural projection. Then $Gx \cap Gy = \emptyset$, and Gx and Gy are closed subsets of X, by Corollary 4.14. Let A be a compact neighborhood of x such that $A \cap Gy = \emptyset$. Then GA is closed in X, by Corollary 4.14, and $GA \cap Gy = \emptyset$ since $A \cap Gy = \emptyset$. Now $\pi(A^{\circ}) = \pi(GA^{\circ})$ and $\pi(X - GA)$ are disjoint open neighborhoods of $\overline{x}$ and $\overline{y}$, respectively, in X/G. □

Corollary 4.16 *Let X be G-space. Then the following are equivalent:*

(i) *X is a Cartan G-space and X/G is Hausdorff.*

(ii) *X is a proper G-space.*

Proof The fact that (i) implies (ii) is given by Lemma 4.12. As we already pointed out before Definitions 4.7 (a)–(c), it follows directly from the definitions that every proper G-space is a Cartan G-space. Hence the fact that (ii) implies (i) now follows by Corollary 4.15. □

Notes Lemma 4.6 is Proposition 1.1.6 in [45] and our proof is basically the same as, but somewhat simpler than, in [45]. Definition 4.7 (a) is from [30], Definition 4.7 (b) is Definition 1.2.2 in [45], and Definition 4.7 (c) occurs in Theorem 1.2.9 in [45]. Proposition 4.13 is taken from [19], Proposition 1.4. Concerning Corollary 4.16, compare with Theorem 1.2.9 in [45].

5 Non-paracompact Cartan G-manifolds

We prove in Theorem 5.1 that in a Cartan G-manifold M, every point has a paracompact G-invariant open neighborhood and thus M is locally a paracompact Cartan G-manifold. In fact it then follows by Lemma 4.11 that M is locally a paracompact proper G-manifold, see Addendum 5.2.

Let us here first note that if A and B are connected subsets of a G-space X, then the set

$$B \cup G[B,A]A = B \cup \bigcup_{g \in G[B,A]} gA \tag{i}$$

is connected since B is connected and each gA is connected, and $gA \cap B \neq \emptyset$ for all $g \in G[B,A]$.

Theorem 5.1 *Let M be a Cartan G-manifold, where G is a Lie group, and let $x_0 \in M$. Then there exists a paracompact G-invariant open neighborhood V^* of x_0 in M.*

Proof Since M is a Cartan G-manifold there exists an open neighborhood V of x_0 in M, such that V is homeomorphic to $\mathbb{R}^m$, $A = \overline{V}$ is compact, and $G[A]$ is a compact subset of G. Since V is homeomorphic to $\mathbb{R}^m$ it follows that V is connected and second countable. We claim that $V^* = GV$ is paracompact. It is enough to prove that each connected component of V^* is second countable, see Proposition 2.1.

Let W be the connected component of V^* which contains V. It is easy to see that each connected component of V^* is of the form $g'W$, for some $g' \in G$. Hence it is enough to prove that W is second countable.

Let us denote $J = G[V] = G[V,V]$. Then $J^2 = J \cdot J = JG[V] = JG[V,V] = G[JV,V]$, and by induction, i.e., by repeated use of the identity $G[JB,A] = JG[B,A]$, see (vii) in Section 4, we obtain that

$$J^{n+1} = G[J^nV,V], \text{ for all } n \geq 0, \text{ where } J^0 = \{e\}. \tag{1}$$

We claim that the sets

$$J^nV, \;\; n \geq 0 \tag{2}$$

are connected. Assume by induction that J^nV is connected. Since $e \in J$ it follows that $J^nV \subset J^{n+1}V$, and hence, by using (1), we obtain that

$$J^{n+1}V = J^nV \cup J^{n+1}V = J^nV \cup G[J^nV,V]V,$$

and this set is connected by (i). This proves that the sets in (2) are connected, and thus

$$V^\infty = \bigcup_{n \geq 0} J^nV$$

is an open and connected subset of V^*, and $V \subset V^\infty$. We shall prove that V^∞ is closed in V^*, and therefore V^∞ is a connected component of V^*, and hence $V^\infty = W$, since $V \subset V^\infty$.

In order to see that V^∞ is closed in $V^* = GV$, let $y \in \overline{V}^\infty \cap V^*$. Since $y \in V^* = GV$ we have that $y = gx$, for some $g \in G$ and some $x \in V$. Now gV is an open neighborhood of $gx = y$ in M, and hence $gV \cap V^\infty \neq \emptyset$, since $y \in \overline{V}^\infty$. Therefore $gV \cap J^kV \neq \emptyset$, for some $k \geq 0$, and hence $g \in G[J^kV,V] = J^{k+1}$, where the equality is given by (1). Thus $y = gx \in J^{k+1}V \subset V^\infty$. We have shown that $\overline{V}^\infty \cap V^* \subset V^\infty \cap V^*$, and hence V^∞ is closed in V^*. Thus we conclude that $V^\infty = W$.

We can now deduce that W is second countable in the following way. Since $G[V] \subset G[A]$ and $G[A]$ is compact it follows that $J = G[V]$ lies in a subset P_0 of G which is a finite union of connected components of G. It then follows that $J^\infty = \bigcup_{n\geq 0} J^n$ lies in a subset P of G which is the union of at most a countable number of connected components of G. Since each connected component of G is second countable, see Propositions 2.1 and 2.5, it follows that P is second countable.

The product space $P \times V$ is second countable, since both P and V are, and the map $P \times V \to PV$, $(g, x) \mapsto gx$, is an open map onto PV. Consequently PV is second countable. Since $W = V^\infty = \bigcup_{n\geq 0} J^n V = (\bigcup_{n\geq 0} J^n)V = J^\infty V \subset PV$, it now follows that W is second countable. □

Addendum 5.2 *In Theorem 5.1 the paracompact G-invariant open neighborhood V^* of x in M may in be chosen such that the action of G on V^* is proper.*

Proof Let the notation be the same as in the proof of Theorem 5.1. Since the open neighborhood V of x in M is such that $A = \overline{V}$ is compact and $G[A]$ is a compact subset of G it follows, by the proof of Lemma 4.11, that the action of G on $V^* = GV$ is proper. □

In connection with Theorem 5.1 we may add the following two remarks.

Remark 5.3 Suppose that X is a second-countable Cartan G-space. We claim that G can have at most a countable number of connected components, i.e., that G is second-countable. This is seen as follows. Let $x \in X$, then $\alpha : G/G_x \to Gx,\ \ gG_x \mapsto gx$, is a homeomorphism, see Corollary 4.5. Since X is second-countable it follows that Gx is second-countable and hence G/G_x is second-countable. Thus G/G_x has at most a countable number of connected components, and since G_x is compact it follows from this that G has at most a countable number of connected components.

Remark 5.4 Suppose that M is a paracompact Cartan G-manifold, and let M_1 be a connected component of M. Then M_1 is a second-countable manifold. Let us denote $G_1 = \{g \in G \mid gM_1 = M_1\}$. Then G_1 is a subgroup of G and $G_0 \subset G_1$, where G_0 denotes the connected component of G containing the identity element $e \in G$, and moreover G_1 is a union of connected components of G. It follows by Remark 5.3 that G_1 has at most a countable number of connected components. Furthermore the action of G on GM_1 is completely determined by the action of G_1 on M_1, and the set of components of GM_1 is in a natural one-to-one correspondence with G/G_1.

Notes Theorem 5.1 is Proposition 1.3 in [16].

6 Homogeneous spaces of Lie groups

We begin with the following fundamental and well known result.

Theorem 6.1 *Let G be a Lie group and H a closed subgroup of G. Then there exists a unique real analytic structure on G/H, making G/H into a real analytic manifold, for which the natural action $\Phi : G \times G/H \to G/H$, $(g', gH) \mapsto g'gH$, of G on G/H, is real analytic.*

Proof See e.g. [13], Theorem II.4.2. □

We will always consider G/H as a real analytic manifold with the real analytic structure given by Theorem 6.1. The natural projection $\pi : G \to G/H$, $g \mapsto gH$, is then a real analytic map. Furthermore we have the following well known result.

Lemma 6.2 *Let G be a Lie group and H a closed subgroup of G. Then each point $gH \in G/H$ has an open neighborhood U so that there exists a real analytic cross section $\sigma : U \to G$ of the projection $\pi : G \to G/H$, i.e., we have that $\pi \circ \sigma = \mathrm{id}_U$.*

Proof See e.g. [8], §V in Chapter IV, or [13], Lemma II.4.1. □

Lemma 6.3 *Let G be a Lie group, and let K be a compact subgroup of G. Then there exist a K-invariant open neighborhood U of eK in G/K and a real analytic cross section*

$$\sigma : U \to G, \tag{1}$$

for which $\sigma(eK) = e$, and which is a K-equivariant map in the sense that

$$\sigma(ku) = k\sigma(u)k^{-1}, \text{ for every } k \in K \text{ and every } u \in U. \tag{2}$$

Furthermore U can be chosen so that there exists a K-equivariant real analytic isomorphism

$$h : \mathbb{R}^d(\tau) \to U, \tag{3}$$

where $\mathbb{R}^d(\tau)$ denotes an orthogonal representation space for K. In particular the number of K-isotropy types in U is finite.

Proof The compact Lie group K acts on the homogeneous space G/K by the action $K \times G/K \to G/K$, $(k, gK) \mapsto kgK$. Since this action is induced by the natural action of G on G/K, it is real analytic. The point $eK \in G/K$ is a fixed point, and hence we know by Corollary 3.3 that there exist arbitrarily small K-invariant open neighborhoods of eK in G/K.

Let K act on G by conjugation, i.e., by $K \times G \to G$, $(k, g) \mapsto k * g = kgk^{-1}$. The natural projection $\pi : G \to G/K$ is a real analytic submersion, and π is a K-equivariant map, since $\pi(k * g) = \pi(kgk^{-1}) = \pi(kg) = k\pi(g)$, for all $k \in K$ and $g \in G$. The points $e \in G$ and $\pi(e) = eK \in G/K$ are fixed points of K, and hence the tangent spaces $\mathrm{T}_e(G)$ and $\mathrm{T}_{eK}(G/K)$ at $e \in G$ and $eK \in G/K$, respectively, are finite-dimensional linear representation spaces for K and we may assume that they are orthogonal representation spaces since K is compact, see Theorem 3.5. The differential $\mathrm{d}(\pi)_e : \mathrm{T}_e(G) \to \mathrm{T}_{eK}(G/K)$, of π at $e \in G$, is a K-equivariant surjective linear map.

The tangent space $\mathrm{T}_e(K)$ to K at $e \in K$ is a K-invariant linear subspace of $\mathrm{T}_e(G)$. Let L be the orthogonal complement to $\mathrm{T}_e(K)$ in $\mathrm{T}_e(G)$. Then L is a K-invariant linear subspace of $\mathrm{T}_e(G)$ and $\mathrm{T}_e(G) = \mathrm{T}_e(K) \oplus L$. Furthermore $\dim L = d = \dim G - \dim K = \dim G/K$, and the restriction of the differential $\mathrm{d}(\pi)_e| : L \to \mathrm{T}_e(G/K)$ is a K-equivariant linear isomorphism. Using the exponential map at $e \in G$ one constructs an d-dimensional K-invariant real analytic submanifold V^* of G, such that $e \in V^*$ and $\mathrm{T}_e(V^*) = L$. Then $\pi| : V^* \to G/K$ is a K-equivariant real analytic map and the differential $\mathrm{d}(\pi|V^*)_e : \mathrm{T}_e(V^*) \to \mathrm{T}_e(G/K)$ is a K-equivariant linear isomorphism. It now follows, using the real analytic inverse function theorem (see e.g. [42], Theorem 2.2.10), that there exists a K-invariant open neighborhood V of e in V^* such that $\pi| : V \to \pi(V)$ is a K-equivariant real analytic isomorphism onto a K-invariant open neighborhood $U = \pi(V)$ of eK in G/K. Then

$$\sigma = (\pi|V)^{-1} : U \to V \hookrightarrow G$$

is a K-equivariant real analytic cross section over U, of the projection $\pi : G \to G/K$, and $\sigma(eK) = e$.

By using the exponential map at $eK \in G/K$ we obtain arbitrarily small K-invariant open neighborhoods of eK in G/K, which are K-equivariantly and real analytically isomorphic to an open disk, of some small radius, in an orthogonal representation space $\mathbb{R}^d(\tau) \cong \mathrm{T}_{eK}(G/K)$ for K. Here $\tau : K \to \mathrm{O}(d)$ denotes an orthogonal representation of K and $\mathbb{R}^d(\tau)$ denotes the corresponding orthogonal representation space for K. Thus it follows that we can always choose the K-invariant open neighborhood U in (1) such that there exists a K-equivariant real analytic isomorphism $h : \mathbb{R}^d(\tau) \to U$, as in (3). In particular we have in this case that the number of K-isotropy types occurring in U is finite, see Theorem 3.6. □

Notes All results described in Section 6 are well-known. Our exposition follows the one in [16], Section 3.

7 Twisted products

Let G be a Lie group and H a closed subgroup of G. Suppose N is a C^t H-manifold, where $1 \le t \le \omega$. We consider the space $G \times N$ as an H-space, where H acts on $G \times N$ by the (left) action $H \times (G \times N) \to G \times N,\ (h, (g, x)) \mapsto (gh^{-1}, hx)$. The orbit space of this H-space is denoted by $G \times_H N$, and is called the *twisted product, of G and N over H*. Let $\Pi : G \times N \to G \times_H N$ be the natural projection, and denote $\Pi(g, x) = [g, x]$. Thus $[gh, x] = [g, hx]$, for all $h \in H$. There is a canonical action of G on $G \times_H N$ given by $g'[g, x] = [g'g, x]$. We shall here below show that the space $G \times_H N$ can, in a natural way, be given the structure of a C^t manifold, and that the canonical action of G on $G \times_H N$ is a C^t action, $1 \le t \le \omega$.

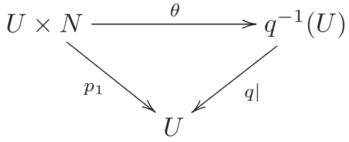
The map

$$q : G \times_H N \to G/H,\ \ [g, x] \mapsto gH, \tag{i}$$

is clearly a well-defined G-equivariant map.

Lemma 7.1 *Let the notation be as above. Then the map $q : G \times_H N \to G/H$, $[g, x] \mapsto gH$, is a locally trivial projection, with fiber N.*

Proof Let $gH \in G/H$. We choose an open neighborhood U of gH in G/H such that there exists a real analytic cross section $\sigma : U \to G$, see Lemma 6.2. Then we have a commutative diagram



(1)

where $\theta(u, x) = [\sigma(u), x]$, and p_1 denotes projection onto the first factor. Furthermore θ is a homeomorphism, with inverse given by $[g, x] \mapsto (\pi(g), \sigma(\pi(g))^{-1}gx)$. Observe that $\sigma(\pi(g))^{-1}g \in H$, since $\pi(\sigma(\pi(g))) = \pi(g)$. Here $\pi : G \to G/H,\ g \mapsto gH$, denotes the natural projection. □

Proposition 7.2 *Let H be a closed subgroup of a Lie group G, and let N be a C^t H-manifold, where $1 \leq t \leq \omega$. Then $G \times_H N$ has a C^t manifold structure, $1 \leq t \leq \omega$. Furthermore the projection map $q : G \times_H N \to G/H$, $[g,x] \mapsto gH$, is a G-equivariant C^t map, and the map $i : N \to G \times_H N$, $x \mapsto [e,x]$, is an H-equivariant closed C^t imbedding, $1 \leq t \leq \omega$.*

Proof Let $\sigma : U \to G$ be a real analytic cross section as in Lemma 7.1. Now suppose $\sigma' : U' \to G$ is another real analytic cross section over an open subset U' of G/H, and let $\theta' : U' \times N \to q^{-1}(U')$, $(u',x) \mapsto [\sigma'(u'),x]$, be the corresponding trivialization over U' of the projection q in (i). Assume $U \cap U' \neq \emptyset$. Then the transition function $\tau_{\theta',\theta} = (\theta'^{-1} \circ \theta)|((U \cap U') \times N)$, from the trivialization θ to the trivilization θ', is given by

$$\tau_{\theta',\theta} : (U \cap U') \times N \to (U \cap U') \times N, \;\; (u,x) \mapsto (u, \sigma'(u)^{-1}\sigma(u)x).$$

Thus we see that the transition functions $\tau_{\theta',\theta}$ are C^t diffeomorphisms, $1 \leq t \leq \omega$. Hence it follows that there exists a C^t structure on $G \times_H N$, making $G \times_H N$ into a C^t manifold, such that each trivialization map $\theta : U \times N \to q^{-1}(U)$, $(u,x) \mapsto [\sigma(u),x]$, as in (1) in Lemma 7.1, is a C^t diffeomorphism from $U \times N$ onto the open subset $q^{-1}(U)$ of $G \times_H N$.

The facts that the projection $q : G \times_H N \to G/H$, $[g,x] \mapsto gH$, is a C^t map, in fact a C^t submersion, and that $i : N \to G \times_H N$, $x \mapsto [e,x]$, is a C^t immersion, $1 \leq t \leq \omega$, are direct consequences of the above description of the C^t structure on $G \times_H N$. Moreover the image $i(N)$ is closed in $G \times_H N$, and it is not difficult to see that $i\colon N \to i(N)$ is a homeomorphism. Thus $i\colon N \to G \times_H N$, $x \mapsto [e,x]$, is a closed C^t imbedding, $1 \leq t \leq \omega$. Clearly q is a G-equivariant map, and i is an H-equivariant map. □

Proposition 7.3 *Let H, G, N, and t be as in Proposition 7.2. Then the canonical action,*

$$\Phi : G \times (G \times_H N) \to G \times_H N, \;\; (g',[g,x]) \mapsto [g'g,x], \tag{1}$$

of G on $G \times_H N$ is a C^t action, $1 \leq t \leq \omega$.

Proof Let $(g_0',[g_0,x_0]) \in G \times (G \times_H N)$, then $q([g_0,x_0]) = g_0H \in G/H$. Now recall that the action $G \times G/H \to G/H$, $(g',gH) \mapsto g'gH$, of G on G/H, is real analytic, see Theorem 6.1. Let U_1 be an open neighborhood of $g_0'g_0H$ in G/H such that there exists a real analytic cross section $\sigma_1 : U_1 \to G$. Next we choose an open neighborhood V of g_0' in G and an open neighborhood U of g_0H in G/H such that $VU \subset U_1$, and such that there exists a real analytic cross section $\sigma : U \to G$. Now $V \times q^{-1}(U)$ is an open neighborhood of $(g_0',[g_0,x_0])$ in $G \times (G \times_H N)$ and $\Phi(V \times q^{-1}(U)) \subset q^{-1}(U_1)$. Furthermore we have the commutative diagram

$$\begin{array}{ccc} V \times q^{-1}(U) & \xrightarrow{\Phi|} & q^{-1}(U_1) \\ {\scriptstyle \mathrm{id}\times\theta}\uparrow\cong & & \cong\uparrow{\scriptstyle \theta_1} \\ V \times U \times N & \xrightarrow{\Psi} & U_1 \times N \end{array} \tag{2}$$

Here the trivializations $\theta : U \times N \to q^{-1}(U)$, $(u,x) \mapsto [\sigma(u),x]$, and $\theta_1 : U_1 \times N \to q^{-1}(U_1)$, $(u,x) \mapsto [\sigma_1(u),x]$, corresponding to the cross sections σ and σ_1, respectively, are C^t diffeomorphisms, $1 \leq t \leq \omega$. The map Ψ is given by

$$\Psi(g,u,x) = (gu, \sigma_1(gu)^{-1}g\sigma(u)x), \;\; \text{for all } \; (g,u,x) \in V \times U \times N,$$

and thus we see that Ψ is a C^t map. Hence it follows that $\Phi|$ in (2) is a C^t map, and thus we have proved that Φ in (1) is a C^t map, $1 \le t \le \omega$. □

Lemma 7.4 *Let G be a Lie group and H a closed subgroup of G. Suppose M is a C^t G-manifold and N is a C^t H-manifold, where $1 \le t \le \omega$. Let $f : N \to M$ be an H-equivariant C^t map. Then*

$$\mu(f) : G \times_H N \to M, \ \ [g, x] \mapsto gf(x),$$

is a G-equivariant C^t map, $1 \le t \le \omega$.

Proof It is easy to see that $\mu(f)$ is a well-defined G-equivariant map. Now let $[g_0, x_0] \in G \times_H N$, and choose an open neighborhood U of $q([g_0, x_0]) = g_0H$ in G/H, such that there is a real analytic cross section $\sigma : U \to G$. Then $q^{-1}(U)$ is an open neighborhood of $[g_0, x_0]$ in $G \times_H N$ and $\theta : U \times N \to q^{-1}(U), \ (u, x) \mapsto [\sigma(u), x]$, is a C^t diffeomorphism, $1 \le t \le \omega$. The composite map $(\mu(f)|) \circ \theta : U \times N \to M$ is given by $(u, x) \mapsto \sigma(u)f(x)$, and hence it is a C^t map. Since θ is a C^t diffeomorphism it follows that $\mu(f)| : q^{-1}(U) \to M$ is a C^t map, $1 \le t \le \omega$, and this proves our claim. □

Corollary 7.5 *Let H be a closed subgroup of a Lie group G. Suppose that N and P are C^t H-manifolds, and let $f : N \to P$ be a H-equivariant C^t map, where $1 \le t \le \omega$. Then*

$$\mathrm{id} \times_H f : G \times_H N \to G \times_H P\ , \ [g, x] \mapsto [g, f(x)]$$

is a G-equivariant C^t map, $1 \le t \le \omega$.

Proof By Propositions 7.2 and 7.3 the twisted products $G \times_H N$ and $G \times_H P$ are C^t G-manifolds, $1 \le t \le \omega$. The H-equivariant map $i : P \to G \times_H P\ , \ y \mapsto [e, y]$, is a closed C^t imbedding, $1 \le t \le \omega$, by Proposition 7.2. Thus $i \circ f : N \to G \times_H P\ , \ x \mapsto [e, f(x)]$, is an H-equivariant C^t map. By Lemma 7.4 the map $\mu(i \circ f) : G \times_H N \to G \times_H P$ is a G-equivariant C^t map. Now $\mu(i \circ f) = \mathrm{id} \times_H f$, since $\mu(i \circ f)([g, x]) = g(i \circ f)(x) = g[e, f(x)] = [g, f(x)] = (\mathrm{id} \times_H f)([g, x])$, and this completes the proof. □

Corollary 7.6 *Let the notation be as in Corollary 7.5 above, and suppose that $f : N \to P$ is an H-equivariant C^t diffeomorphism, where $1 \le t \le \omega$. Then*

$$\mathrm{id} \times_H f : G \times_H N \to G \times_H P\ , \ [g, x] \mapsto [g, f(x)],$$

is a G-equivariant C^t diffeomorphism, $1 \le t \le \omega$. □

Remark 7.7 Suppose that K is a *compact* subgroup of G. By Lemma 6.3 there exist a K-invariant open neighborhood U of eK in G/K and a real analytic cross section $\sigma : U \to G$, with $\sigma(eK) = e$, which is a K-equivariant map in the sense that $\sigma(ku) = k\sigma(u)k^{-1}$, for every $k \in K$ and every $u \in U$. In this case the corresponding trivialization

$$\theta : U \times N \to q^{-1}(U), \ \ (u, x) \mapsto [\sigma(u), x], \tag{1}$$

is in addition a K-equivariant map. Here K acts diagonally on $U \times N$, and the fact that θ is a K-equivariant map is seen by $\theta(k(u, x)) = \theta(ku, kx) = [\sigma(ku), kx] = [k\sigma(u)k^{-1}, kx] = [k\sigma(u), x] = k[\sigma(u), x] = k\theta(u, x)$. □

Notes Everything presented here in Section 7 is well-known. Our exposition mainly follows the one in [16], Section 4.

8 Slices

Definition 8.1 Let M be a C^t G-manifold, where G is a Lie group and $1 \leq t \leq \omega$, and let H be a closed subgroup of G. We say that an H-invariant C^t submanifold S of M is a C^t *H-slice in M*, if GS is open in M, and the map

$$\mu : G \times_H S \to GS, \ \ [g, x] \to gx,$$

is a G-equivariant C^t diffeomorphism, $1 \leq t \leq \omega$. We call GS *the tube corresponding to the H-slice S*.

By a C^t *presentation of* a C^t H-slice S in M, we mean a C^t H-manifold P together with an H-equivariant C^t imbedding, $1 \leq t \leq \omega$,

$$j : P \to M, \tag{i}$$

such that $j(P) = S$. Then $j : P \to S$ is an H-equivariant C^t-diffeomorphism, and hence $(\mathrm{id} \times_H j)\colon G \times_H P \to G \times_H S$ is a G-equivariant C^t diffeomorphism by Corollary 7.6. Thus

$$\mu(j) = \mu \circ (id \times_H j) : G \times_H P \to GS$$

is a G-equivariant C^t diffeomorphism, $1 \leq t \leq \omega$.

Suppose that $z \in M$. By a C^t *slice at z in M* we mean a C^t G_z-slice S in M, such that $z \in S$. In this case GS is a G-invariant open neighborhood of the orbit Gz in M, and GS is called a *G-tube about Gz*. A C^t slice S at z in M is said to be *linear* if S has a C^t presentation of the form

$$j : \mathbb{R}^q(\rho) \xrightarrow{\cong} S \subset M, \tag{ii}$$

where $j(0) = z$. Here $\mathbb{R}^q(\rho)$ denotes a linear representation space for H.

Lemma 8.2 *Let M, G, H and t be as in Definition 8.1, and suppose S is a C^t H-slice in M. Then there exists a G-equivariant C^t map $p\colon GS \to G/H$ such that $p^{-1}(eH) = S$.*

Proof By assumption GS is open in M, and $\mu\colon G \times_H S \to GS$, $[g, x] \mapsto gx$, is a G-equivariant C^t diffeomorphism, $1 \leq t \leq \omega$. Furthermore the natural projection $q\colon G \times_H S \to G/H$, $[g, x] \mapsto gH$, is a G-equivariant C^t map, see Proposition 7.2. Hence the composite map $p = q \circ \mu^{-1}\colon GS \to G/H$, is a G-equivariant C^t map, and $p(gx) = gH$, for every $x \in S$ and every $g \in G$. Clearly $p^{-1}(eH) = S$. □

Definition 8.3 Let M, G, H and t be as in Definition 8.1. An H-invariant C^t submanifold S of M is said to be a C^t *near H-slice in M*, if there exist an open neighborhood U of eH in G/H and a real analytic cross section $\sigma : U \to G$, with $\sigma(eH) = e$, such that $\sigma(U)S$ is open in M and

$$\gamma : U \times S \to \sigma(U)S, \quad (u, x) \mapsto \sigma(u)x,$$

is a C^t diffeomorphism, $1 \leq t \leq \omega$.

The notions, *a presentation of a C^t near H-slice, a C^t near slice at $z \in M$, and a linear C^t near slice at $z \in M$*, are defined in complete analogy with the corresponding notions for C^t slices, $1 \leq t \leq \omega$. For example the assertion that S is a linear C^t near slice at $z \in M$, means that S is a C^t near G_z-slice in M, such that $z \in S$, and there exists a G_z-equivariant C^t diffeomorphism $j\colon \mathbb{R}^q(\rho) \to S$, as in (ii). In this case

$$\gamma(j) = \gamma \circ (\mathrm{id} \times j)\colon U \times \mathbb{R}^q(\rho) \to \sigma(U)S\ ,\ (u,a) \mapsto \sigma(u)j(a),$$

is a C^t diffeomorphism, $1 \leq t \leq \omega$.

Remark 8.4 Let the notation be as above, and suppose that S is a C^t near H-slice in M. Then the set GS is open in M, since $GS = G\sigma(U)S = \bigcup_{g\in G} g\sigma(U)S$ is a union of open subsets of M. □

Remark 8.5 Let the notation be as in Definition 8.3. Suppose S is a C^t near H-slice in M, and let $g_0H \in G/H$. Then we can find an open neighborhood U_0 of g_0H in G/H and a real analytic cross section $\sigma_0 : U_0 \to G$, such that $\sigma_0(U_0)S$ is open in M and

$$\gamma_0 : U_0 \times S \to \sigma_0(U_0)S,\ (u',x) \mapsto \sigma_0(u')x,$$

is a C^t diffeomorphism, $1 \leq t \leq \omega$. We choose $U_0 = g_0U$ and let $\sigma_0\colon U_0 \to G$, $u' \mapsto g_0\sigma(g_0^{-1}u')$. Then $\gamma_0 = g_0 \circ \gamma \circ (g_0^{-1} \times \mathrm{id})\colon U_0 \times S \to \sigma_0(U_0)S = g_0\sigma(U)S$, is as required.

Remark 8.6 Suppose $\sigma\colon U \to G$ and $\sigma'\colon U \to G$ are two real analytic cross sections over an open subset U of G/H, and let $\gamma\colon U \times S \to \sigma(U)S\ ,\ (u,x) \mapsto \sigma(u)x$, and $\gamma'\colon U \times S \to \sigma'(U)S,\ (u,x) \mapsto \sigma'(u)x$. Then $\kappa(u) = \sigma(u)^{-1}\sigma'(u) \in H$, for every $u \in U$, and $\kappa\colon U \to H$ is real analytic. Thus $\sigma'(U)S = (\sigma \cdot \kappa)(U)S = \sigma(U)\kappa(U)S = \sigma(U)S$, and furthermore $\gamma' = \gamma \circ \kappa^*$, where $\kappa^*\colon U \times S \to U \times S,\ (u,x) \mapsto (u, \kappa(u)x)$. Hence $\sigma'(U)S$ is open in M if and only if $\sigma(U)S$ is, and γ' is a C^t diffeomorphism if and only if γ is, $1 \leq t \leq \omega$. In particular the definition of a C^t near slice in Definition 8.3 is independent of the choice of the real analytic section $\sigma\colon U \to G$.

Lemma 8.7 *Let M, G, H and t be as in Definition 8.1, and let S be a C^t H-slice in M. Suppose $\sigma : U \to G$ is a real analytic cross section over an open subset U of G/H. Then $\sigma(U)S$ is open in M and*

$$\gamma : U \times S \to \sigma(U)S, \quad (u,x) \mapsto \sigma(u)x, \tag{1}$$

is a C^t diffeomorphism. Thus every C^t H-slice S in M, is a C^t near H-slice in M.

Proof As we saw in Section 7 (see Lemma 7.1 and Proposition 7.2) the cross section σ gives rise to a C^t diffeomorphism $\theta : U \times S \to q^{-1}(U),\ (u,x) \mapsto [\sigma(u), x]$, where $q : G \times_H S \to G/H\ ,\ [g,x] \mapsto gH$. Let μ be as in Definition 8.1. Then $(\mu \circ \theta)(u,x) = \mu([\sigma(u),x]) = \sigma(u)x = \gamma(u,x)$, for all $(u,x) \in U \times S$. Hence $\sigma(U)S = \gamma(U \times S) = \mu(q^{-1}(U))$, and $\gamma = \mu \circ \theta\colon U \times S \to \sigma(U)S$. Since $\mu\colon G \times_H S \to GS\ ,\ [g,x] \mapsto gx$, is a C^t diffeomorphism it follows that $\sigma(U)S$ is open in GS, and hence also open in M, and that $\gamma = \mu \circ \theta\colon U \times S \to \sigma(U)S$ is a C^t diffeomorphism. □

Remark 8.8 Let M be a C^t G-manifold, where G is a Lie group and $1 \leq t \leq \omega$. Suppose S is a C^t near K-slice in M, where K is a compact subgroup of G, and let $j\colon P \to S \hookrightarrow M$ be a C^t presentation of S, i.e., P is a C^t K-manifold and $j\colon P \to M$ is a K-equivariant C^t imbedding with $j(P) = S$. In this case, when K is compact, we may in Definition 8.3 take the open neighborhood U, of eK in G/K, to be K-invariant, and the real analytic cross section $\sigma\colon U \to G$, with $\sigma(eK) = e$, to be a K-equivariant map, see Corollary 3.3, Lemma 6.3 and Remark 8.6. We then have that the C^t diffeomorphism

$$\gamma(j) : U \times P \to \sigma(U)S,\ (u,x) \mapsto \sigma(u)j(x),$$

is a K-equivariant map, where K acts diagonally on $U \times P$. This holds since $\gamma(j)(k(u,x)) = \gamma(j)(ku,kx) = \sigma(ku)j(kx) = k\sigma(u)k^{-1}kj(x) = k\sigma(u)j(x) = k\gamma(j)(u,x)$, for every $k \in K$ and every $(u,x) \in U \times S$. Thus $W = \sigma(U)S$ is a K-invariant open neighborhood of S in M, and W is called a *K-invariant product neighborhood of S in M*. We call the K-equivariant C^t diffeomorphism $\gamma(j) : U \times P \to \sigma(U)S = W$, a *$K$-equivariant C^t presentation of W*, $1 \le t \le \omega$.

Remark 8.9 Let the notation be as in Remark 8.8. Suppose S is a linear C^t near slice at $z \in M$, where $G_z = K$ is compact, and let $j : \mathbb{R}^q(\rho) \to S$ be a C^t presentation of S. We may choose the K-invariant open neighborhood U of eK in G/K, such that there is a K-equivariant real analytic isomorphism $h : \mathbb{R}^d(\tau) \to U$, where $\mathbb{R}^d(\tau)$ denotes an orthogonal representation space for K, see Lemma 6.3. Then

$$\hat{\gamma}(j) = \gamma(j) \circ (h \times \mathrm{id}) : \mathbb{R}^d(\tau) \times \mathbb{R}^q(\rho) \to W, \quad (b,a) \mapsto \sigma(h(b))j(a), \tag{1}$$

is a K-equivariant C^t diffeomorphism. Since $\mathbb{R}^d(\tau) \times \mathbb{R}^q(\rho) = \mathbb{R}^{d+q}(\tau \oplus \rho)$, and the number of K-isotropy types occurring in a linear representation space for K is finite, see Theorem 3.6, it follows that the number of K-isotropy types occurring in W is finite.

Lemma 8.10 *Let M be a C^t G-manifold, where G is a Lie group and $1 \le t \le \omega$, and let H be a closed subgroup of G. Then the following assertions are equivalent:*

(a) *S is a C^t H-slice in M.*

(b) *S is a C^t near H-slice in M and $gS \cap S = \emptyset$, for all $g \in G - H$.*

(c) *S is a C^t near H-slice in M and there exists a G-equivariant map $p\colon GS \to G/H$ such that $p^{-1}(eH) = S$.*

Proof Suppose that (a) holds. Then it follows directly by Definition 8.1 that $gS \cap S = \emptyset$, for all $g \in G - H$. Furthermore we have by Lemma 8.7 that S is a C^t near H-slice in M. Thus (a) implies (b).

Now assume that (b) holds. We shall show that (a) holds. By Remark 8.4 the set GS is open in M, and our claim is that

$$\mu : G \times_H S \to GS,\ [g,x] \mapsto gx \tag{1}$$

is a G-equivariant C^t diffeomorphism, $1 \le t \le \omega$. We know by Lemma 7.4 that μ is a G-equivariant C^t map. The fact that $gS \cap S = \emptyset$, for all $g \in G - H$, implies that μ is injective, and hence μ in (1) is a bijective map. Therefore it is enough to prove that μ is a local C^t diffeomorphism. Let $[g_0, x_0] \in G \times_H S$, then $q([g_0, x_0]) = g_0H \in G/H$, where $q\colon G \times_H S \to G/H\ ,\ [g,x] \mapsto gH$. By Remark 8.5 we find an open neighborhood U_0 of g_0H in G/H and a real analytic cross section $\sigma_0 : U_0 \to G$ such that $\sigma_0(U_0)S$ is open in M and

$$\gamma_0 : U_0 \times S \to \sigma_0(U_0)S,\ (u,x) \mapsto \sigma_0(u)x,$$

is a C^t diffeomorphism. Now $q^{-1}(U_0)$ is an open neighborhood of $[g_0, x_0]$ in $G \times_H S$ and

$$\theta_0 : U_0 \times S \to q^{-1}(U_0),\ (u, x) \mapsto [\sigma_0(u), x],$$

is a C^t diffeomorphism, see Lemma 7.1 and Proposition 7.2. Since γ_0 and θ_0 are C^t diffeomorphisms and $\gamma_0 = (\mu|) \circ \theta_0$ it follows that $\mu| : q^{-1}(U_0) \to \sigma_0(U_0)S$ is a C^t diffeomorphism, $1 \leq t \leq \omega$. Thus (b) implies (a).

The fact that (a) implies (c) follows by Lemmas 8.2 and 8.7. Clearly (c) implies (b). This completes the proof. □

Proposition 8.11 *Let M be a C^t Cartan G-manifold, where G is a Lie group and $1 \leq t \leq \omega$, and let $z \in M$. Suppose S^* is a C^t near slice at z. Then there exists a C^t slice S at z. Moreover we can choose S to be a G_z-invariant open neighborhood of z in S^*. If S^* is a linear C^t near slice at z, we can choose S to be a linear C^t slice at $z \in M$.*

Proof The isotropy subgroup $G_z = K$ at $z \in M$ is a compact subgroup of G. We let U be an open neighborhood of eK in G/K for which there is a real analytic cross section $\sigma\colon U \to G$, with $\sigma(eK) = e$, such that $\sigma(U)S^*$ is open in M and $\gamma\colon U \times S^* \to \sigma(U)S^*$, $(u, x) \mapsto \sigma(u)x$, is a C^t diffeomorphism, $1 \leq t \leq \omega$. Let $\pi : G \to G/K$ denote the natural projection. By Lemma 4.6 there exists a K-invariant open neighborhood V of z in M, such that $G[V] \subset \pi^{-1}(U)$. We define

$$S = V \cap S^*, \tag{1}$$

and claim that S is a C^t K-slice at z in M.

Since S is a K-invariant open subset of S^* it follows directly by Definition 8.3 that S is a C^t near K-slice at z in M. Hence it is enough, by Lemma 8.10, to show that $gS \cap S = \emptyset$, for all $g \in G - K$. Suppose that $gS \cap S \neq \emptyset$, where $g \in G$, and let $x_1, x_2 \in S$ be such that $gx_1 = x_2$. Then $g \in G[S] \subset G[V] \subset \pi^{-1}(U)$, and hence $\pi(g) \in U$. Thus $g = \sigma(\pi(g))k$, for some $k \in K$. Now $\gamma(\pi(g), kx_1) = \sigma(\pi(g))(kx_1) = gk^{-1}kx_1 = gx_1 = x_2$ and $\gamma(eK, x_2) = \sigma(eK)x_2 = ex_2 = x_2$. Since γ is an injective map it follows that $\pi(g) = eK$, and hence $g \in K$.

Suppose S^* is a linear C^t near slice at $z \in M$, and let $j\colon \mathbb{R}^q(\rho) \to S^*$ be a presentation of S^*, where $j(0) = z$ and $\mathbb{R}^q(\rho)$ is an orthogonal representation for $K = G_z$. Let S denote the C^t slice at $z \in M$ constructed above, i.e., S is as in (1). Then $j^{-1}(S)$ is a K-invariant open neighborhood of the origin in $\mathbb{R}^q(\rho)$. Hence there exists $\varepsilon > 0$, such that $\mathring{D}^q_\varepsilon(\rho) \subset j^{-1}(S)$. Then $S' = j(\mathring{D}^q_\varepsilon(\rho))$ is a linear C^t slice at $z \in M$. Moreover S' is a K-invariant open neighborhood of z in S^*. □

Proposition 8.12 *Let M be a paracompact C^s Cartan G-manifold, where G is a Lie group and $1 \leq s \leq \infty$, and let $x \in M$. Then there exists a linear C^s near slice S at x in M.*

Proof See [29], p. 139. Compare also with [45], Lemma in Section 2.2 and Proposition 2.2.1. □

Theorem 8.13 (The C^s slice theorem) *Let M be a C^s Cartan G-manifold, where G is a Lie group and $1 \leq s \leq \infty$, and let $x \in M$. Then there exists a linear C^t slice S at x in M.*

Proof By Theorem 5.1 there exists a paracompact G-invariant open neighborhood V^* of x in M. Thus V^* is a paracompact C^s Cartan G-manifold, and $x \in V^*$. By Proposition 8.12 there exists a linear C^s near slice S^* at x in V^*. Since V^* is an open and G-invariant subset of M it follows that S^* is a linear C^s near slice at x in M. Hence we have by Proposition 8.11 that there exists a linear C^s slice at x in M. □

Notes The exposition here in Section 8 follows the one in [16], Section 5. In [45], Palais uses the assertion in Lemma 8.10 (c) as the definition of a C^t H-slice in M. Thus Lemma 8.10 shows that the definition by Palais is equivalent to the one we use, i.e., the one in Definition 8.1. The definition of a near slice at a point x, is in the topological case, given in Definition 2.1.6 in [45]. Proposition 8.11 corresponds to Proposition 2.1.7 in [45].

9 The strong C^r topologies, $1 \le r < \infty$, and the very-strong C^∞ topology

Suppose U is an open subset of $\mathbb{R}^m$, or of $\mathbb{R}^m_\# = \{x = (x_1, \ldots, x_m) \in \mathbb{R}^m \mid x_m \ge 0\}$, and let $f\colon U \to \mathbb{R}^n$ be a C^r map, where $1 \le r < \infty$, and let $A \subset U$. Then we define

$$\|f\|^r_A = \sup\left\{|\mathrm{D}^\alpha f_j(a)| \;\middle|\; a \in A,\;\; 1 \le j \le n,\; 0 \le |\alpha| \le r\right\}.$$

Here $f_j : U \to \mathbb{R}$ denotes the j:th component function of f, $1 \le j \le n$, and $\alpha = (\alpha_1, \ldots, \alpha_m)$ is an m-tuple of non-negative integers, and

$$\mathrm{D}^\alpha f_j(a) = \frac{\partial^{|\alpha|} f_j(a)}{\partial x_1^{\alpha_1} \ldots \partial x_m^{\alpha_m}}, \quad \text{for each } a \in A,$$

where $|\alpha| = \alpha_1 + \ldots + \alpha_m$.

Now suppose that M and N are C^r manifolds, with or without boundary, where $1 \le r < \infty$. By $\mathrm{C}^r(M, N)$ we denote the set of all C^r maps from M to N, $1 \le r < \infty$. Suppose that $f \in \mathrm{C}^r(M, N)$, and let (U, φ) be a chart in M, B a compact subset of U, and (V, ψ) a chart in N, such that $f(B) \subset V$. We then define, for each $\varepsilon > 0$,

$$\begin{aligned} &\mathcal{N}^r(f; B, (U, \varphi), (V, \psi), \varepsilon) \\ &\;= \{h \in \mathrm{C}^r(M, N) \mid h(B) \subset V \text{ and } \|\psi \circ h \circ \varphi^{-1} - \psi \circ f \circ \varphi^{-1}\|^r_{\varphi(B)} < \varepsilon\}. \end{aligned} \tag{i}$$

Note that here $\psi \circ h \circ \varphi^{-1} - \psi \circ f \circ \varphi^{-1}\colon \varphi(f^{-1}(V) \cap h^{-1}(V) \cap U) \to \mathbb{R}^n$ is a C^r map, and that $\varphi(B) \subset \varphi(f^{-1}(V) \cap h^{-1}(V) \cap U)$. We call a set $\mathcal{N}^r(f; B, (U, \varphi), (V, \psi), \varepsilon)$ as in (i) an *elementary* C^r *neighborhood of* f *in* $\mathrm{C}^r(M, N)$, $1 \le r < \infty$.

In the case when $N = \mathbb{R}^n$, and the chart (V, ψ) equals $(\mathbb{R}^n, \mathrm{id})$, we instead of the full notation $\mathcal{N}^r(f; B, (U, \varphi), (\mathbb{R}^n, \mathrm{id}), \varepsilon)$ use the simpler notation $\mathcal{N}^r(f; B, (U, \varphi), \varepsilon)$. If furthermore U is an open subset of $\mathbb{R}^m$, or of the halfspace $\mathbb{R}^m_\#$, and the chart (U, φ) equals (U, id), we denote $\mathcal{N}^r(f; B, (U, \mathrm{id}), \varepsilon)$ by $\mathcal{N}^r(f; B, \varepsilon)$. Thus

$$\mathcal{N}^r(f; B, \varepsilon) = \{h \in \mathrm{C}^r(U, \mathbb{R}^n) \mid \|h - f\|^r_B < \varepsilon\},$$

where $B \subset U \subset \mathbb{R}^m$, and B is compact.

In this article $\mathcal{N}^r(f; B, (U, \varphi), (V, \psi), \varepsilon)$ *will always denote a set of the form in* (i).

Let $f \in \mathrm{C}^r(M, N)$, where $1 \le r < \infty$, and M and N are paracompact C^r manifolds, with or without boundary. By a *basic strong* C^r *neighborhood of* f *in* $\mathrm{C}^r(M, N)$, we mean a set of the form

$$\mathcal{U}^r = \bigcap_{i \in \Lambda} \mathcal{N}^r(f; B_i, (U_i, \varphi_i), (V_i, \psi_i), \varepsilon_i), \tag{ii}$$

where the family $\{B_i\}_{i\in\Lambda}$ is locally finite in M. It is easy to see, compare with Lemma 9.4, that the family of all basic strong C^r neighborhoods of f in $\mathrm{C}^r(M,N)$, for all $f \in \mathrm{C}^r(M,N)$, is a basis for a topology on $\mathrm{C}^r(M,N)$. This topology is called *the strong* C^r *topology on* $\mathrm{C}^r(M,N)$. Let us *temporarily* denote $\mathrm{C}^r(M,N)$ with the strong C^r topology by $\mathrm{C}^r_{\mathrm{S}}(M,N)$. In the definition of the strong C^r topology some authors require that the family $\{U_i\}_{i\in\Lambda}$ is locally finite, see e.g. [15], Section 2.1. However, this definition and the one given above are equivalent, see Lemma 1.1 in [22].

Now let $1 \le s \le \infty$, and assume that $1 \le r < \infty$ is such that $r \le s$. Suppose M and N are paracompact C^s manifolds, with or without boundary, and let $f \in \mathrm{C}^s(M,N)$. Then we define

$$\mathcal{N}^{s,r}(f; B, (U,\varphi), (V,\psi), \varepsilon) = \mathrm{C}^s(M,N) \cap \mathcal{N}^r(f; B, (U,\varphi), (V,\psi), \varepsilon). \tag{iii}$$

Occasionally we will use the shorter notation $\mathcal{N}^{s,r} = \mathrm{C}^s(M,N) \cap \mathcal{N}^r$ instead of the complete form in (iii). We call a set $\mathcal{N}^{s,r}(f; B, (U,\varphi), (V,\psi), \varepsilon)$ as in (iii) an *elementary* C^r *neighborhood of f in* $\mathrm{C}^s(M,N)$. Then we define a *basic strong* C^r *neighborhood of f in* $\mathrm{C}^s(M,N)$ to be a set of the form

$$\mathcal{U}^{s,r} = \bigcap_{i\in\Lambda} \mathcal{N}^{s,r}(f; B_i, (U_i,\varphi_i), (V,\psi_i), \varepsilon_i), \tag{iv}$$

where the family $\{B_i\}_{i\in\Lambda}$ is locally finite in M. We define *the strong* C^r *topology on* $\mathrm{C}^s(M,N)$ to be the topology which as a basis has the family of all sets of the form (iv), for all $f \in \mathrm{C}^s(M,N)$. We shall temporarily denote the set $\mathrm{C}^s(M,N)$ with the strong C^r topology by $\mathrm{C}^{s,r}_{\mathrm{S}}(M,N)$, where $1 \le s \le \infty$, $1 \le r < \infty$ and $r \le s$.

Since $\mathcal{N}^{s,r} = \mathrm{C}^s(M,N) \cap \mathcal{N}^r$, by the definition in (iii), we also have that $\mathcal{U}^{s,r} = \mathrm{C}^s(M,N) \cap \mathcal{U}^r$, where $\mathcal{U}^{s,r}$ is as in (iv) and $\mathcal{U}^r$ is as in (ii). Thus we see that the strong C^r topology on $\mathrm{C}^s(M,N)$ is nothing but the relative topology that $\mathrm{C}^s(M,N)$ obtains as a subset of the space $\mathrm{C}^r_{\mathrm{S}}(M,N)$. Thus we have, as topological spaces,

$$\mathrm{C}^{s,r}_{\mathrm{S}}(M,N) = \mathrm{C}^s(M,N) \cap \mathrm{C}^r_{\mathrm{S}}(M,N).$$

Suppose that M and N are paracompact C^∞ manifolds, with or without boundary. *Let us here, for comparison only, mention a topology on* $\mathrm{C}^\infty(M,N)$ *that we will not use, it is inadequate for our purposes.* It is *the strong* C^∞ *topology on* $\mathrm{C}^\infty(M,N)$, introduced by Mather in [33], Section 2. The strong C^∞ topology on $\mathrm{C}^\infty(M,N)$ has as a basis the union of all strong C^r topologies on $\mathrm{C}^\infty(M,N)$, $1 \le r < \infty$. See also [15], Section 2.1. We use the notation $\mathrm{C}^\infty_{\mathrm{S}}(M,N)$ for $\mathrm{C}^\infty(M,N)$ with the strong C^∞ topology.

Remark 9.1 Let $1 \le r < s \le \infty$, and suppose that M and N are paracompact C^s manifolds, with or without boundary, and let $f \in \mathrm{C}^s(M,N)$. When we defined an elementary C^r neighborhood $\mathcal{N}^{s,r}(f; B, (U,\varphi), (V,\psi), \varepsilon)$ of f is $\mathrm{C}^s(M,N)$, see (iii) and (i), we considered M and N as C^r manifolds and used C^r charts (U,φ) and (V,ψ) in M and N, respectively, $1 \le r < s \le \infty$. Since M and N are C^s manifolds it would also be natural to instead only use charts $(\tilde{U},\tilde{\varphi})$ and $(\tilde{V},\tilde{\psi})$ in M and N, respectively, that are C^s charts. The version of an elementary C^r neighborhood of f in $\mathrm{C}^s(M,N)$ that one obtains in this way we could denote by

$$\mathcal{N}^{s,r}_s(f; B, (\tilde{U},\tilde{\varphi}), (\tilde{V},\tilde{\psi}), \varepsilon)$$

and use the notation $\mathcal{U}_s^{s,r}$ to denote a corresponding form of a basic strong C^r neighborhood of f in $\mathrm{C}^s(M,N)$. It is not difficult to see that the topology on $\mathrm{C}^s(M,N)$, which has all sets of the form $\mathcal{U}_s^{s,r}$ as a basis, equals the topology on $\mathrm{C}^s(M,N)$, which has all sets of the form $\mathcal{U}^{s,r}$, given in (iv), as a basis, i.e., equals the strong C^r topology on $\mathrm{C}^s(M,N)$, $1 \leq r < s \leq \infty$.

We now note that

$$\mathcal{N}_s^{s,r+1}(f;B,(\tilde{U},\tilde{\varphi}),(\tilde{V},\tilde{\psi}),\varepsilon) \subset \mathcal{N}_s^{s,r}(f;B,(\tilde{U},\tilde{\varphi}),(\tilde{V},\tilde{\psi}),\varepsilon),$$

for $1 \leq r < s \leq \infty$. Therefore a set of the form $\mathcal{U}_s^{s,r}$ is open in the corresponding $\mathcal{U}_s^{s,r+1}$, and hence it follows by the above that the identity map

$$\mathrm{id}\colon \mathrm{C}_{\mathrm{S}}^{s,r+1}(M,N) \to \mathrm{C}_{\mathrm{S}}^{s,r}(M,N)$$

is continuous. Starting with $s = r+1 < \infty$ we obtain that the identity map on the set $\mathrm{C}^s(M,N)$ gives us a sequence of continuous maps

$$\mathrm{C}_{\mathrm{S}}^{s}(M,N) = \mathrm{C}_{\mathrm{S}}^{s,s}(M,N) \to \mathrm{C}_{\mathrm{S}}^{s,s-1}(M,N) \to \ldots \to \mathrm{C}_{\mathrm{S}}^{s,1}(M,N).$$

For $s = \infty$ we obtain an infinite sequence of continuous maps

$$\ldots \to \mathrm{C}_{\mathrm{S}}^{\infty,r+1}(M,N) \to \mathrm{C}_{\mathrm{S}}^{\infty,r}(M,N) \to \mathrm{C}_{\mathrm{S}}^{\infty,r-1}(M,N) \to \ldots \to \mathrm{C}_{\mathrm{S}}^{\infty,1}(M,N)$$

and

$$\mathrm{C}_{\mathrm{S}}^{\infty}(M,N) \to \mathrm{C}_{\mathrm{S}}^{\infty,r}(M,N), \quad \text{for all } 1 \leq r < \infty.$$

In the case when r is finite, and M and N are paracompact C^r manifolds, with or without boundary, it is clear that the strong C^r topology is the right topology to use on $\mathrm{C}^r(M,N)$.

However, one should note that, in the case when M and N are paracompact C^∞ manifolds, with or without boundary, the strong C^∞ topology on $\mathrm{C}^\infty(M,N)$ has some serious drawbacks. For example a key lemma, the glueing lemma, that is Lemma 12.1 is this article, does not hold for $s = \infty$ if one uses the strong C^∞ topology, but it holds if one uses the very-strong C^∞ topology defined below. Cf. Lemma 2.2.8 in [15], which holds and is given only in the C^r case, where $1 \leq r < \infty$, since the topology used in [15], in the C^∞ case, is the strong C^∞ topology. Mather [33] calls the strong C^∞ topology the Whitney C^∞ topology, but this choice of terminology is not well founded. In fact du Plessis and Wall, see [48], p. 59, propose that the strong C^∞ topology on $\mathrm{C}^\infty(M,N)$ be named the Mather topology. It is only in the case when r is finite that the strong C^r topology should be named the Whitney C^r topology. One should also note that the strong C^∞ topology on $\mathrm{C}^\infty(M,N)$ is not really a genuine C^∞ topology, since it is completely determined by the strong C^r topologies on $\mathrm{C}^\infty(M,N)$, for all finite r.

There is however a genuine C^∞ topology on $\mathrm{C}^\infty(M,N)$, namely the *very-strong* C^∞ *topology*. This topology was introduced by Cerf in [6], Definition I.4.3.1. We give the definition of the very-strong C^∞ on $\mathrm{C}^\infty(M,N)$, in Definition 9.2 below, in a slightly different way than Cerf does. Cerf does not give this topology any special name, but it is named the 'very strong topology' by du Plessis and Wall, see [48], p. 59. We will temporarily denote $\mathrm{C}^\infty(M,N)$ with the very-strong C^∞ topology by $\mathrm{C}_{\mathrm{vS}}^{\infty}(M,N)$.

The very-strong C^∞ topology is the right topology to use on $C^\infty(M,N)$. It does give the means to express a classical result by Whitney, concerning approximation of C^∞ maps by real analytic maps, Lemma 6 in [54], see Lemma 13.1 (a) in this article. The result by Whitney involves approximation of partial derivatives of increasingly high order as one moves out towards infinity. This phenomenon is captured by the very-strong C^∞ topology, but *not* by the strong C^∞ topology. This is the reason why the strong C^∞ topology on $C^\infty(M,N)$ *should not* be called the Whitney C^∞ topology, it is the very-strong C^∞ topology on $C^\infty(M,N)$ that deserves to be called the Whitney C^∞ topology. We shall however in this article avoid terminology like "Mather topology" and "Whitney topology" in order to avoid any misunderstandings. We will use the terminology "strong C^r topology", $1 \le r < \infty$, "strong C^∞ topology" and "very-strong C^∞ topology".

We shall now define the *very-strong* C^∞ *topology* on $C^\infty(M,N)$. Suppose $f \in C^\infty(M,N)$, where M and N are paracompact C^∞ manifolds, with or without boundary. By a *basic very-strong* C^∞ *neighborhood of* f *in* $C^\infty(M,N)$ we mean a set of the form

$$\mathcal{U}^{\infty,\infty} = \bigcap_{i\in\Lambda} \mathcal{N}^{\infty,r_i}(f; B_i, (U_i,\varphi), (V,\psi_i), \varepsilon_i), \tag{v}$$

where $1 \le r_i < \infty$, $i \in \Lambda$, and the family $\{B_i\}_{i\in\Lambda}$ is locally finite in M. By Lemma 9.4 below the family of all basic very-strong C^∞ neighborhoods of f in $C^\infty(M,N)$, for all $f \in C^\infty(M,N)$, is a basis for a topology on $C^\infty(M,N)$. Thus we can give the following definition.

Definition 9.2 Let M and N be paracompact C^∞ manifolds, with or without boundary. The *very-strong* C^∞ *topology* on $C^\infty(M,N)$ is the topology which as a basis has the family of all basic very-strong C^∞ neighborhoods of f in $C^\infty(M,N)$,for all $f \in C^\infty(M,N)$.

In (v) it may very well be that

$$\sup\{r_i \mid i \in \Lambda\} = \infty, \tag{vi}$$

and this fact is the crucial point here. The fact that (vi) is allowed to hold makes the very-strong C^∞ topology on $C^\infty(M,N)$ to differ from the strong C^∞ topology on $C^\infty(M,N)$. Temporarily we denote $C^\infty(M,N)$ with the very-strong C^∞ topology by $C^\infty_{\mathrm{vS}}(M,N)$.

Note that the very-strong C^∞ topology on $C^\infty(M,N)$ is at least as fine as the strong C^∞ topology on $C^\infty(M,N)$. Thus the identity map on the set $C^\infty(M,N)$ gives us a continuous map

$$\mathrm{id}\colon C^\infty_{\mathrm{vS}}(M,N) \to C^\infty_{\mathrm{S}}(M,N). \tag{vii}$$

Let us now give the promised Lemma 9.4. We begin with the following fact.

Lemma 9.3 *Let* $\mathcal{N} = \mathcal{N}^{\infty,r}(f; B, (U,\varphi), (V,\psi), \varepsilon)$ *be an elementary* C^r *neighborhood of* $f \in C^\infty(M,N)$*, and let* $f_0 \in \mathcal{N}$*. Then there exists* $\varepsilon_0 > 0$ *such that if we set* $\mathcal{N}_0 = \mathcal{N}^{\infty,r}(f_0; B, (U,\varphi), (V,\psi), \varepsilon_0)$*, then* $\mathcal{N}_0 \subset \mathcal{N}$.

Proof We have that $d = \|\psi\circ f_0\circ\varphi^{-1} - \psi\circ f\circ\varphi\|^r_{\varphi(B)} < \varepsilon$, and by choosing $\varepsilon_0 = \varepsilon - d$ the claim follows. □

Lemma 9.4 *Let* $f, f' \in C^\infty(M,N)$*´, and suppose that* $\mathcal{U}$ *and* $\mathcal{U}'$ *are basic very-strong* C^∞ *neighborhoods of* f *and* f' *respectively, in* $C^\infty(M,N)$*. If* $f_0 \in \mathcal{U}\cap\mathcal{U}'$*, then there exists a basic very-strong* C^∞ *neighborhood* $\mathcal{U}_0$ *of* f_0 *in* $C^\infty(M,N)$*, such that* $\mathcal{U}_0 \subset \mathcal{U}\cap\mathcal{U}'$.

Proof Here $\mathcal{U} = \bigcap_{i\in\Lambda} \mathcal{N}_i$ and $\mathcal{U}' = \bigcap_{j\in\Gamma} \mathcal{N}'_j$, where each $\mathcal{N}_i = \mathcal{N}^{\infty,r_i}(f; B_i, (U_i,\varphi_i), (V_i,\psi_i), \varepsilon_i)$ is an elementary C^{r_i} neighborhood of f in $\mathrm{C}^\infty(M,N)$, $1 \leq r_i < \infty$, $i \in \Lambda$, and each $\mathcal{N}'_j = \mathcal{N}^{\infty,s_j}(f'; B'_j, (U'_j,\varphi'_j), (V'_j,\psi'_j), \varepsilon'_j)$ is an elementary C^{s_j} neighborhood of f' in $\mathrm{C}^\infty(M,N)$, $1 \leq s_j < \infty$, $j \in \Gamma$, and the families $\{B_i\}_{i\in\Lambda}$ and $\{B'_j\}_{j\in\Gamma}$ are locally finite in M. By Lemma 9.2 there exists for each $i \in \Lambda$ an $\varepsilon_{0,i} > 0$ such that $\mathcal{N}_{0,i} = \mathcal{N}^{\infty,r_i}(f_0; B_i, (U_i,\varphi_i), (V_i,\psi_i), \varepsilon_{0,i}) \subset \mathcal{N}_i$, and also for each $j \in \Gamma$ an $\varepsilon'_{0,j} > 0$ such that $\mathcal{N}'_{0,j} = \mathcal{N}^{\infty,s_j}(f_0; B'_j, (U'_j,\varphi'_j), (V'_j,\psi'_j), \varepsilon'_{0,j}) \subset \mathcal{N}'_j$. Since the family $\{B_i, B'_j\}_{i\in\Lambda,j\in\Gamma}$ is locally finite in M, it follows that $\mathcal{N}_0 = \bigcap_{i\in\Lambda} \mathcal{N}_{0,i} \cap \bigcap_{j\in\Gamma} \mathcal{N}'_{0,j}$ is a basic very-strong C^∞ neighborhood of f_0 in $\mathrm{C}^\infty(M,N)$, and we have that $\mathcal{U}_0 \subset \mathcal{U} \cap \mathcal{U}'$. □

The following easy lemma will be used later on in the paper.

Lemma 9.5 *Let M and N be paracompact C^s manifolds, with or without boundary, $1 \leq s \leq \infty$, and let W be an open subset of N. Then the set $\mathrm{C}^s(M,W)$ is open in $\mathrm{C}^s_{\mathrm{S}}(M,N)$, for $1 \leq s < \infty$, and in $\mathrm{C}^\infty_{\mathrm{vS}}(M,N)$, for $s = \infty$.*

Proof Let $f \in \mathrm{C}^s(M,W)$. We choose a locally finite family $\{B_i\}_{i\in\Lambda}$ of compact subsets of M such that:

(a) $\bigcup_{i\in\Lambda} B_i = M$,

(b) $B_i \subset U_i$, where (U_i,φ_i) is a chart in M, $i \in \Lambda$,

(c) $f(B_i) \subset V_i \subset W$, where (V_i,ψ_i) is a chart in N, $i \in \Lambda$.

Let $\varepsilon_i > 0$ be arbitrary, e.g., $\varepsilon_i = \infty$, $i \in \Lambda$. Then $\mathcal{U}^* = \bigcap_{i\in\Lambda} \mathcal{N}^{s,1}(f; B_i, (U_i,\varphi_i), (V_i,\psi_i), \varepsilon_i)$ is a basic strong C^1 neighborhood of f in $\mathrm{C}^s(M,N)$. Furthermore $\mathcal{U}^* \subset \mathrm{C}^s(M,W)$, since if $f' \in \mathcal{U}^*$, then $f'(M) = \bigcup_{i\in\Lambda} f'(B_i) \subset \bigcup_{i\in\Lambda} V_i \subset W$. Thus $\mathrm{C}^s(M,W)$ is open in $\mathrm{C}^{s,1}_{\mathrm{S}}(M,N)$, and hence also in $\mathrm{C}^s_{\mathrm{S}}(M,N)$, for $1 \leq s \leq \infty$, see Remark 9.1. By (vii) it now follows that $\mathrm{C}^\infty(M,W)$ is open in $\mathrm{C}^\infty_{\mathrm{vS}}(M,N)$. □

Notes Our exposition here in Section 9 follows the one we gave, for the very-strong C^∞ topology, in Section 1 of [21]. We have here simply also included the presentation of the strong C^r topologies, $1 \leq r < \infty$, into the same pattern.

10 Continuity of induced maps in the strong C^r topologies, $1 \leq r < \infty$, and in the very-strong C^∞ topology

CONVENTION: *From now on we will denote $\mathrm{C}^r(M,N)$ with the strong C^r topology, $1 \leq r < \infty$, simply by $\mathrm{C}^r(M,N)$, instead of by $\mathrm{C}^r_{\mathrm{S}}(M,N)$. Likewise we denote $\mathrm{C}^s(M,N)$ with the strong C^r topology, where $1 \leq r < s \leq \infty$, by $\mathrm{C}^{s,r}(M,N)$, instead of $\mathrm{C}^{s,r}_{\mathrm{S}}(M,N)$. Furthermore we denote $\mathrm{C}^\infty(M,N)$ with the very–strong topology by $\mathrm{C}^\infty(M,N)$, instead of $\mathrm{C}^\infty_{\mathrm{vS}}(M,N)$.*

The purpose of this section is to establish the continuity results, for induced maps between function spaces, given in Propositions 10.4 and 10.5. We use Lemma 10.1 below to prove Proposition 10.4, and Lemma 10.3 is used in the proof of Proposition 10.5.

Lemma 10.1 *Suppose that M, N and P are C^s manifolds, with or without boundary, where $1 \leq s \leq \infty$. Let $1 \leq r < \infty$ be such that $r \leq s$. Let $w\colon N \to P$ be a C^s map. Suppose $f \in \mathrm{C}^s(M,N)$, and let $\mathcal{P} = \mathcal{N}^{s,r}(w \circ f; B, (U,\varphi), (W,\xi), \varepsilon)$ be an* *elementary* C^r *neighborhood of* $w_*(f) = w \circ f$ *in* $\mathrm{C}^s(M,P)$. *Then there exist finitely many elementary* C^r *neighborhoods* $\mathcal{M}_j = \mathcal{N}^{s,r}(f; B_j, (U_j, \varphi_j), (V_j, \psi_j), \varepsilon_j)$ *of* f *in* $\mathrm{C}^s(M,N)$, $1 \le j \le q$, *such that* $w_*(\bigcap_{j=1}^q \mathcal{M}_j) \subset \mathcal{P}$.

Proof It follows directly by the definitions given in Section 9 that it is enough to prove Lemma 10.1 in the case when $s = r$, and $1 \le r < \infty$. We shall anyhow give a formal proof of this fact, so we show here that Lemma 10.1 follows from Lemma 10.1* below. After that we prove the version given in Lemma 10.1*.

Thus let us assume that Lemma 10.1* holds. We have that $\mathcal{P} = \mathrm{C}^s(M,P) \cap \mathcal{P}'$, where $\mathcal{P}' = \mathcal{N}^r(w \circ f; B, (U,\varphi), (W,\xi), \varepsilon)$ is an elementary C^r neighborhood of $w_*(f) = w \circ f$ in $\mathrm{C}^r(M,P)$. By Lemma 10.1* there then exist finitely many elementary C^r neighborhoods $\mathcal{M}'_j = \mathcal{N}^r(f; B_j, (U_j,\varphi_j), (V_j,\psi_j), \varepsilon_j)$ of f in $\mathrm{C}^r(M,N)$, $1 \le j \le q$, such that $w_*(\bigcap_{j=1}^q \mathcal{M}'_j) \subset \mathcal{P}'$. Now, for each $1 \le j \le q$, the set $\mathcal{M}_j = \mathrm{C}^s(M,N) \cap \mathcal{M}'_j$ is an elementary C^r neighborhood of f in $\mathrm{C}^s(M,N)$, and $w_*(\bigcap_{j=1}^q \mathcal{M}_j) = w_*(\bigcap_{j=1}^q (\mathrm{C}^s(M,N) \cap \mathcal{M}'_j)) = w_*(\mathrm{C}^s(M,N) \cap (\bigcap_{j=1}^q \mathcal{M}'_j)) \subset \mathrm{C}^s(M,P) \cap w_*(\bigcap_{j=1}^q \mathcal{M}'_j) \subset \mathrm{C}^s(M,P) \cap \mathcal{P}' = \mathcal{P}$. □

Lemma 10.1* *Let* M, N *and* P *be* C^r *manifolds, with or without boundary, and let* $w\colon N \to P$ *be a* C^r *map, where* $1 \le r < \infty$. *Suppose* $f \in \mathrm{C}^r(M,N)$, *and let* $\mathcal{P} = \mathcal{N}^r(w \circ f; B, (U,\varphi), (W,\xi), \varepsilon)$ *be an elementary* C^r *neighborhood of* $w_*(f) = w \circ f$ *in* $\mathrm{C}^r(M,P)$. *Then there exist finitely many elementary* C^r *neighborhoods* $\mathcal{M}_j = \mathcal{N}^r(f; B_j, (U,\varphi), (V_j,\psi_j), \varepsilon_j)$ *of* f *in* $\mathrm{C}^r(M,N)$, $1 \le j \le q$, *such that* $w_*(\bigcap_{j=1}^q \mathcal{M}_j) \subset \mathcal{P}$.

Proof Since B is compact, and $B \subset U$ and $(w \circ f)(B) \subset W$, we can find finitely many compact subsets B_j of B and charts (V'_j, ψ'_j) in N, $1 \le j \le q$, such that

(a) $B = \bigcup_{j=1}^q B_j$,

(b) $f(B_j) \subset V_j \subset \overline{V}_j \subset V'_j \subset w^{-1}(W)$, where V_j is open in N, and $\overline{V}_j$ is compact.

We denote $\psi_j = \psi'_j \mid V_j$, $1 \le j \le q$. Since $\psi'_j(\overline{V}_j)$ is compact and $\psi(V_j) \subset \psi'_j(\overline{V}_j)$, it follows that $\|\xi \circ w \circ (\psi_j)^{-1}\|^r_{\psi_j(V_j)} < \infty$, for each $1 \le j \le q$. Hence there exists $\varepsilon_j > 0$ such that if $h\colon M \to N$ is a C^r map, with $h(B_j) \subset V_j$, and

$$\|\psi_j \circ h \circ \varphi^{-1} - \psi_j \circ f \circ \varphi^{-1}\|^r_{\varphi(B_j)} < \varepsilon_j$$

then $\|(\xi \circ w \circ \psi_j^{-1})(\psi_j \circ h \circ \varphi^{-1}) - (\xi \circ w \circ \psi_j^{-1})(\psi_j \circ f \circ \varphi^{-1})\|^r_{\varphi(B_j)} < \varepsilon$, i.e., then

$$\|\xi \circ w \circ h \circ \varphi^{-1} - \xi \circ w \circ f \circ \varphi^{-1}\|^r_{\varphi(B_j)} < \varepsilon, \quad 1 \le j \le q. \tag{1}$$

We then define

$$\mathcal{M}_j = \mathcal{N}^r(f; B_j, (U,\varphi), (V_j,\psi_j), \varepsilon_j), \quad 1 \le j \le q.$$

Now suppose $h \in \bigcap_{j=1}^q \mathcal{M}_j$. Then $(w \circ h)(B) \subset W$, and by (1) we have that

$$\|\xi \circ w \circ h \circ \varphi^{-1} - \xi \circ w \circ f \circ \varphi^{-1}\|^r_{\varphi(B)} < \varepsilon.$$

Hence $w_*(h) = w \circ h \in \mathcal{N}^r(w \circ f; B, (U,\varphi), (W,\xi), \varepsilon) = \mathcal{P}$. □

If we take $N = P$ and $w = \mathrm{id}_N$ in Lemma 10.1, we see that the proof of Lemma10.1 proves the following.

Remark 10.2 Suppose M and N are C^s manifolds, with or without boundary, where $1 \leq s \leq \infty$. Let $1 \leq r < \infty$ be such that $r \leq s$. Suppose $f \in \mathrm{C}^s(M, N)$ and let $\mathcal{M} = \mathcal{N}^{s,r}(f; B, (U, \varphi), (V, \psi), \varepsilon)$ be an elementary C^r neighborhood of f in $\mathrm{C}^s(M, N)$. Suppose we are given C^r charts (V_j, ψ_j) and (V'_j, ψ'_j), where $\psi_j = \psi'_j | V_j$, in N and compact subsets B_j of B, $1 \leq j \leq q$, such that

(a) $B = \bigcup_{j=1}^q B_j$,

(b) $f(B_j) \subset V_j \subset \overline{V}_j \subset V'_j \subset V$, and $\overline{V}_j$ is compact, $1 \leq j \leq q$.

Then there exist $\varepsilon_j > 0$, $1 \leq j \leq q$, so that if $\mathcal{M}_j = \mathcal{N}^{s,r}(f; B_j, (U, \varphi), (V_j, \psi_j), \varepsilon_j)$, then $\bigcap_{j=1}^q \mathcal{M}_j \subset \mathcal{M}$.

Lemma 10.3 *Let M, N and P be C^s manifolds, with or without boundary, where $1 \leq s \leq \infty$, and let $1 \leq r < \infty$ be such that $r \leq s$. Let $v\colon M \to N$ be a C^s map. Let $f \in \mathrm{C}^s(N, P)$, and let $\mathcal{P} = \mathcal{N}^{s,r}(f \circ v; B, (U, \varphi), (W, \xi), \varepsilon)$ be an elementary C^r neighborhood of $v^*(f) = f \circ v$ in $\mathrm{C}^s(M, P)$. Then there exist finitely many elementary C^r neighborhoods $\mathcal{N}_j = \mathcal{N}^{s,r}(f; D_j, (V_j, \psi_j), (W, \xi), \varepsilon_j)$ of f in $\mathrm{C}^s(N, P)$, $1 \leq j \leq q$, such that $v^*(\bigcap_{j=1}^q \mathcal{N}_j) \subset \mathcal{P}$.*

Proof In the same way as in Lemma 10.1 we see that it is enough to prove Lemma 10.3 in the case when $s = r$, and $1 \leq r < \infty$. So let $s = r$, and assume that $1 \leq r < \infty$.

Since $v(B)$ is compact we can find finitely many compact subsets D_j of $v(B)$, and charts (V_j, ψ_j) in N, $1 \leq j \leq q$, such that

(a) $\bigcup_{j=1}^q D_j = v(B)$,

(b) $D_j \subset V_j$, $1 \leq j \leq q$.

Let us denote $B_j = B \cap v^{-1}(D_j)$, $1 \leq j \leq q$. Then each B_j is compact, $\bigcup_{j=1}^q B_j = B$ and $v(B_j) = D_j$. For each $1 \leq j \leq q$ we have that $\|\psi_j \circ v \circ \varphi^{-1}\|^r_{\varphi(B_j)} < \infty$, since $\varphi(B_j)$ is compact, and hence there exists $\varepsilon_j > 0$ such that the following holds. If $h\colon N \to P$ is a C^r map with $h(D_j) \subset W$ and

$$\|\xi \circ h \circ \psi_j^{-1} - \xi \circ f \circ \psi_j^{-1}\|^r_{\psi_j(D_j)} < \varepsilon_j,$$

then $\|(\xi \circ h \circ \psi_j^{-1} - \xi \circ f \circ \psi_j^{-1})(\psi_j \circ v \circ \varphi^{-1})\|^r_{\varphi(B_j)} < \varepsilon, \quad 1 \leq j \leq q$, i.e., then

$$\|\xi \circ h \circ v \circ \varphi^{-1} - \xi \circ f \circ v \circ \varphi^{-1}\|_{\varphi(B_j)} < \varepsilon, \quad 1 \leq j \leq q. \tag{1}$$

We define

$$\mathcal{N}_j = \mathcal{N}^r(f; D_j, (V_j, \psi_j), (W, \xi), \varepsilon_j), \quad 1 \leq j \leq q.$$

If $h \in \bigcap_{j=1}^q \mathcal{N}_j$, then $(h \circ v)(B) \subset W$ and by (1) we have

$$\|\xi \circ h \circ v \circ \varphi^{-1} - \xi \circ f \circ v \circ \varphi^{-1}\|^r_{\varphi(B)} < \varepsilon.$$

Hence $v^*(h) = h \circ v \in \mathcal{N}^r(f \circ v; B, (U, \varphi), (W, \xi), \varepsilon) = \mathcal{P}$. □

In Propositions 10.4 and 10.5 below K denotes a compact Lie group. However, since the role of the group K, with respect to the claims in Propositions 10.4 and 10.5, is completely formal, K could equally well be any Lie group, or in fact any topological group.

Proposition 10.4 *Let M, N and P be paracompact* C^s *K-manifolds, with or without boundary, and let $w\colon N \to P$ be a K-equivariant* C^s *map, where $1 \le s \le \infty$. Then the induced map*

$$w_*\colon \mathrm{C}^{s,K}(M,N) \to \mathrm{C}^{s,K}(M,P), \;\; f \mapsto w \circ f,$$

is continuous, $1 \le s \le \infty$.

Proof Since the function space $\mathrm{C}^{s,K}(M,N)$, of all K-equivariant C^s maps from M to N, has the relative topology from $\mathrm{C}^s(M,N)$, and similarly for $\mathrm{C}^{s,K}(M,P)$, it is enough to prove the result when $K = \{e\}$. We shall first give the proof in the case $s = \infty$. Recall that the topology on $\mathrm{C}^\infty(\cdot,\cdot)$ is the very-strong C^∞ topology.

Suppose $f \in \mathrm{C}^\infty(M,N)$, and let $\mathcal{W} = \bigcap_{i\in\Lambda}\mathcal{P}_i$ be a basic very-strong C^∞ neighborhood of $w_*(f) = w \circ f$ in $\mathrm{C}^\infty(M,P)$. Each $\mathcal{P}_i = \mathcal{N}^{\infty,r_i}(w\circ f; B_i, (U_i,\varphi_i), (W_i,\zeta_i), \varepsilon_i)$, is an elementary C^{r_i} neighborhood of $w \circ f$ in $\mathrm{C}^\infty(M,P)$, where $1 \le r_i < \infty$, $i \in \Lambda$. Furthermore the family $\{B_i\}_{i\in\Lambda}$ is locally finite in M. By Lemma 10.1 there exist for each $\mathcal{P}_i$, $i \in \Lambda$, finitely many elementary C^{r_i} neighborhoods $\mathcal{M}_{i,j} = \mathcal{N}^{\infty,r_i}(f; B_{i,j}, (U_i,\varphi_i), (V_{i,j},\psi_{i,j}), \varepsilon_{i,j})$ of f in $\mathrm{C}^\infty(M,N)$, $1 \le j \le q(i)$, such that $w_*(\bigcap_{j=1}^{q(i)}\mathcal{M}_{i,j}) \subset \mathcal{P}_i$. Furthermore we have, see the proof of Lemma 10.1*, that $\bigcup_{j=1}^{q(i)} B_{i,j} = B_i$, for each $i \in \Lambda$. Thus the family $\{B_{i,j} \mid i \in \Lambda,\ 1 \le j \le q(i)\}$ is locally finite in M, and hence $\mathcal{U} = \bigcap_{i\in\Lambda}\bigcap_{j=1}^{q(i)}\mathcal{M}_{i,j}$ is a basic very-strong C^∞ neighborhood of f in $\mathrm{C}^\infty(M,N)$. Furthermore we have that $w_*(\mathcal{U}) \subset \bigcap_{i\in\Lambda} w_*(\bigcap_{j=1}^{q(i)}\mathcal{M}_{i,j}) \subset \bigcap_{i\in\Lambda}\mathcal{P}_i = \mathcal{W}$. This completes the proof in the case when $s = \infty$.

In the case when $1 \le s < \infty$, the proof is completely similar. In this case we simply have that $r_i = s$, for all $i \in \Lambda$. □

Proposition 10.5 *Let M, N and P be paracompact* C^s *K-manifolds, with or without boundary, where $1 \le s \le \infty$. Suppose that $v\colon M \to N$ is a K-equivariant proper* C^s *map. Then the induced map*

$$v^*\colon \mathrm{C}^{s,K}(N,P) \to \mathrm{C}^{s,K}(M,P), \;\; f \mapsto f \circ v,$$

is continuous, $1 \le s \le \infty$.

Proof As in Proposition 10.4 it is enough to prove the result for $K = \{e\}$. We begin by proving the proposition in the case $s = \infty$. Recall that the topology on $\mathrm{C}^\infty(\cdot,\cdot)$ is the very-strong C^∞ topology.

Let $f \in \mathrm{C}^\infty(N,P)$, and suppose $\mathcal{W} = \bigcap_{i\in\Lambda}\mathcal{P}_i$ is any basic very-strong C^∞ neighborhood of $v^*(f) = f \circ v$ in $\mathrm{C}^\infty(M,P)$. Here each $\mathcal{P}_i = \mathcal{N}^{\infty,r_i}(f \circ v; B_i, (U_i,\varphi_i), (W_i,\xi_i), \varepsilon_i)$ is an elementary C^{r_i} neighborhood of $f \circ v$ in $\mathrm{C}^\infty(M,P)$, and $1 \le r_i < \infty$, $i \in \Lambda$. Furthermore the family $\{B_i\}_{i\in\Lambda}$ is locally finite in M. By Lemma 10.3 there exist for each $\mathcal{P}_i$, $i \in \Lambda$, finitely many elementary C^{r_i} neighborhoods $\mathcal{N}_{i,j} = \mathcal{N}^{r_i}(f; D_{i,j}, (V_{i,j},\psi_{i.j}), (W_i,\xi_i), \varepsilon_i)$ of f in $\mathrm{C}^\infty(N,P)$, $1 \le j \le q(i)$, such that $v^*(\bigcap_{j=1}^{q(i)}\mathcal{N}_{i,j} \subset \mathcal{P}_i$. Furthermore we have, by the proof of Lemma 10.3, that $\bigcup_{j=1}^{q(i)} D_{i,j} = v(B_i)$, for each $i \in \Lambda$. Since $\{B_i\}_{i\in\Lambda}$ is locally finite in M and the map $v : M \to N$ is proper it follows that the family $\{v(B_i)\}_{i\in\Lambda}$ is locally finite in N. Hence the family $\{D_{i,j} \mid i \in \Lambda,\ 1 \le j \le q(i)\}$ is locally finite in N. Therefore $\mathcal{V} = \bigcap_{i\in\Lambda}\bigcap_{j=1}^{q(i)}\mathcal{N}_{i,j}$ is a basic very-strong C^∞ neighborhood of f in $\mathrm{C}^\infty(N,P)$. Furthermore we have that

$v^*(\mathcal{V}) \subset \bigcap_{i \in \Lambda} v^*(\bigcap_{j=1}^{q(i)} \mathcal{N}_{i.j}) \subset \bigcap_{i \in \Lambda} \mathcal{P}_i = \mathcal{W}$. This completes the proof in the case $s = \infty$. In the case when $1 \leq s < \infty$, the proof is completely similar. We have in this case that $r_i = s$, for all $i \in \Lambda$. $\square$

Remark 10.6 Proposition 10.4 and 10.5 also hold in the real analytic case, for the simple reason that the topology on $\mathrm{C}^\omega(\cdot,\cdot)$ is the relative topology from $\mathrm{C}^\infty(\cdot,\cdot)$ equipped with the very-strong C^∞ topology. This is why we have chosen to not include the C^ω case in the formulations of Propositions 10.4 and 10.5, and the same applies for analogous situations later on in the article.

Remark 10.7 Let M and N be paracompact C^s manifolds, where $1 \leq s \leq \infty$, and let $M_i,\ i \in \Gamma$, and $N_j,\ j \in \Lambda$, be the connected components of M and N, respectively. Then M equals the topological union of the manifolds $M_i,\ i \in \Gamma$, which we denote by $M = \coprod_{i \in \Gamma} M_i$, and similarly $N = \coprod_{j \in \Lambda} N_j$. If $f \in \mathrm{C}^s(M,N),\ 1 \leq s \leq \infty$, then we denote $f^{(i)} = f| : M_i \to N,\ i \in \Gamma$. This gives us a canonical bijection

$$\mathrm{D} : \mathrm{C}^s(M,N) \to \prod_{i \in \Gamma} \mathrm{C}^s(M_i,N)\ ,\ f \mapsto (f^{(i)})_{i \in \Gamma} \tag{1}$$

It is an immediate consequence of the definitions of the topology on $\mathrm{C}^s(\cdot,\cdot),\ 1 \leq s \leq \infty$, that both D and its inverse are continuous. Thus D in (1) is a homeomorphism and we may use it to identify the two sides in (1).

Now suppose that $i \in \Gamma$ is fixed, and let $\tilde{f} \in \mathrm{C}^s(M_i,N),\ 1 \leq s \leq \infty$. Since $\tilde{f}(M_i)$ is connected there exists a unique $j = \tilde{f}(i) \in \Lambda$, such that $\tilde{f}(M_i) \subset N_{\tilde{f}(i)}$. Since each N_j, is open in N it follows, see Lemma 9.5, that each $\mathrm{C}^s(M_i,N_j),\ j \in \Lambda$, is open in $\mathrm{C}^s(M_i,N)$, and hence each $\mathrm{C}^s(M_i,N_j),\ j \in \Lambda$, is both open and closed in $\mathrm{C}^s(M_i,N)$. Therefore the natural bijection

$$\mathrm{E} : \mathrm{C}^s(M_i,N) \to \coprod_{j \in \Lambda} \mathrm{C}^s(M_i,N_j), \tag{2}$$

where $\mathrm{E}(\tilde{f} \colon M_i \to N) = \tilde{f} \colon M_i \to N_{\tilde{f}(i)}$, is a homeorphism. Note that both M_i and N_j in $\mathrm{C}^s(M_i,N_j)$, in (2), are second countable, see Theorem 2.1.

Remark 10.8 Let M and N be paracompact C^s K-manifolds, where K is a compact Lie group and $1 \leq s \leq \infty$. By a K-component of M we mean a set M_μ of the form $M_\mu = KM_i$, where M_i is a connected component of M. Since K is compact each K-component M_μ of M is a finite union of connected components of M, and hence M_μ is second countable. If $M_\mu\ ,\ \mu \in \Theta$, are the K-components of M, we have that $M = \coprod_{\mu \in \Theta} M_\mu$. Similarly we have that $N = \coprod_{\nu \in \Omega} N_\nu$, where $N_\nu\ ,\ \nu \in \Omega$, are the K-components of N.

Completely analogously to the case in Remark 10.7, we obtain a canonical homeomorphism

$$\mathrm{D} : \mathrm{C}^s(M,N) \to \prod_{\mu \in \Theta} \mathrm{C}^s(M_\mu,N),\ f \mapsto (f^{(\mu)})_{\mu \in \Theta},$$

where $f^{(\mu)} = f| : M_\mu \to N,\ \mu \in \Theta$. Since every $M_\mu,\ \mu \in \Theta$, is a C^s K-manifold the homeomorphism D induces a homeomorphism

$$\overline{\mathrm{D}} : \mathrm{C}^{s,K}(M,N) \to \prod_{\mu \in \Theta} \mathrm{C}^{s,K}(M_\mu,N),\ f \mapsto (f^{(\mu)})_{\mu \in \Theta}. \tag{1}$$

Now suppose that $\mu \in \Theta$ is fixed, and let $\tilde{f} \in \mathrm{C}^{s,K}(M_\mu, N),\ 1 \leq s \leq \infty$. Then $\tilde{f}(M_\mu) = \tilde{f}(KM_i) = K\tilde{f}(M_i) \subset KN_{\tilde{f}(i)}$, for some $i \in \Gamma$, and $KN_{\tilde{f}(i)} = N_{\tilde{f}(\mu)}$ is a K-component of N. Thus we see that for $\mu \in \Theta$ there exists a unique $\nu = \tilde{f}(\mu) \in \Omega$, such that $\tilde{f}(M_\mu) \subset N_{\tilde{f}(\mu)}$. Analogously to the case in (2) in Remark 10.8 we obtain a homeomorphism

$$\overline{\mathrm{E}} : \mathrm{C}^{s,K}(M_\mu, N) \to \coprod_{\nu\in\Omega} \mathrm{C}^{s,K}(M_\mu, N_\nu), \tag{2}$$

where $\overline{\mathrm{E}}(\tilde{f} \colon M_\mu \to N) = \tilde{f} \colon M_\mu \to N_{\tilde{f}(\mu)}$. In $\mathrm{C}^{s,K}(M_\mu, N_\nu)$, in (2), both M_μ and N_ν are second countable C^s K-manifolds.

Notes The proofs of Lemmas 10.1 and 10.3, and Propositions 10.4 and 10.5 follow the ones we gave, for the very-strong C^∞ topology, in [21], Section 2.

11 The product theorem

By the product theorem we mean the result in Proposition 11.1 below. Using the product theorem and also Proposition 10.5 we prove Corollary 11.2, which will be important for us later on in the paper. Corollary 11.2 is used in the proof of Theorem 15.4 and in the proof of Lemma 17.1.

Proposition 11.1 *Let M, N_1 and N_2 be paracompact C^s K-manifolds, with or without boundary, where $1 \leq s \leq \infty$, and suppose that either $\partial N_1 = \emptyset$ or $\partial N_2 = \emptyset$. Let $q_j \colon N_1 \times N_2 \to N_j,\ j = 1, 2$, denote the projection maps. Then the natural bijection*

$$\iota \colon \mathrm{C}^{s,K}(M, N_1 \times N_2) \to \mathrm{C}^{s,K}(M, N_1) \times \mathrm{C}^{s,K}(M, N_2),\quad f \mapsto (q_1 \circ f, q_2 \circ f),$$

is a homeomorphism, $1 \leq s \leq \infty$.

Proof Since $\mathrm{C}^{s,K}(M, N_1 \times N_2)$ has the relative topology from $\mathrm{C}^s(M, N_1 \times N_2)$, and since the cartesian product topology on $\mathrm{C}^{s,K}(M, N_1) \times \mathrm{C}^{s,K}(M, N_2)$ equals the relative topology from the product $\mathrm{C}^s(M, N_1) \times \mathrm{C}^s(M, N_2)$, it is enough to give the proof in the case when $K = \{e\}$.

First we give the proof in the case $s = \infty$. It follows by Proposition 10.4 that ι is continuous. The fact that ι^{-1} is continuous is seen as follows. Let $(f_1, f_2) \in \mathrm{C}^\infty(M, N_1) \times \mathrm{C}^\infty(M, N_2)$, and denote $\iota^{-1}(f_1, f_2) = f$. Let $\mathcal{V} = \bigcap_{i\in\Lambda} \mathcal{N}_i$ be any basic very-strong C^∞ neighborhood of f in $\mathrm{C}^\infty(M, N_1 \times N_2)$. Here each $\mathcal{N}_i = \mathcal{N}^{\infty, r_i}(f; B_i, (U_i, \varphi_i), (V_i, \psi_i), \varepsilon_i),\ i \in \Lambda$, is an elementary C^{r_i} neighborhood of f in $\mathrm{C}^\infty(M, N_1 \times N_2)$, where $1 \leq r_i < \infty,\ i \in \Lambda$, and the family $\{B_i\}_{i\in\Lambda}$ is locally finite in M. It follows by Remark 10.2 that we may assume that each $\mathcal{N}_i,\ i \in \Lambda$, is of the form where the chart in $N_1 \times N_2$ is a product chart. That is, we can assume that each $\mathcal{N}_i$ is of the form

$$\mathcal{N}_i = \mathcal{N}^{\infty, r_i}(f; B_i, (U_i, \varphi_i), (V_i^{(1)} \times V_i^{(2)}, \psi_i^{(1)} \times \psi_i^{(2)}), \varepsilon_i),\ \ i \in \Lambda. \tag{1}$$

It is readily seen that if $\mathcal{N}_i$ is as in (1), and we set

$$\mathcal{N}_i^{(j)} = \mathcal{N}^{\infty, r_i}(f_j; B_i, (U_i, \varphi_i), (V_i^{(j)}, \psi_i^{(j)}), \varepsilon_i),\ \ j = 1, 2,$$

then $\iota^{-1}(\mathcal{N}_i^{(1)} \times \mathcal{N}_i^{(2)}) \subset \mathcal{N}_i,\ i \in \Lambda$. Now $\mathcal{U}^{(j)} = \bigcap_{i\in\Lambda}\mathcal{N}_i^{(j)}$ is a basic very-strong C^∞ neighborhood of $f_j = q_j \circ f$ in $\mathrm{C}^\infty(M, N_j),\ j = 1, 2$. Furthermore $\iota^{-1}(\mathcal{U}^{(1)} \times \mathcal{U}^{(2)}) \subset \bigcap_{i\in\Lambda}\mathcal{N}_i = \mathcal{V}$. This proves that ι^{-1} is continuous. This completes the proof in the case when $s = \infty$. In the case when s is finite the proof is completely similar. In this case we simply have that $r_i = s$, for all $i \in \Lambda$. □

Corollary 11.2 *Suppose that M and N are paracompact C^s K-manifolds, and let Q be a compact C^s K-manifold, with or without boundary, $1 \le s \le \infty$. Then the map*

$$\chi\colon \mathrm{C}^{s,K}(M,N) \to \mathrm{C}^{s,K}(Q\times M, Q\times N),\ \ f \mapsto \mathrm{id}\times f,$$

is continuous, $1 \le s \le \infty$.

Proof It follows by Proposition 11.1 that it is enough to prove that the maps

$$\mathrm{C}^{s,K}(M,N) \to \mathrm{C}^{s,K}(Q\times M, Q),\quad f \mapsto q_1 \circ (\mathrm{id}\times f) \tag{1}$$

and

$$\mathrm{C}^{s,K}(M,N) \to \mathrm{C}^{s,K}(Q\times M, N),\ \ f \mapsto q_2 \circ (\mathrm{id}\times f) \tag{2}$$

are continuous. Here $q_1\colon Q\times N \to Q$ and $q_2\colon Q\times N \to N$ denote the projection maps. The map in (1) is the constant map from $\mathrm{C}^{s,K}(M,N)$ onto the element $p_1 \in \mathrm{C}^{s,K}(Q\times M, Q)$, where $p_1\colon Q\times M \to Q$ is the projection, and hence (1) is continuous.

Observe that $q_2 \circ (\mathrm{id}\times f) = f \circ p_2$, where $p_2\colon Q\times M \to M$ is the projection. Thus the map in (2) equals the map $p_2^*\colon \mathrm{C}^{s,K}(M,N) \to \mathrm{C}^{s,K}(Q\times M, N),\ f \mapsto f\circ p_2$. Since Q is compact, the projection p_2 is a proper map, and hence p_2^* is continuous by Proposition 10.5. □

Notes The exposition here in Section 11 follows the one, given for the very-strong C^∞ topology, in Section 3 in [21].

12 The equivariant glueing lemma

Lemma 12.1 *Let $f\colon M \to N$ be a K-equivariant C^s map between C^s K-manifolds, where $1 \le s \le \infty$, and let U be a K-invariant open subset of M. Then there exists an open neighborhood $\mathcal{N}$ of $f|U$ in $\mathrm{C}^{s,K}(U,N)$ such that the following holds: If $h \in \mathcal{N}$ and we define $E(h)\colon M \to N$ by*

$$E(h)(x) = \begin{cases} h(x), & x \in U \\ f(x), & x \in M - U, \end{cases}$$

then $E(h)$ is a K-equivariant C^s map, $1 \le s \le \infty$. Furthermore $E\colon \mathcal{N} \to \mathrm{C}^{s,K}(M,N),\ h \mapsto E(h)$, is continuous.

Proof It is clear that it is enough to prove the lemma in the case when $K = \{e\}$, and this is given in [6], I.4.3.4.4, for $s = \infty$. If $1 \le s < \infty$, the topology on $\mathrm{C}^s(\cdot,\cdot)$ is the strong C^s topology, and one can for example refer to Lemma 2.2.8 in [15]. □

13 Whitney approximation

The following two basic results were proved by H. Whitney in 1932-33.

Lemma 13.1 (a) (H. Whitney) *Let U be an open subset of $\mathbb{R}^m$, and let $U_1, U_2, \ldots$ be open subsets of U (some of which may be empty) such that $\overline{U}_q$ is compact and $\overline{U}_q \subset U_{q+1}$ for all $q \geq 1$, and $\bigcup_{q=1}^{\infty} U_q = U$. Then if $f\colon U \to \mathbb{R}^n$ is a C^∞ map, and $\varepsilon_1 \geq \varepsilon_2 \geq \ldots$ are given positive real numbers, and $r_1 \leq r_2 \leq \ldots$ are given positive integers, there is real analytic map $h\colon U \to \mathbb{R}^n$ such that, for each $1 \leq j \leq n$, we have that*

$$|\mathrm{D}^\alpha (h-f)_j(x)| < \varepsilon_q, \text{ for all } x \in U - U_q,$$

and all $\alpha = (\alpha_1, \ldots, \alpha_m)$ with $|\alpha| \leq r_q,\ q = 1, 2, \ldots$. (Here $(h-f)_j : U \to \mathbb{R}$, denotes the j:th component of the map $h - f : U \to \mathbb{R}^n$.)

Proof See [54], Lemma 6. In [54] the formulation of this result is given in the case when $r_q = q$, for $q \geq 1$. The above form of the result is an immediate consequence of this one. □

Lemma 13.1 (**b**) (H. Whitney) *Let U be an open subset of $\mathbb{R}^m$, and let U_1, U_2 be open subsets of U (some of which may be empty) such that $\overline{U}_q$ is compact and $\overline{U}_q \subset U_{q+1}$, for all $q \geq 1$, and $\bigcup_{q=1}^{\infty} U_q = U$. Let $f : U \to \mathbb{R}^n$ be a C^r map, where $1 \leq r < \infty$, and let $\varepsilon_1 \geq \varepsilon_2 \geq \ldots$ be given positive real numbers. Then there exists a real analytic map $h : U \to \mathbb{R}^n$ such that, for each $1 \leq j \leq n$, we have that*

$$|\mathrm{D}^\alpha (h-f)_j(x)| < \varepsilon_q \ , \text{ for all } x \in U - U_q \ , \ q \geq 1,$$

and all $\alpha = (\alpha_1, \ldots, \alpha_m)$, with $|\alpha| \leq r$.

Proof See [54], Lemma 6. □

We begin by showing that Lemmas 13.1(a) and (b) give us the following result.

Proposition 13.2 *Let U be an open subset of $\mathbb{R}^m$. Then the set $\mathrm{C}^\omega(U, \mathbb{R}^n)$ is dense in $\mathrm{C}^s(U, \mathbb{R}^n)$, where $1 \leq s \leq \infty$. (Recall that in the case $s = \infty$, the topology on $\mathrm{C}^\infty(U, \mathbb{R}^n)$ is the very-strong C^∞ topology.)*

Proof Let us first give the proof in the case $s = \infty$. Suppose $f \in \mathrm{C}^\infty(U, \mathbb{R}^n)$ and let $\mathcal{U} = \bigcap_{i \in \Lambda} \mathcal{N}^{\infty, r_i}(f; B_i, \varepsilon_i)$ be any basic very-strong C^∞ neighborhood of f in $\mathrm{C}^\infty(U, \mathbb{R}^n)$. Here $1 \leq r_i < \infty$, and $\varepsilon_i > 0,\ i \in \Lambda$, and each $B_i,\ i \in \Lambda$, is a non-empty compact subset of U, and the family $\{B_i\}_{i \in \Lambda}$ is locally finite in U.

First we choose bounded open subsets $\emptyset = U_1, U_2, \ldots$ of U such that

(a) $\bigcup_{q=1}^{\infty} U_q = U$,

(b) $\overline{U}_q \subset U_{q+1},\ q = 1, 2, \ldots$.

Next we define subsets Λ_q of $\Lambda,\ q \geq 1$, in the following way. We set

$$\Lambda_q = \{i \in \Lambda \mid B_i \cap U_{q+1} \neq \emptyset\}, \ \ q = 1, 2, \ldots.$$

Clearly $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_q \subset \Lambda_{q+1} \subset \cdots$, and $\bigcup_{q=1}^{\infty} \Lambda_q = \Lambda$. Since $\overline{U}_{q+1}$ is a compact subset of U and the family $\{B_i\}_{i \in \Lambda}$ is locally finite in U, it follows that $B_i \cap \overline{U}_{q+1} \neq \emptyset$, and hence also $B_i \cap U_{q+1} \neq \emptyset$, for only finitely many $i \in \Lambda$. Thus each Λ_q is a finite set. We define

$$\begin{aligned} \overline{r}_q &= \max\{r_i \mid i \in \Lambda_q\}, \quad q = 1, 2, \ldots, \\ \overline{\varepsilon}_q &= \min\{\varepsilon_i \mid i \in \Lambda_q\}, \quad q = 1, 2, \ldots. \end{aligned} \tag{1}$$

Then $\overline{r}_1 \leq \overline{r}_2 \leq \ldots$, and $\overline{\varepsilon}_1 \geq \overline{\varepsilon}_2 \geq \ldots$.

By Lemma 13.1(a) there exists a real analytic map $h\colon U \to \mathbb{R}^n$ such that, for each $1 \leq j \leq n$,

$$|\mathrm{D}^\alpha (h - f)_j(x)| < \overline{\varepsilon}_q, \quad \text{for all } x \in U - U_q, \tag{2}$$

and all $\alpha = (\alpha_1, \ldots, \alpha_m)$ with $|\alpha| \leq \overline{r}_q,\ q = 1, 2, \ldots$.

Now consider a fixed compact set $B_i,\ i \in \Lambda$. We let $q(i)$ be the least integer for which $i \in \Lambda_{q(i)}$. Thus

$$i \in \Lambda_{q(i)} - \Lambda_{q(i)-1},$$

where $\Lambda_0 = \emptyset$. Since $i \notin \Lambda_{q(i)-1}$ we have that

$$B_i \subset U - U_{q(i)}. \tag{3}$$

For any $i \in \Lambda$ have that $i \in \Lambda_{q(i)}$, and hence we have by (1) that

$$\overline{r}_{q(i)} \geq r_i, \quad \text{and } \overline{\varepsilon}_{q(i)} \leq \varepsilon_i, \quad \text{for every } i \in \Lambda.$$

It now follows by (2) and (3) that, for each $1 \leq j \leq n$,

$$|\mathrm{D}^\alpha (h - f)_j(x)| < \overline{\varepsilon}_{q(i)} \leq \varepsilon_i, \quad \text{for all } x \in B_i,$$

and all $\alpha = (\alpha_1, \ldots, \alpha_m)$ with $|\alpha| \leq \overline{r}_{q(i)}$, and hence in particular for all α with $|\alpha| \leq r_i,\ i \in \Lambda$. Thus we have that

$$\|h - f\|_{B_i}^{r_i} < \varepsilon_i, \quad i \in \Lambda.$$

Hence $h \in \bigcap_{i \in \Lambda} \mathcal{N}^{\infty, r_i}(f; B_i, \varepsilon_i) = \mathcal{U}$. Now $h \in \mathcal{U} \cap \mathrm{C}^\omega(U, \mathbb{R}^n)$, and this proves that the set $\mathrm{C}^\omega(U, \mathbb{R}^n)$ is dense in $\mathrm{C}^\infty(U, \mathbb{R}^n)$.

In the case when $1 \leq s < \infty$, the proof is entirely similar to the above one, and uses Lemma 13.1(b). In fact it is conceptually simpler in this case, since we then have that $r_i = s$, for all $i \in \Lambda$. □

We will need the following deep result, the Grauert-Morrey imbedding theorem, concerning the existence of real analytic imbeddings of real analytic manifolds into euclidean space.

Theorem 13.3 *Let M be a second countable real analytic manifold. Then there exists a real analytic closed imbedding $h : M \to \mathbb{R}^u$, into some euclidean space $\mathbb{R}^u$.*

Proof See Theorem 3 in [12]. In the case when M is compact, this was proved in [41]. □

Using Proposition 13.2 and Theorem 13.3 we prove the following.

Theorem 13.4 *Let M be a paracompact real analytic manifold. Then the set $\mathrm{C}^\omega(M, \mathbb{R}^n)$ is dense in the space $\mathrm{C}^s(M, \mathbb{R}^n),\ 1 \leq s \leq \infty$.*

Proof It follows by Remark 10.7 that is suffices to prove Theorem 13.4 in the case when M is connected. In this case M is second countable, see Proposition 2.1, and hence we may, by the Grauert-Morrey imbedding theorem, Theorem 13.3 above, consider M as a real analytic closed submanifold of some euclidean space $\mathbb{R}^u$. Let $i\colon M \hookrightarrow \mathbb{R}^u$ denote the inclusion. Now, let $1 \leq s \leq \infty$. By Proposition 10.5 the induced map

$$i^*\colon \mathrm{C}^s(\mathbb{R}^u, \mathbb{R}^n) \to \mathrm{C}^s(M, \mathbb{R}^n), \quad f \mapsto f|M,$$

is continuous. Furthermore i^* is surjective, since each C^s map $f'\colon M \to \mathbb{R}^n$ can be extended to a C^s map $f\colon \mathbb{R}^u \to \mathbb{R}^n$, see e.g. [43], Proposition 2.5.14.

Let $\mathcal{U}$ be a non-empty open subset of $\mathrm{C}^s(M, \mathbb{R}^n)$. Then $(i^*)^{-1}(\mathcal{U})$ is a non-empty open subset of $\mathrm{C}^s(\mathbb{R}^u, \mathbb{R}^n)$, and hence we have by Proposition 13.2 that there exists a real analytic map $h\colon \mathbb{R}^u \to \mathbb{R}^n$ such that $h \in (i^*)^{-1}(\mathcal{U})$. Then $h \circ i = h|\colon M \to \mathbb{R}^n$ is real analytic, and thus $h \circ i = i^*(h) \in \mathcal{U} \cap \mathrm{C}^\omega(M, \mathbb{R}^n)$. □

Notes The C^∞ case of Theorem 13.4 is Proposition 4.3 in [21], and our exposition here in Section 13 follows the one in Section 4 in [21].

14 Haar integrals of C^s maps, $1 \leq s \leq \infty$, and of real analytic maps

We will freely use the basic properties of the Haar integral for a compact group, see e.g. Theorem 0.3.1 in [5]. For an elementary proof of the existence and uniqueness of the Haar integral for compact groups, due to von Neumann, we refer to [49].

Let K be a compact Lie group and let M be any C^t manifold, where $1 \leq t \leq \omega$. Suppose

$$\zeta\colon K \times M \to \mathbb{R} \tag{i}$$

is a real-valued C^t map, $1 \leq t \leq \omega$. Then we define

$$\hat{\mathrm{A}}(\zeta)\colon M \to \mathbb{R} \tag{ii}$$

by

$$\hat{\mathrm{A}}(\zeta)(x) = \int_K \zeta(k, x) dk, \quad \text{for each } x \in M.$$

Here the integral is the Haar integral. For the fact that $\hat{\mathrm{A}}(\zeta)\colon M \to \mathbb{R}$ is continuous, see e.g. [5], Proposition 0.3.2. We will need to know that if $\zeta\colon K \times M \to \mathbb{R}$ is a C^t map, where $1 \leq t \leq \omega$, then $\hat{\mathrm{A}}(\zeta)\colon M \to \mathbb{R}$ is also a C^t map. This result is given in Proposition 14.4. Let us first record two standard results, given in Lemmas 14.1 and 14.2 below.

Lemma 14.1 *Let K be a compact Lie group, and let U be an open subset of $\mathbb{R}^m$, $m \geq 1$. Suppose $\zeta\colon K \times U \to \mathbb{R}$ is a real-valued C^s map, where $1 \leq s \leq \infty$. Then $\hat{\mathrm{A}}(\zeta)\colon U \to \mathbb{R}$ is a C^s map, $1 \leq s \leq \infty$.*

Proof See e.g. [5], Theorem 0.3.3. □

Lemma 14.2 *Let U be an open subset of $\mathbb{R}^m$, $m \geq 1$, and let $f\colon U \to \mathbb{R}$ be a real-valued C^∞ map. Then f is real analytic if and only if for each compact subset B of U there exists a constant $Q > 0$, such that $|\mathrm{D}^\alpha f(x)| \leq Q^{|\alpha|+1}\alpha!$, for every $x \in B$ and all m-tuples $\alpha = (\alpha_1, \ldots, \alpha_m)$. (Here $|\alpha| = \alpha_1 + \cdots + \alpha_m$, and $\alpha! = \alpha_1!\alpha_2! \cdots \alpha_m!$.)*

Proof See [42], Proposition 1.1.14. □

We can now prove the following.

Lemma 14.3 *Let K be a compact Lie group, and let U be an open subset of $\mathbb{R}^m$, $m \geq 1$. Suppose $\zeta\colon K \times U \to \mathbb{R}$ is a real analytic real-valued map. Then $\hat{\mathrm{A}}(\zeta)\colon U \to \mathbb{R}$ is real analytic.*

Proof We know by Lemma 14.1 that $\hat{\mathrm{A}}(\zeta)\colon U \to \mathbb{R}$ is a C^∞ map. Thus it only remains to verify that the condition in Lemma 14.2 holds. Let B be any compact subset of U. Since the Lie group K is compact we can find finitely many charts (V_j, ψ_j) in K and compact subsets $E_j \subset V_j$, $1 \leq j \leq q$, such that $\bigcup_{j=1}^q E_j = K$. Now $\psi_j(V_j) \times U$ is an open subset of $\mathbb{R}^{l+m}$, where $l = \dim K$, and $\zeta \circ (\psi_j^{-1} \times \mathrm{id})\colon \psi_j(V_j) \times U \to \mathbb{R}$ is a real analytic real-valued map, $1 \leq j \leq q$. Since $\psi_j(E_j) \times B$ is compact there exists, by Lemma 14.2, for each $1 \leq j \leq q$, a constant $Q_j > 0$, such that

$$\left|\mathrm{D}^\gamma \zeta(\psi_j^{-1}(y), x)\right| \leq Q_j^{|\gamma|+1} \gamma!$$

for every $(y, x) \in \psi_j(E_j) \times B$, and all $(l+m)$-tuples $\gamma = (\beta_1, \ldots, \beta_l, \alpha_1, \ldots, \alpha_m)$. By taking $Q = \max\{Q_j | 1 \leq j \leq q\}$, and choosing $\beta_1 = \cdots = \beta_l = 0$ we see that

$$|\mathrm{D}^\alpha \zeta(k, x)| \leq Q^{|\alpha|+1} \alpha!$$

for every $k \in \bigcup_{j=1}^q E_j = K$, and every $x \in B$, and all m-tuples $\alpha = (\alpha_1, \ldots, \alpha_m)$. Since

$$\mathrm{D}^\alpha \hat{\mathrm{A}}(\zeta)(x) = \mathrm{D}^\alpha \int_K \zeta(k, x) dk = \int_K \mathrm{D}^\alpha \zeta(k, x) dk$$

it now follows that

$$\left|\mathrm{D}^\alpha \hat{\mathrm{A}}(\zeta)(x)\right| = \left|\int_K \mathrm{D}^\alpha \zeta(k, x) dk\right| \leq \int_K |\mathrm{D}^\alpha \zeta(k, x)|\, dk \leq \int_K Q^{|\alpha|+1} \alpha! dk = Q^{|\alpha|+1} \alpha!$$

for every $x \in B$, and all m-tuples $\alpha = (\alpha_1, \ldots, \alpha_m)$. Hence we have, by Lemma 14.2, that $\hat{\mathrm{A}}(\zeta)\colon U \to \mathbb{R}$ is real analytic. □

Proposition 14.4 *Let K be a compact Lie group, and let M be any C^t manifold, where $1 \leq t \leq \omega$. If $\zeta\colon K \times M \to \mathbb{R}$ is a C^t map, then $\hat{\mathrm{A}}(\zeta)\colon M \to \mathbb{R}$ is also a C^t map, $1 \leq t \leq \omega$.*

Proof Suppose (U, φ) is a chart in the C^t manifold M, $1 \leq t \leq \omega$. We need to show that $\hat{\mathrm{A}}(\zeta) \circ \varphi^{-1}\colon \varphi(U) \to \mathbb{R}$ is a C^t map, $1 \leq t \leq \omega$. Now

$$\begin{aligned}(\hat{\mathrm{A}}(\zeta) \circ \varphi^{-1})(x) &= \int_K \zeta(k, \varphi^{-1}(x)) dk \\ &= \int_K (\zeta \circ (\mathrm{id} \times \varphi^{-1}))(k, x) dk = \hat{\mathrm{A}}(\zeta \circ (\mathrm{id} \times \varphi^{-1}))(x)\end{aligned}$$

for every $x \in \varphi(U)$. Thus $\hat{\mathrm{A}}(\zeta) \circ \varphi^{-1} = \hat{\mathrm{A}}(\zeta \circ (\mathrm{id} \circ \varphi^{-1}))\colon \varphi(U) \to \mathbb{R}$, and since $\zeta \circ (\mathrm{id} \times \varphi^{-1}) : K \times \varphi(U) \to \mathbb{R}$ is a C^t map, $1 \leq t \leq \omega$, our claim follows by Lemma 14.1 or 14.3. □

Next suppose that we are given a C^t map

$$\zeta : K \times M \to \mathbb{R}^n, \tag{iii}$$

where $n \geq 1$ and $1 \leq t \leq \omega$. We define

$$\hat{\mathrm{A}}(\zeta) : M \to \mathbb{R}^n \tag{iv}$$

by

$$\hat{\mathrm{A}}(\zeta)(x) = \int_K \zeta(k,x)dk \text{ , for each } x \in M.$$

Here the integral is the Haar integral, obtained by integrating each coordinate function of ζ. That is if $\zeta(k,x) = (\zeta_1(k,x), \ldots, \zeta_n(k,x))$, for every $(k,x) \in K \times M$, then $\hat{\mathrm{A}}(\zeta)(x) = (\hat{\mathrm{A}}(\zeta_1)(x), \ldots, \hat{\mathrm{A}}(\zeta_n)(x))$. Thus we have, by applying Proposition 14.4 to each coordinate function of $\zeta : K \times M \to \mathbb{R}^n$, that the following holds.

Corollary 14.5 *Let K be a compact Lie group, and let M be any C^t manifold, where $1 \leq t \leq \omega$. Suppose $\zeta : K \times M \to \mathbb{R}^n$ is a C^t map, $1 \leq t \leq \omega$. Then $\hat{\mathrm{A}}(\zeta) : M \to \mathbb{R}^n$ is also a C^t map, $1 \leq t \leq \omega$. That is we have a map of sets*

$$\hat{\mathrm{A}} : \mathrm{C}^t(K \times M, \mathbb{R}^n) \to \mathrm{C}^t(M, \mathbb{R}^n) \text{ , } \zeta \mapsto \hat{\mathrm{A}}(\zeta),$$

for each compact Lie group K and any C^t manifold M, where $1 \leq t \leq \omega$. □

We call the map $\hat{\mathrm{A}}$ in Corollary 14.5 the *outer averaging map.*

Let us now consider the situation, where K is a compact Lie group and M is a C^t K-manifold, where $1 \leq t \leq \omega$. We denote the given C^t action of K on M by

$$\Phi\colon K \times M \to M, \ (k,x) \mapsto kx. \tag{v}$$

Suppose furthermore that $\mathbb{R}^n(\theta)$ is a linear representation space for K, and let

$$\Theta\colon K \times \mathbb{R}^n(\theta) \to \mathbb{R}^n(\theta), \ (k,y) \mapsto \theta(k)y = ky, \tag{vi}$$

denote the corresponding action of K on $\mathbb{R}^n(\theta)$. Note that Θ is a real analytic action.

Now suppose

$$f\colon M \to \mathbb{R}^n(\theta) \tag{vii}$$

is any C^t map, where $1 \leq t \leq \omega$. We define

$$\mathrm{A}(f)\colon M \to \mathbb{R}^n(\theta) \tag{viii}$$

by

$$\mathrm{A}(f)(x) = \int_K \theta(k^{-1})f(kx)dk = \int_K k^{-1}f(kx)dk, \ \ x \in M.$$

Proposition 14.6 *Let M be a C^t K-manifold, where K is a compact Lie group and $1 \leq t \leq \omega$, and let $\mathbb{R}^n(\theta)$ be a linear representation space for K. Suppose $f\colon M \to \mathbb{R}^n(\theta)$ is any C^t map, $1 \leq t \leq \omega$. Then $\mathrm{A}(f)\colon M \to \mathbb{R}^n(\theta)$ is a K-equivariant C^t map, $1 \leq t \leq \omega$. Furthermore $\mathrm{A}(f) = f$, if f is K-equivariant. In short we have a retraction map, of sets,*

$$\mathrm{A}\colon \mathrm{C}^t(M, \mathbb{R}^n(\theta)) \to \mathrm{C}^{t,K}(M, \mathbb{R}^n(\theta)), \ \ f \mapsto \mathrm{A}(f), \ \ 1 \leq t \leq \omega.$$

Proof We define $f_{(\Phi,\Theta)}$ to be the composite map

$$f_{(\Phi,\Theta)} = \Theta \circ (\iota \times \mathrm{id}) \circ (\mathrm{id} \times f) \circ \ \Phi_\Delta : K \times M \to \mathbb{R}^n(\theta).$$

Here $\Phi_\Delta : K \times M \to K \times M,\ (k,x) \mapsto (k, \Phi(k,x)) = (k, kx)$, where Φ is as in (v). Furthermore Θ is as in (vi), and $\iota : K \to K\ ,\ k \mapsto k^{-1},$. Since the action $\Phi : K \times M \to M$ is a C^t map, it follows that Φ_Δ is a C^t map, $1 \le t \le \omega$. The maps ι and Θ are both real analytic. Thus we see that if $f : M \to \mathbb{R}^n(\theta)$ is a C^t map, then $f_{(\Phi,\Theta)} : K \times M \to \mathbb{R}^n(\theta)$ is also a C^t map, $1 \le t \le \omega$. Since $f_{(\Phi,\Theta)}(k,x) = k^{-1} f(kx)$, for all $(k,x) \in K \times M$, we have that

$$\mathrm{A}(f) = \hat{\mathrm{A}}(f_{\Phi,\Theta}) : M \to \mathbb{R}^n(\theta), \tag{1}$$

and hence $\mathrm{A}(f)$ is a C^t map, $1 \le t \le \omega$, by Corollary 14.5. Furthermore $\mathrm{A}(f) : M \to \mathbb{R}^n(\theta)$ is K-equivariant, since if $k' \in K$ is fixed, then

$$\mathrm{A}(f)(k'x) = \int_K k^{-1} f(kk'x)dk = \int_K k'(kk')^{-1} f(kk'x)dk$$
$$= k' \int_K (kk')^{-1} f(kk'x)dk = k' \int_K k^{-1} f(kx)dk = k'\mathrm{A}(f)(x),$$

for all $x \in M$. Thus we have now shown that if $f \in \mathrm{C}^t(M, \mathbb{R}^n(\theta))$, then

$$\mathrm{A}(f) = \hat{\mathrm{A}}(f_{\Phi,\Theta}) \in \mathrm{C}^{t,K}(M, \mathbb{R}^n(\theta)), \tag{2}$$

$1 \le t \le \omega$. It only remains to show that $\mathrm{A}(f)\big|\mathrm{C}^{t,K}(M, \mathbb{R}^n(\theta)) = \mathrm{id}$. Suppose that $f : M \to \mathbb{R}^n(\theta)$ is K-equivariant. Then, for all $x \in M$,

$$\mathrm{A}(f)(x) = \int_K k^{-1} f(kx)dk = \int_K k^{-1} k f(x) dk = \int_K f(x)dk = f(x) \int_K 1dk = f(x).$$

This completes the proof. □

We call the map A in Proposition 14.6 the *inner averaging map.*

Notes The exposition in Section 14 follows to some extent the one in [21], Section 5. For the real analytic part of Proposition 14.6, see Theorem 1.16 in [25].

15 Continuity of the averaging maps in the strong C^r topologies, $1 \le r < \infty$, and in the very-strong C^∞ topology

Let K be a compact Lie group, and let M be a paracompact C^t manifold, $1 \le t \le \omega$. By Corollary 14.5 we know that the outer averaging map $\hat{\mathrm{A}}$ gives us a map of sets

$$\hat{\mathrm{A}} \colon \mathrm{C}^t(K \times M, \mathbb{R}^n) \to \mathrm{C}^t(M, \mathbb{R}^n), \quad \zeta \mapsto \hat{\mathrm{A}}(\zeta)$$

$1 \le t \le \omega$. We shall now prove that $\hat{\mathrm{A}}$ is continuous, see Corollary 15.3. That is, $\hat{\mathrm{A}}$ is continuous in the strong C^t topology, for $1 \le t < \infty$, and for $t = \infty$ we prove that $\hat{\mathrm{A}}$ is continuous in the very-strong C^∞ topology. Since the topology on $\mathrm{C}^\omega(\cdot,\cdot)$ is the relative topology from $\mathrm{C}^\infty(\cdot,\cdot)$, equipped with the very-strong C^∞ topology, there is nothing further to prove in the C^ω case, it follows directly from the C^∞ case. We begin with the following lemma.

Lemma 15.1 *Let M be a C^s manifold, where $1 \leq s \leq \infty$, and let $1 \leq r < \infty$ be such that $r \leq s$. Let K be a compact Lie group. Suppose $\zeta \in \mathrm{C}^s(K \times M, \mathbb{R})$, and let $\mathcal{P} = \mathcal{N}^{s,r}(\hat{\mathrm{A}}(\zeta); B, (U, \varphi), \varepsilon)$ be any elementary C^r neighborhood of $\hat{\mathrm{A}}(\zeta)$ in $\mathrm{C}^s(M, \mathbb{R})$. Then there exist finitely many elementary C^r neighborhoods $\mathcal{N}_j = \mathcal{N}^{s,r}(\zeta; E_j \times B, (V_j \times U, \psi_j \times \varphi), \varepsilon)$ of ζ in $\mathrm{C}^s(K \times M, \mathbb{R})$, $1 \leq j \leq q$, such that $\hat{\mathrm{A}}(\bigcap_{j=1}^q \mathcal{N}_j) \subset \mathcal{P}$.*

Proof In the same way as in Lemma 10.1 we see that it is enough to prove Lemma 15.1 in the case when $s = r$, and $1 \leq r < \infty$. So let $s = r$, and assume that $1 \leq r < \infty$.

Since K is compact we can find finitely many charts (V_j, ψ_j) in K and compact subsets E_j of V_j, $1 \leq j \leq q$, such that $\bigcup_{j=1}^q E_j = K$. We set

$$\mathcal{N}_j = \mathcal{N}^r(\zeta; E_j \times B,\ (V_j \times U, \psi_j \times \varphi),\ \varepsilon),\ \ 1 \leq j \leq q.$$

If $\eta \in \mathcal{N}_j$, where $1 \leq j \leq q$, then

$$|\mathrm{D}^\gamma(\eta - \zeta)(\psi_j^{-1}(y), \varphi^{-1}(x))| < \varepsilon,$$

for every $(y, x) \in \psi_j(E_j) \times \varphi(B) \subset \psi_j(V_j) \times \varphi(U) \subset \mathbb{R}^l \times \mathbb{R}^m$, and all $(l+m)$-tuples $\gamma = (\beta_1, \ldots, \beta_l, \alpha_1, \ldots, \alpha_m)$, with $|\gamma| \leq r$. Here $l = \dim K$ and $m = \dim M$. In particular

$$|\mathrm{D}^\alpha(\eta - \zeta)(k, \varphi^{-1}(x))| < \varepsilon, \tag{1}$$

for each $k \in E_j$, and every $x \in \varphi(B)$, and every m-tuple $\alpha = (\alpha_1, \ldots, \alpha_m)$, with $|\alpha| \leq r$.

If $\eta \in \bigcap_{j=1}^q \mathcal{N}_j$, then (1) holds for all $k \in \bigcup_{j=1}^q E_j = K$, and every $x \in \varphi(B)$, and m-tuple α with $|\alpha| \leq r$. Thus if $\eta \in \bigcap_{j=1}^q \mathcal{N}_j$ then, for every $x \in \varphi(B)$ and every $\alpha = (\alpha_1, \ldots, \alpha_m)$ with $|\alpha| \leq r$, the following holds,

$$\begin{aligned}\left|\mathrm{D}^\alpha(\hat{\mathrm{A}}(\eta) - \hat{\mathrm{A}}(\zeta))(\varphi^{-1}(x))\right| &= \left|\mathrm{D}^\alpha \int_K (\eta - \zeta)(k, \varphi^{-1}(x))dk\right| \\ = \left|\int_K \mathrm{D}^\alpha(\eta - \zeta)(k, \varphi^{-1}(x))dk\right| &\leq \int_K \left|\mathrm{D}^\alpha(\eta - \zeta)(k, \varphi^{-1}(x))dk\right| < \int_K \varepsilon dk = \varepsilon\end{aligned}$$

Thus $\hat{\mathrm{A}}(\eta) \in \mathcal{N}^r(\hat{\mathrm{A}}(\zeta); B, (U, \varphi), \varepsilon) = \mathcal{P}$, and this completes the proof. □

Proposition 15.2 *Let M be a paracompact C^s manifold, where $1 \leq s \leq \infty$, and let K be a compact Lie group. Then the map $\hat{\mathrm{A}}\colon \mathrm{C}^s(K \times M, \mathbb{R}) \to \mathrm{C}^s(M, \mathbb{R}),\ \zeta \mapsto \hat{\mathrm{A}}(\zeta)$, is continuous, $1 \leq s \leq \infty$.*

Proof We shall first give the proof in the case $s = \infty$. Let $\zeta \in \mathrm{C}^\infty(K \times M, \mathbb{R})$ and let $\mathcal{W} = \bigcap_{i \in \Lambda} \mathcal{P}_i$ be any basic very-strong C^∞ neighborhood of $\hat{\mathrm{A}}(\zeta)$ in $\mathrm{C}^\infty(M, \mathbb{R})$. Here each $\mathcal{P}_i = \mathcal{N}^{\infty, r_i}(\hat{\mathrm{A}}(\zeta); B_i, (U_i, \varphi_i), \varepsilon_i),\ i \in \Lambda$, is an elementary C^{r_i} neighborhood of $\hat{\mathrm{A}}(\zeta)$ in $\mathrm{C}^\infty(M, \mathbb{R})$, where $1 \leq r_i < \infty$, for $i \in \Lambda$, and the family $\{B_i\}_{i \in \Lambda}$ is locally finite in M. By Lemma 15.1 we find for each $\mathcal{P}_i,\ i \in \Lambda$, finitely many elementary C^{r_i} neighborhoods $\mathcal{N}_{i,j}$ of ζ in $\mathrm{C}^\infty(K \times M, \mathbb{R})$, $1 \leq j \leq q(i)$, such that $\hat{\mathrm{A}}(\bigcap_{j=1}^{q(i)} \mathcal{N}_{i,j}) \subset \mathcal{P}_i$. Here $\mathcal{N}_{i,j} = \mathcal{N}^{\infty, r_i}(\zeta; E_j \times B_i, (V_j \times U_i, \psi_j \times \varphi_i), \varepsilon_i)$, $i \in \Lambda$, $1 \leq j \leq q(i)$, and $\bigcup_{j=1}^{q(i)} E_j = K$. Now the family $\{E_j \times B_i \mid 1 \leq j \leq q(i),\ i \in \Lambda\}$ is locally finite in $K \times M$, and hence $\mathcal{V} = \bigcap_{i \in \Lambda} \bigcap_{j=1}^{q(i)} \mathcal{N}_{i,j}$ is a basic very-strong C^∞ neighborhood of ζ in $\mathrm{C}^\infty(K \times M, \mathbb{R})$. Furthermore we have that $\hat{\mathrm{A}}(\mathcal{V}) \subset \bigcap_{i \in \Lambda} \hat{\mathrm{A}}(\bigcap_{j=1}^{q(i)} \mathcal{N}_{i,j}) \subset \bigcap_{i \in \Lambda} \mathcal{P}_i =$

$\mathcal{W}$. This completes the proof in the case when $s = \infty$. In the case when $1 \le s < \infty$, the proof is completely similar. In this case we have that $r_i = s$, for all $i \in \Lambda$. □

Corollary 15.3 *Let M and K be as in Proposition 15.2, and let $n \ge 1$. Then the map $\hat{\mathrm{A}} : \mathrm{C}^s(K \times M, \mathbb{R}^n) \to \mathrm{C}^s(M, \mathbb{R}^n)$, is continuous, $1 \le s \le \infty$.*

Proof Follows from Proposition 15.2, by applying the product theorem, Proposition 11.1, to both $\mathrm{C}^s(K \times M, \mathbb{R}^n)$ and $\mathrm{C}^s(M, \mathbb{R}^n)$. □

Now suppose that M is a paracompact C^t K-manifold, where K is compact Lie group and $1 \le t \le \omega$, and let $\mathbb{R}^n(\theta)$ be a linear representation space for K. We know, by Proposition 14.6, that the inner averaging map A gives us a retraction map, of sets,

$$\mathrm{A} : \mathrm{C}^t(M, \mathbb{R}^n(\theta)) \to \mathrm{C}^{t,K}(M, \mathbb{R}^n(\theta)) \ , \ f \mapsto \mathrm{A}(f),$$

where $1 \le t \le \omega$. We shall prove that A is continuous, see Theorem 15.4. That is, A is continuous in the strong C^t topology for $1 \le t < \infty$, and in the very-strong C^∞ topology for $t = \infty$. In the case $t = \omega$ the topology on $\mathrm{C}^\omega(\cdot,\cdot)$ is the relative topology from $\mathrm{C}^\infty(\cdot,\cdot)$, so there is nothing further to prove in this case.

Theorem 15.4 *Let M be a paracompact C^s K-manifold, where K is a compact Lie group and $1 \le s \le \infty$, and let $\mathbb{R}^n(\theta)$ be a linear representation space for K. Then the map*

$$\mathrm{A}\colon \mathrm{C}^s(M, \mathbb{R}^n(\theta)) \to \mathrm{C}^{s,K}(M, \mathbb{R}^n(\theta)), \quad f \mapsto \mathrm{A}(f),$$

is continuous, $1 \le s \le \infty$.

Proof By (2) in the proof of Proposition 14.6, we know that if $f \in \mathrm{C}^s(M, \mathbb{R}^n(\theta))$, $1 \le s \le \infty$, then

$$\mathrm{A}(f) = \hat{\mathrm{A}}(f_{\Phi,\Theta}) \in \mathrm{C}^{s,K}(M, \mathbb{R}^n(\theta)).$$

Here $f_{(\Phi,\Theta)} = \Theta \circ (\iota \times \mathrm{id}) \circ (\mathrm{id} \ \times f) \circ \Phi_\triangle \colon K \times M \to \mathbb{R}^n(\theta)$, where

$$\Phi_\triangle \colon K \times M \to K \times M, \ (k, x) \mapsto (k, kx),$$

$$\iota : K \to K \ , \ k \mapsto k^{-1},$$

$$\Theta\colon K \times \mathbb{R}^n(\theta) \to \mathbb{R}^n(\theta), \ (k, y) \mapsto ky.$$

Hence the map $\mathrm{A} : \mathrm{C}^s(M, \mathbb{R}^n(\theta)) \to \mathrm{C}^{s,K}(M, \mathbb{R}^n(\theta)) \hookrightarrow \mathrm{C}^s(M, \mathbb{R}^n(\theta))$ equals the composite map

$$\mathrm{C}^s(M, \mathbb{R}^n(\theta)) \xrightarrow{\chi} \mathrm{C}^s(K \times M, K \times \mathbb{R}^n(\theta)) \xrightarrow{\Phi_\triangle^*} \mathrm{C}^s(K \times M, K \times \mathbb{R}^n(\theta))$$

$$\xrightarrow{(\iota \times \mathrm{id})_*} \mathrm{C}^s(K \times M, K \times \mathbb{R}^n(\theta)) \xrightarrow{\Theta_*} \mathrm{C}^s(K \times M, \mathbb{R}^n(\theta)) \xrightarrow{\hat{\mathrm{A}}} \mathrm{C}^s(M, \mathbb{R}^n(\theta))$$

Here $\chi(f) = \mathrm{id} \times f$, and χ is continuous by Corollary 11.2. The map $\Phi_\triangle$ is easily seen to be proper, and hence $\Phi_\triangle^*$ is continuous by Proposition 10.5. Furthermore $(\iota \times \mathrm{id})_*$ and Θ_* are continuous by Proposition 10.4, and $\hat{\mathrm{A}}$ is continuous by Corollary 15.3. □

Theorem 15.5 *Let M be a paracompact real analytic K-manifold, where K is a compact Lie group, and let $\mathbb{R}^n(\theta)$ be a linear representation space for K. Then the set* $\mathrm{C}^{\omega,K}(M,\mathbb{R}^n(\theta))$ *is dense in the space* $\mathrm{C}^{s,K}(M,\mathbb{R}^n(\theta))$, $1 \leq s \leq \infty$.

Proof Let $\mathcal{U}$ be a non-empty, open subset of $\mathrm{C}^{s,K}(M,\mathbb{R}^n(\theta))$, where $1 \leq s \leq \infty$. By Theorem 15.4 we know that $\mathrm{A}^{-1}(\mathcal{U})$ is an open subset of $\mathrm{C}^s(M,\mathbb{R}^n(\theta))$, and since A is surjective $\mathrm{A}^{-1}(\mathcal{U})$ is non-empty. Hence we have by Theorem 13.4 that there exists a real analytic map $f\colon M \to \mathbb{R}^n(\theta)$, such that $f \in \mathrm{A}^{-1}(\mathcal{U})$. Since $f\colon M \to \mathbb{R}^n(\theta)$ is real analytic, we have Proposition 14.6 that $\mathrm{A}(f)\colon M \to \mathbb{R}^n(\theta)$ is real analytic K-equivariant map. Thus $\mathrm{A}(f) \in \mathcal{U} \cap \mathrm{C}^{\omega,K}(M,\mathbb{R}^n(\theta))$, and this completes the proof. □

Notes The C^∞ cases, i.e., the very-strong C^∞ topology cases, of Theorems 15.4 and 15.5 equal Theorem 6.4 and Proposition 7.1, respectively, in [21].

16 Approximation of K-equivariant C^s maps, $1 \leq s \leq \infty$, by K-equivariant real analytic maps, in the strong C^s topologies, $1 \leq s < \infty$, and in the very-strong C^∞ topology

As in previous sections we shall also in this section use K to denote a compact Lie group. We will use the K-equivariant C^∞ imbedding result, Proposition 16.1 below in the proof of Theorem 16.3. Proposition 16.1 also holds in the C^r cases, $1 \leq r < \infty$, but we have chosen to give it here only in the C^∞ case for two reasons. First of all we only need the C^∞ case, and the other reason is that we do not know of any good reference for the C^r cases, $1 \leq r < \infty$, when M is non-compact.

Proposition 16.1 *Let M be a second countable C^∞ K-manifold, where K is a compact Lie group, and assume that the number of K-isotropy types in M is finite. Then there exist a linear representation space $\mathbb{R}^v(\lambda)$ for K and a K-equivariant closed C^∞ imbedding $j : M \to \mathbb{R}^v(\lambda)$.*

Proof See [53], §1. □

Proposition 16.2 *Let M be a paracompact C^s K-manifold, with or without boundary, and let N be a paracompact C^s K-manifold, $1 \leq s \leq \infty$. Then the set $\mathrm{Imb}^{s,K}_{\mathrm{c}}(M,N)$, of all K-equivariant closed C^s imbeddings of M into N, is an open subset of $\mathrm{C}^{s,K}(M,N)$.*

Proof It is a standard and well-known result that the set $\mathrm{Imb}^s_{\mathrm{c}}(M,N)$, of all closed C^s imbeddings of M into N, is open in $\mathrm{C}^s(M,N)$, with the strong C^s topology, $1 \leq s \leq \infty$, see e.g. [15], Corollary 2.1.6. In the case when $s = \infty$ the very-strong C^∞ topology on $\mathrm{C}^\infty(M,N)$ is at least as fine as the strong C^∞ topology on $\mathrm{C}^\infty(M,N)$), see (vii) in Section 9. Hence $\mathrm{Imb}^{s,K}_{\mathrm{c}}(M,N) = \mathrm{Imb}^s_{\mathrm{c}}(M,N) \cap \mathrm{C}^{s,K}(M,N)$ is open in $\mathrm{C}^{s,K}(M,N)$, $1 \leq s \leq \infty$. (Recall our CONVENTION in Section 10.) □

Theorem 16.3 *Let M be a second countable real analytic K-manifold, where K is a compact Lie group, and assume that the number of K-isotropy types in M is finite. Then there exist a linear representation space $\mathbb{R}^v(\lambda)$ for K and a K-equivariant real analytic closed imbedding $h : M \to \mathbb{R}^v(\lambda)$.*

Proof By Proposition 16.1 there exist a linear representation space $\mathbb{R}^v(\lambda)$ and a K-equivariant closed C^∞-imbedding $j : M \to \mathbb{R}^v(\lambda)$. We have by Proposition 16.2 that $\mathrm{Imb}^{\infty,K}_{\mathrm{c}}(M,\mathbb{R}^v(\lambda))$ is an open subset of $\mathrm{C}^{\infty,K}(M,\mathbb{R}^v(\lambda))$, and it is non-empty since

$j \in \mathrm{Imb}_{\mathrm{c}}^{\infty,K}(M, \mathbb{R}^v(\lambda))$. Hence there exists, by Theorem 15.5, a K-equivariant real analytic map $h : M \to \mathbb{R}^v(\lambda)$, such that $h \in \mathrm{Imb}_{\mathrm{c}}^{\infty,K}(M, \mathbb{R}^v(\lambda))$. Thus $h\colon M \to \mathbb{R}^v(\lambda)$ is a K-equivariant real analytic closed imbedding. □

Theorem 16.4 *Let M be a paracompact real analytic K-manifold, where K is a compact Lie group. Then M has a K-invariant real analytic Riemannian metric.*

Proof Suppose M_0 is a second countable real analytic manifold. By the Grauert-Morrey imbedding theorem, see Theorem 13.3, we may consider M_0 as a real analytic closed submanifold of some euclidean space $\mathbb{R}^u$. Hence the standard Riemannian metric on $\mathbb{R}^u$ induces a real analytic Riemannian metric on M_0.

Now let M be a paracompact real analytic K-manifold. Then each connected component M_i of M, $i \in \Lambda$, is a second countable real analytic manifold, see Proposition 2.1. Thus each M_i , $i \in \Lambda$, has a real analytic Riemannian metric, by the above, and this gives M a real analytic Riemannian metric

$$\Xi\colon \mathrm{T}M \oplus \mathrm{T}M \to \mathbb{R}$$

Here $\mathrm{T}M$ denotes the tangent bundle of M. The Whitney sum $\mathrm{T}M \oplus \mathrm{T}M$ is a real analytic K-manifold, and Ξ is a real analytic map, such that for each $x \in M$ the restriction $\Xi\big|\colon \mathrm{T}_xM \oplus \mathrm{T}_xM \to \mathbb{R}$ is an inner product on the tangent space T_xM. Now let

$$\mathrm{A}(\Xi) : \mathrm{T}M \oplus \mathrm{T}M \to \mathbb{R}$$

be as in Section 14. That is, for each $(a, b) \in \mathrm{T}M \oplus \mathrm{T}M$, we have that

$$\mathrm{A}(\Xi)(a, b) = \int_K \Xi(ka, kb)dk.$$

By Proposition 14.6 we know that $\mathrm{A}(\Xi)\colon \mathrm{T}M \oplus \mathrm{T}M \to \mathbb{R}$ is a K-invariant real analytic map, and it is straightforward to verify that for each $x \in M$ the restriction $\mathrm{A}(\Xi)\big|\colon \mathrm{T}_xM \oplus \mathrm{T}_xM \to \mathbb{R}$ is an inner product on T_xM. Thus $\mathrm{A}(\Xi)\colon \mathrm{T}M \oplus \mathrm{T}M \to \mathbb{R}$ is a K-invariant real analytic Riemannian metric on M. □

Theorem 16.4 is used in the proof of Theorem 16.5 below.

Theorem 16.5 *Let K be a compact Lie group, and let M be a paracompact real analytic K-manifold and N a K-invariant real analytic closed submanifold of M. Then there exist a K-invariant open neighborhood V of N in M, and a K-equivariant real analytic retraction $p\colon V \to N$.*

Proof See e.g. Theorem I in [23], where this is proved in a more general situation than we here have. □

Theorem 16.6 *Let M and N be paracompact real analytic K-manifolds, where K is a compact Lie group, and assume that the number of K-isotropy types in N is finite. Then the set $\mathrm{C}^{\omega,K}(M, N)$ is dense in the space $\mathrm{C}^{s,K}(M, N)$, $1 \leq s \leq \infty$.*

Proof It follows by Remark 10.8 that it is enough to prove Theorem 16.6 in the case when M and N are second countable. In this case we may, by Theorem 16.3, consider N as a K-invariant real analytic closed submanifold of some linear representation space $\mathbb{R}^v(\lambda)$ for K. By Theorem 16.5 we find a K-invariant open neighborhood V of N in $\mathbb{R}^v(\lambda)$ and a K-equivariant real analytic retraction $p\colon V \to N$. By Proposition 10.4 the induced map

$$p_*\colon \mathrm{C}^{s,K}(M, V) \to \mathrm{C}^{s,K}(M, N), \;\; f \mapsto p \circ f,$$

is continuous, for each $1 \leq s \leq \infty$, and moreover p_* is surjective.

Let $\mathcal{U}$ be a non-empty open subset of $\mathrm{C}^{s,K}(M,N)$. Then $p_*^{-1}(\mathcal{U})$ is a non-empty set, and it is open in $\mathrm{C}^{s,K}(M,V)$, and hence also open in $\mathrm{C}^{s,K}(M,\mathbb{R}^v(\lambda))$, see Lemma 9.5. By Theorem 15.5 there exists a real analytic K-equivariant map $f\colon M \to \mathbb{R}^v(\lambda)$ such that $f \in p_*^{-1}(\mathcal{U})$. Then $f\colon M \to V \subset \mathbb{R}^v(\lambda)$, and $p \circ f\colon M \to N$ is a K-equivariant real analytic map. Now $p \circ f = p_*(f) \in \mathcal{U} \cap \mathrm{C}^{\omega,K}(M,N)$, and this completes the proof. □

Notes Theorem 7.2 in [21] is the very-strong C^∞ topology case of Theorem 16.6. A strong C^∞ topology version of Theorem 16.6 is proved in Theorem 1.2 in [34]. For Theorem 16.4, see Theorem 1.17 in [25].

17 Approximation of C^s K-slices, $1 \leq s \leq \infty$

In Lemma 17.1 below G denotes a Lie group and K is a compact subgroup of G. By U^* we denote a K-invariant open neighborhood of eK in G/K such that there is a K-equivariant real analytic cross section

$$\sigma\colon U^* \to G$$

with $\sigma(eK) = e$. Thus $\sigma(ku) = k\sigma(u)k^{-1}$, for every $k \in K$ and each $u \in U^*$. Moreover we choose U^* such that there is a K-equivariant real analytic isomorphism

$$h\colon \mathbb{R}^d(\tau) \to U^*,$$

where $\mathbb{R}^d(\tau)$ is an orthogonal representation space for K, see Lemma 6.3.

Lemma 17.1 here below is Lemma 4.1 in [17]. It is a C^s, $1 \leq s \leq \infty$, version of the corresponding Lemma 6.1 in [16], which is given only in C^∞ case. The proof of Lemma 17.1, although a more general result, is somewhat shorter than that of Lemma 6.1 in [16].

Lemma 17.1 *Let the notation be as above, and let M be a paracompact C^s G-manifold, and let P be a paracompact connected C^s K-manifold, where $1 \leq s \leq \infty$. Suppose $i\colon P \to M$ is a K-equivariant C^s imbedding such that $i(P) = S$ is a C^s K-slice in M, $1 \leq s \leq \infty$. Then there exist K-invariant open neighborhoods U and U_1 of eK in G/K, where $U \subset U_1 \subset U^*$, such that if we denote $W = \sigma(U)S$ then the following holds. There exists an open neighborhood $\mathcal{W}$ of $i\colon P \to W$ in $\mathrm{C}^{s,K}(P,W)$ such that if $j \in \mathcal{W}$ then $j\colon P \to W$ is a K-equivariant C^s imbedding, $1 \leq s \leq \infty$, with the following properties:*

(a) *$S' = j(P)$ is a C^s K-slice in M,*

(b) *$GS' = GS$,*

(c) *$W = \sigma(U)S \subset \sigma(U_1)S'$.*

Proof Let $\sigma\colon U^* \to G$ and $h\colon \mathbb{R}^d(\tau) \to U^*$ be as above. If we let K act diagonally on $U^* \times P$, then $\gamma(i)\colon U^* \times P \to GS$, $(u,x) \mapsto \sigma(u)i(x)$, is a K-equivariant C^s-imbedding, and its image $\sigma(U^*)S = W^*$ is a K-invariant open subset of GS, see Remark 8.8. We set

$$U_1 = h(\mathring{\mathrm{D}}^d(\tau)) .$$

Here $\mathring{\mathrm{D}}^d(\tau)$ denotes the open unit disk in the orthogonal representation space $\mathbb{R}^d(\tau)$ for K, and $\mathrm{D}^d(\tau)$ denotes the corresponding closed unit disk. Then U_1 is a K-invariant open neighborhood of eK in G/K, and $\overline{U}_1 = h(\mathrm{D}^d(\tau))$ is a K-invariant real analytic compact manifold with boundary. We note that

$$\gamma(i)\colon \overline{U}_1 \times P \to GS,\ (u, x) \mapsto \sigma(u)i(x) \tag{1}$$

is a closed imbedding into GS.

Let $\pi\colon G \to G/K$ denote the natural projection. Then $\pi^{-1}(U_1)$ is an open neighborhood of K in G, and hence there exists, by Lemma 4.6, a K-invariant open neighborhood U of eK in G/K such that $G[U] \subset \pi^{-1}(U_1)$. By using (ix) in Section 4 we obtain that

$$G[U] = \pi^{-1}(U)(\pi^{-1}(U))^{-1} \subset \pi^{-1}(U_1)\ . \tag{2}$$

Clearly we may choose U so that $U = h(\mathring{\mathrm{D}}^d_\varepsilon(\tau))$, where $0 < \varepsilon < 1$, and then

$$U \subset \overline{U} \subset U_1 \subset \overline{U}_1 \subset U^* \subset G/K\ . \tag{3}$$

We set

$$W = \sigma(U)S,\ \text{and}\ W_1 = \sigma(U_1)S\ .$$

Then $\sigma(\overline{U})S = \overline{W}$ and $\sigma(\overline{U}_1)S = \overline{W}_1$, where the closures $\overline{W}$ and $\overline{W}_1$ are taken in GS, and

$$S \subset W \subset \overline{W} \subset W_1 \subset \overline{W}_1 \subset W^* \subset GS\ .$$

Let $p\colon GS \to G/K$ be the G-equivariant C^s-smooth map onto G/K determined by the C^s K-slice S, i.e., $p^{-1}(eK) = S$, see Lemma 8.2. Then $W = \sigma(U)S = p^{-1}(U)$, and hence $G[W] = G[U]$, by Lemma 4.1. Thus we have by (2) that

$$G[W] \subset \pi^{-1}(U_1)\ . \tag{4}$$

Since $\overline{U}_1$ is compact the map $\chi\colon \mathrm{C}^{s,K}(P, W) \to \mathrm{C}^{s,K}(\overline{U}_1 \times P, \overline{U}_1 \times W),\ j \mapsto \mathrm{id} \times j$, is continuous, by Corollary 11.2. The K-equivariant C^s map $\eta\colon \overline{U}_1 \times W \to GS,\ (u, y) \mapsto \sigma(u)y$, induces a continuous map $\eta_*\colon \mathrm{C}^{s,K}(\overline{U}_1 \times P, \overline{U}_1 \times W) \to \mathrm{C}^{s,K}(\overline{U}_1 \times P, GS)$, see Proposition 10.4, and hence

$$\Gamma = \eta_* \circ \chi\colon \mathrm{C}^{s,K}(P, W) \to \mathrm{C}^{s,K}(\overline{U}_1 \times P, GS)\ , \quad j \mapsto \eta \circ (\mathrm{id} \times j)\ ,$$

is continuous. If $j \in C^{s,K}(P, W)$ then $\Gamma(j)(u, x) = \sigma(u)j(x),$ for all $(u, x) \in \overline{U}_1 \times P$. In particular we have that $\Gamma(i) = \gamma(i)\colon \overline{U}_1 \times P \to GS$, where $\gamma(i)$ is as in (1), and hence $\Gamma(i)$ is a closed imbedding, i.e.,

$$\Gamma(i) \in \mathrm{Imb}^{s,K}_{\mathrm{c}}(\overline{U}_1 \times P, GS)\ .$$

Next we observe that the restriction map

$$\mathrm{res}\colon \mathrm{Imb}^{s,K}_{\mathrm{c}}(\overline{U}_1 \times P, GS) \to \mathrm{Imb}^{s,K}_{\mathrm{c}}(\partial\overline{U}_1 \times P, GS),\quad h \mapsto h|\ ,$$

is continuous. This follows by Proposition 10.5, since the inclusion map $\mathrm{incl}\colon \partial\overline{U}_1 \times P \to \overline{U}_1 \times P$ is proper, and $\mathrm{res} = (\mathrm{incl})^*$. By (3) we have that $\partial\overline{U}_1 \subset G/K - \overline{U}$, and hence

$\Gamma(i)(\partial\overline{U}_1 \times P) \subset GS - \overline{W}$. By Lemma 9.5 we can find an open neighborhood $\mathcal{N}_1$ of $\Gamma(i)|$ in $C^{s,K}(\partial\overline{U}_1 \times P, GS)$ such that if $h' \in \mathcal{N}_1$ then $\mathrm{Im}(h') \cap \overline{W} = \emptyset$. We set

$$\mathcal{M}_1 = (\mathrm{res})^{-1}(\mathcal{N}_1) .$$

Then $\mathcal{M}_1$ is open in $\mathrm{Imb}^{s,K}_{\mathrm{c}}(\overline{U}_1 \times P, GS)$, and hence also in $C^{s,K}(\overline{U}_1 \times P, GS)$, by Theorem 16.2. Thus

$$\mathcal{W} = \Gamma^{-1}(\mathcal{M}_1)$$

is an open neighborhood of i in $C^{s,K}(P, W)$.

If $j \in \mathcal{W}$ then $\Gamma(j)\colon \overline{U}_1 \times P \to GS$ is a K-equivariant C^s closed imbedding, with the additional property that $\Gamma(j)(\partial\overline{U}_1 \times P) \cap \overline{W} = \emptyset$. Hence $\Gamma(j)(\overline{U}_1 \times P) \cap \overline{W} = \Gamma(j)(U_1 \times P) \cap \overline{W}$ and therefore also

$$\Gamma(j)(\overline{U}_1 \times P) \cap W = \Gamma(j)(U_1 \times P) \cap W .$$

This shows that $\Gamma(j)(U_1 \times P) \cap W$ is closed in W. Furthermore $\Gamma(j)(U_1 \times P)$ is an open set, for example by invariance of domain (see e.g. [10], Theorem XI. 3.11), and thus $\Gamma(j)(U_1 \times P) \cap W$ is both open and closed in W. Since $W = \sigma(U)i(P)$ is connected it follows that $\Gamma(j)(U_1 \times P) \cap W = W$. Therefore $W \subset \Gamma(j)(U_1 \times P)$, that is

$$W \subset \sigma(U_1)j(P) .$$

Thus $GS \subset GW \subset G\sigma(U_1)j(P) = Gj(P) \subset GS$, and hence

$$GS = Gj(P) .$$

Thus we have shown that if $j \in \mathcal{W}$ then $S' = j(P)$ satisfies (b) and (c), and hence it only remains to show that $j(P)$ is a C^s K-slice in M. We already know that $\Gamma(j)\colon U_1 \times P \to \sigma(U_1)j(P)$, $(u, x) \mapsto \sigma(u)j(x)$, is a K-equivariant C^s diffeomorphism onto $\sigma(U_1)j(P)$. Thus $j(P)$ is a C^s near K-slice in M, and hence it is enough to show that

$$gj(P) \cap j(P) = \emptyset \text{ , for all } g \in G - K , \tag{5}$$

see Lemma 8.10. Suppose $g \in G$ is such that $gj(P) \cap j(P) \neq \emptyset$. Since $j(P) \subset W$ it follows that $gW \cap W \neq \emptyset$, and hence $g \in G[W]$. Thus (4) implies that $g \in \pi^{-1}(U_1)$. Therefore $\pi(g) \in U_1$, and $g = \sigma(\pi(g))k$, for some $k \in K$. Now $gj(P) = \sigma(\pi(g))kj(P) = \sigma(\pi(g))j(P) = \Gamma(j)(\{\pi(g)\} \times P)$. Since $\Gamma(j)\colon U_1 \times P \to GS$ is injective we have that $\Gamma(j)(\{u\} \times P)$ is disjoint from $\Gamma(j)(\{eK\} \times P) = j(P)$, for all $u \in U_1$, except $u = eK$. Thus we see that $gj(P) \cap j(P) \neq \emptyset$ implies that $\pi(g) = eK$, and hence $g \in K$. This proves that (5) holds, and completes the proof of the lemma. □

18 Proof of the main theorem

We are now ready to give the proof of the main theorem, Theorem 1.2. The proof is practically the same as the proof of Theorem 7.1 in [16], but we have now included all cases

$1 \le s \le \infty$, whereas in [16] the proof is given in the smooth, i.e., in the C^∞ case. Compared with the proof in [16] there are however two changes in the proof. We now, correctly, use the glueing lemma, Lemma 12.1, in its very-strong C^∞ topology form. Furthermore we now need to employ, in the C^∞ case, a result for the very-strong C^∞ topology concerning approximation of K-equivariant C^∞ maps by K-equivariant real analytic maps, where K denotes a compact Lie group, see Theorem 7.2 in [21], Theorem 16.6 in the present article.

Theorem 18.1 *Let M be a Cartan C^s G-manifold, where G is any Lie group and $1 \le s \le \infty$. Then there exists a real analytic structure β on M, compatible with the given C^s structure on M, such that the action of G on M_β is real analytic.*

Proof We define $\mathcal{B}$ to be the family of all pairs (B, β), where B is a non-empty G-invariant open subset of M and β is a real analytic structure on B, compatible with the given C^s structure on B, such that the action of G on $B_\beta = (B, \beta)$ is real analytic.

Let us first show that the family $\mathcal{B}$ is non-empty. This is seen as follows. Let $x_0 \in M$ and denote $G_{x_0} = K_0$. Then K_0 is a compact subgroup of G. By the slice theorem, Theorem 8.13, there exists a G-equivariant C^s diffeomorphism

$$\mu_0 : G \times_{K_0} \mathbb{R}^{q_0}(\rho_0) \to GS_0,$$

where $B_0 = GS_0$ is a G-invariant open neighborhood of x_0 in M, and $\mathbb{R}^{q_0}(\rho_0)$ denotes a linear representation space for K_0. As we saw in Section 7, Propositions 7.2 and 7.3, the twisted product $G \times_{K_0} \mathbb{R}^{q_0}(\rho_0)$ is a real analytic G-manifold, and we give B_0 the real analytic structure β_0 induced from $G \times_{K_0} \mathbb{R}^{q_0}(\rho_0)$ through μ_0^{-1}. Since μ_0 is a C^s diffeomorphism it follows that β_0 is compatible with the C^s structure on B_0. Since the action of G on $G \times_{K_0} \mathbb{R}^{q_0}(\rho_0)$ is real analytic and since μ_0 is G-equivariant, it follows that the action of G on $(B_0)_{\beta_0}$ is real analytic. Thus $(B_0, \beta_0) \in \mathcal{B}$, and we have shown that $\mathcal{B}$ is non-empty.

We define an order in $\mathcal{B}$ by setting

$$(B_1, \beta_1) \le (B_2, \beta_2)$$

if and only if:

(i) $B_1 \subset B_2$.

(ii) The real analytic structure β_1 on B_1 is the one induced from the real analytic structure β_2 on B_2.

Now suppose $\mathcal{C}$ is a chain in $\mathcal{B}$, i.e., if (B_1, β_1) and (B_2, β_2) belong to $\mathcal{C}$ then either $(B_1, \beta_1) \le (B_2, \beta_2)$ or $(B_2, \beta_2) \le (B_1, \beta_1)$. Let $\mathcal{C}_1$ denote the family of all B occurring as the first coordinate of a pair in $\mathcal{C}$, and let $\mathcal{C}_2$ be the family of all β occurring as the second coordinate of a pair in $\mathcal{C}$. Using this notation we define

$$B^* = \bigcup_{B \in \mathcal{C}_1} B, \quad \text{and} \quad \beta' = \bigcup_{\beta \in \mathcal{C}_2} \beta.$$

Then B^* is a non-empty G-invariant open subset of M, and β' is a real analytic atlas on B^*, compatible with the C^s structure on B^*. Let $\beta^* = \langle \beta' \rangle$ be the real analytic structure, i.e., the maximal real analytic atlas, on B^* generated by β'. Then the real analytic structure

β^* on B^* has the property that for each $(B, \beta) \in \mathcal{C}$ the real analytic structure that β^* induces on B equals β. It can now easily be seen that the action of G on $B^*_{\beta^*}$ is real analytic. For if $x \in B^*$, then there exists an open neighborhood U of x in B^* with $\overline{U}$ is compact and $\overline{U} \subset B^*$. Since $\overline{U}$ is compact it follows that there exists $(B, \beta) \in \mathcal{C}$, such that $U \subset \overline{U} \subset B$. Now $GU \subset B$ and the real analytic structure on GU induced from B_β equals the one induced from $B^*_{\beta^*}$. Since the action of G on B_β is real analytic it now follows that the action of G on $(GU)_{\beta^*}$ is real analytic, and hence we have shown that the action of G on $B^*_{\beta^*}$ is real analytic. Thus $(B^*, \beta^*) \in \mathcal{B}$, and furthermore we have that

$$(B, \beta) \leq (B^*, \beta^*), \text{ for all } (B, \beta) \in \mathcal{C}.$$

This shows that (B^*, β^*) is an upper bound for $\mathcal{C}$ in $\mathcal{B}$. Hence we obtain by Zorn's lemma that there exists a maximal element (B, β) in $\mathcal{B}$. We claim that $B = M$.

Suppose the contrary and assume that $B \subsetneq M$. If B is closed in M, then $M - B$ is a non-empty G-invariant open subset of M. In this case we could, as in the beginning of the proof, find a non-empty G-invariant open subset B_0, where $B_0 \subset M - B$, such that B_0 has a real analytic structure β_0, compatible with the C^s structure on B_0, and the action of G on $(B_0)_{\beta_0}$ is real analytic. Now $\beta \dot\cup \beta_0$ is a real analytic atlas on $B \dot\cup B_0$, which is compatible with the C^s structure on $B \dot\cup B_0$, and we let $\langle \beta \dot\cup \beta_0 \rangle$ denote the real analytic structure determined by $\beta \dot\cup \beta_0$. Then $\langle \beta \dot\cup \beta_0 \rangle$ is compatible with the C^s structure on $B \dot\cup B_0$, and it is also clear that the action of G on $(B \dot\cup B_0, \langle \beta \dot\cup \beta_0 \rangle)$ is real analytic. Thus $(B \dot\cup B_0, \langle \beta \dot\cup \beta_0 \rangle) \in \mathcal{B}$ and $(B, \beta) < (B \dot\cup B_0, \langle \beta \dot\cup \beta_0 \rangle)$, which contradicts the fact that (B, β) is a maximal element in $\mathcal{B}$. Thus B is not closed in M, and hence $\overline{B} - B \neq \emptyset$.

Let $x \in \overline{B} - B$, and denote $G_x = K$. By the C^s slice theorem, $1 \leq s \leq \infty$, Theorem 8.13, there exists a linear C^s slice S at x in M, and we let

$$i : \mathbb{R}^q(\rho) \xrightarrow{\cong} S \subset M$$

be a K-equivariant C^s imbedding into M such that $i(\mathbb{R}^q(\rho)) = S$, and $i(0) = x$. Here $\mathbb{R}^q(\rho)$ denotes a linear representation space for K. Then GS is open in M and

$$\mu(i) : G \times_K \mathbb{R}^q(\rho) \xrightarrow{\cong} GS, \quad [g, x] \mapsto gi(x), \tag{1}$$

is a G-equivariant C^s diffeomorphism.

We now choose a K-invariant product neighborhood W of S in M, cf. Remark 8.8, and an open neighborhood $\mathcal{W}$ of

$$i : \mathbb{R}^q(\rho) \to W \tag{2}$$

in $\mathrm{C}^{s,K}(\mathbb{R}^q(\rho), W)$, such that Lemma 17.1 holds for W and $\mathcal{W}$. Furthermore we may assume that the number of K-isotropy types occurring in W is finite, see Remark 8.9.

Since $x \in \overline{B} - B$ and GS is an open neighborhood of x in M it follows that

$$GS \cap B \neq \emptyset, \text{ and } GS \cap (M - B) \neq \emptyset. \tag{3}$$

Furthermore we have that $G(S \cap B) = GS \cap B$, since B in a G-invariant set, and hence $S \cap B \neq \emptyset$. Thus we see that $S \cap B$ is a non-empty K-invariant open subset of S, and therefore

$$V = i^{-1}(S \cap B)$$

is a non-empty K-invariant open subset of $\mathbb{R}^q(\rho)$, and $i(V) = S \cap B \subset W \cap B$.

The fact that V is a K-invariant open subset of $\mathbb{R}^q(\rho)$ implies in particular that V is a real analytic K-manifold. Furthermore, also $W \cap B_\beta$ is a real analytic K-manifold, since $W \cap B_\beta$ is a K-invariant open subset of the real analytic G-manifold B_β. We now consider the K-equivariant C^s imbedding

$$i|V : V \to W \cap B_\beta.$$

By Lemma 12.1 there exists an open neighborhood $\mathcal{N}$ of $i|V$ in $\mathrm{C}^{s,K}(V, W \cap B_\beta)$ such that we obtain a continuous map

$$E : \mathcal{N} \to \mathrm{C}^{s,K}(\mathbb{R}^q(\rho), W)$$

by defining for each $h \in \mathcal{N}$,

$$E(h)(x) = \begin{cases} h(x), & x \in V \\ i(x), & x \in \mathbb{R}^q(\rho) - V. \end{cases} \tag{4}$$

Observe that $E(i|V) = i$. Since E is continuous, and $\mathcal{W}$ is an open neighborhood of i in $\mathrm{C}^{s,K}(\mathbb{R}^q(\rho), W)$, there exists an open neighborhood $\mathcal{N}_1$ of $i|V$ in $\mathcal{N}$ such that

$$E(\mathcal{N}_1) \subset \mathcal{W}. \tag{5}$$

The number of K-isotropy types occurring in W is finite and hence the same holds for $W \cap B_\beta$. By Theorem 16.6 there exists

$$h_1 \in \mathcal{N}_1 \cap \mathrm{C}^{\omega,K}(V, B_\beta \cap W), \tag{6}$$

that is, there exists a K-equivariant real analytic map

$$h_1 : V \to W \cap B_\beta$$

such that $h_1 \in \mathcal{N}_1$. We now define

$$j = E(h_1) : \mathbb{R}^q(\rho) \to W. \tag{7}$$

Then we have by (5) and (6) that $j \in \mathcal{W}$. By the choice of $\mathcal{W}$, Lemma 17.1 holds for $\mathcal{W}$ and hence j is a K-equivariant C^s imbedding such that

$$j(\mathbb{R}^q(\rho)) = S'$$

is a C^s K-slice in M and $GS' = GS$. (In fact $j(0) = x$, so S' is a C^s slice at $x \in M$.) Hence

$$\mu(j) : G \times_K \mathbb{R}^q(\rho) \xrightarrow{\cong} GS' = GS, \quad [g, x] \mapsto gj(x), \tag{8}$$

is a G-equivariant C^s diffeomorphism.

We claim that the restriction

$$\mu(j)| : G \times_K V \to GS \cap B_\beta \tag{9}$$

is a G-equivariant real analytic isomorphism. First of all we claim that $\mu(j)(G \times_K V) = GS \cap B$. It follows directly from the definition (4) that the maps $j = E(h_1)$ in (7) and i in (2) agree on the set $\mathbb{R}^q(\rho) - V$. Hence the maps $\mu(j)$ in (8) and $\mu(i)$ in (1) agree on $G \times_K (\mathbb{R}^q(\rho) - V)$. Since $\mathrm{im}(\mu(j)) = GS' = GS = \mathrm{im}(\mu(i))$, and both $\mu(j)$ and $\mu(i)$ are bijective maps it now follows that $\mu(j)(G \times_K V) = \mu(i)(G \times_K V) = Gi(V) = G(S \cap B) = GS \cap B$. Thus we see that $\mu(j)|$ in (9) is a G-equivariant C^s diffeomorphism onto $GS \cap B_\beta$. Since $j|V = h_1 : V \to W \cap B_\beta \subset B_\beta$ is a K-equivariant real analytic map it follows by Lemma 7.4 that $\mu(j)|$ in (9) is a real analytic map. It now follows by the real analytic inverse function theorem (see e.g. Theorem 2.2.10 in [42]) that $(\mu(j)|)^{-1}$ is real analytic. This proves that $\mu(j)|$ in (9) is a G-equivariant real analytic isomorphism.

Let us now denote $GS = B_1$. Then B_1 is a G-invariant open subset of M. We give B_1 the real analytic structure β_1 induced from $G \times_K \mathbb{R}^q(\rho)$ through $\mu(j)^{-1}$, i.e., the real analytic structure β_1 for which $\mu(j) : G \times_K \mathbb{R}^q(\rho) \to (B_1)_{\beta_1}$ is a real analytic isomorphism. Since $\mu(j)$ in (8) is a C^s diffeomorphism it follows that the real analytic structure β_1 is compatible with the C^s structure on B_1. Since $\mu(j)$ is G-equivarint and the action of G on $G \times_K \mathbb{R}^q(\rho)$ is real analytic it follows that the action of G on $(B_1)_{\beta_1}$ is real analytic. The fact that $\mu(j)|$ in (9) is a real analytic isomorphism onto the open subset $B_1 \cap B_\beta$ of B_β implies that the real analytic structure on $B_1 \cap B$ induced from $(B_1)_{\beta_1}$ is the same as the one induced from B_β. Hence $\beta \cup \beta_1$ is a real analytic atlas on $B \cup B_1$, which is compatible with the C^s structure on $B \cup B_1$, and we let $\langle \beta \cup \beta_1 \rangle$ denote the real analytic structure determined by $\beta \cup \beta_1$. It is clear that the action of G on $(B \cup B_1, \langle \beta \cup \beta_1 \rangle)$ is a real analytic, and thus $(B \cup B_1, \langle \beta \cup \beta_1 \rangle) \in \mathcal{B}$. Furthermore we have by the latter part of (3) that $B \subsetneq B \cup B_1$, and hence $(B, \beta) < (B \cup B_1, \langle \beta \cup \beta_1 \rangle)$, but this contradicts the maximality of (B, β). Thus $B = M$. $\square$

Bibliography

[1] R. H. Bing: A homeomorphism between the 3-sphere and the sum of two solid horned spheres *Ann. of Math.* **56** (1952) 354–362

[2] G. Birkhoff: Lie groups simply isomorphic with no linear group *Bull. Amer. Math. Soc.* **42** (1936) 883–888

[3] G. Birkhoff: Analytical groups *Trans. Amer. Math. Soc.* **43** (1938) 61–101

[4] N. Bourbaki: *General Topology, Part 1* (Paris: Hermann and Reading, Mass.: Addison-Wesley, 1966)

[5] G. Bredon: *Introduction to compact transformation groups* (New York and London: Academic Press, 1972)

[6] J. Cerf: Topologie de certains espaces de plongements *Bull. Soc. Math. France* **89** (1961) 227–380

[7] C. Chevalley: Two theorems on solvable topological groups *In: Michigan Lectures in Topology* (1941) 291–292

[8] C. Chevalley: *Theory of Lie Groups: I* (Princeton, N.J.: Princeton University Press, 1946)

[9] J. Dieudonné: Une généralisation des espaces compacts *J. Math. Pures Appl.* **23** (1944) 65–76

[10] S. Eilenberg and N. Steenrod: *Foundations of Algebraic Topology* (Princeton, N.J.: Princeton University Press, 1952)

[11] A. Gleason: Groups without small subgroups *Ann. of Math.* **56** (1952) 193–212

[12] H. Grauert: On Levi's problem and the imbedding of real analytic manifolds *Ann. of Math.* **68** (1958) 460–472.

[13] S. Helgason: *Differential geometry and symmetric spaces* (New York and London: Academic Press, 1962)

[14] D. Hilbert: Mathematische Probleme *In: Nachr. Akad. Wiss.* (Göttingen, 1900) 253–297

[15] M. Hirsch: *Differential topology* (New York, Heidelberg, Berlin: Springer-Verlag, 1976)

[16] S. Illman: Every proper smooth action of a Lie group is equivalent to a real analytic action: a contribution to Hilbert's fifth problem *Ann. of Math. Stud.* **138** (1995) 189–220

[17] S. Illman: Smoothly equivalent real analytic proper G-manifolds are subanalytically equivalent *Math. Ann.* **306** (1996) 647–673

[18] S. Illman: Hilbert's fifth problem: a survey *In: Proceedings of the International Conference on Mathematics and Applications, dedicated to the 90th Anniversary of L.S. Pontryagin (Moscow State Lomonosov University, 1998)* (Moscow: Itogi Nauki Tehniki, Soveremennaya Matematika, VINITI edition, vol. 68, 1999) 55–64; (In Russian.) [English translation in: *J. Math. Sci.* **105** (2001), 1843–1847

[19] S. Illman: Existence and uniqueness of equivariant triangulations of smooth proper G-manifolds with some applications to equivariant Whitehead torsion *J. reine angew. Math.* **524** (2000) 129–183

[20] S. Illman: Hilbert's fifth problem and proper actions of Lie groups *In: Current Trends in Transformation Groups* (Kluveer Academic Publishers, K-Monographs in Mathematics, Vol.7, 2002) 1–23

[21] S. Illman: The very-strong C^∞ topology on $C^\infty(M,N)$ and K-equivariant maps *Osaka J. Math.* **40** (2003) 409–428

[22] S. Illman and M. Kankaanrinta: A new topology for the set $C^{\infty,G}(M,N)$ of G-equivariant smooth maps *Math. Ann.* **316** (2000) 139–168

[23] S. Illman and M. Kankaanrinta: Three basic results for real analytic proper G-manifolds *Math. Ann.* **316** (2000) 169–183

[24] K. Iwasawa: On some topological groups *Ann. of Math.* **50** (1949) 507–557

[25] M. Kankaanrinta: Proper real analytic actions of Lie groups on manifolds *Ann. Acad. Sci. Fenn., Ser A I Math. Dissertationes* **83** (1991) 1–41

[26] K. Kawakubo: *The Theory of Transformation groups* (Oxford: Oxford University Press, 1991)

[27] M. A. Kervaire: A manifold which does not admit any differentiable structure *Comment. Math. Helv.* **34** (1960) 257–270

[28] W. Koch and D. Puppe: Differenzierbare Strukturen auf Mannigfaltigkeiten ohne abzählbare Basis *Arch. Math.* **19** (1968) 95–102

[29] J. L. Koszul: Sur certains groupes des transformations de Lie *In: Colloque de Géométrie Différentielle* (Strasbourg, 1953) 137–141

[30] J. L. Koszul: *Lectures on Groups of Transformations* (Bombay: Tata Institute of Fundamental Research, 1965)

[31] S. Lie: *Theorie der Transformationsgruppen, Band III* (Leipzig: unter Mitwirkung von F. Engel, 1893)

[32] A. I. Mal'cev: On solvable topological groups *Rec. Math. (Mat. Sbornik) N.S.* **19** (1946) 165–174

[33] J. Mather: Stability of C^∞-mappings: II. Infinitesimal stability implies stability *Ann. of Math.* **89** (1969) 254–291.

[34] T. Matumoto and M. Shiota: Unique triangulation of the orbit space of a differentiable transformation group and its applications, Homotopy Theory and Related Topics *Adv. Stud. Pure Math.* **9** (1986) 41–55

[35] J. Milnor: *Differential topology: Lecture notes* (Princeton, N.J.: Princeton University, 1958)

[36] D. Montgomery: Properties of finite-dimensional groups *In: Proc. International Congress of Mathematicians, Vol. II* (Cambridge, U.S.A., 1950) 442–446

[37] D. Montgomery: Topological transformation groups *In: Proc. International Congress of Mathematicians, Vol. III* (Amsterdam, 1954) 185–188

[38] D. Montgomery and L. Zippin: Small subgroups of finite-dimensional groups *Ann. of Math.* **56** (1952) 213–241

[39] D. Montgomery and L. Zippin: Examples of transformation groups *Proc. Amer. Math. Soc.* **5** (1954) 460–465

[40] D. Montgomery and L. Zippin: *Topological Transformation Groups* (New York and London: Interscience Publishers, 1955)

[41] C. B. Morrey, Jr.: The analytic embedding of abstract real-analytic manifolds *Ann. of Math.* **68** (1958) 159–201

[42] R. Narasimhan: *Analysis on real and complex manifolds* (Amsterdam: North-Holland, 1968)

[43] J. von Neumann: Die Einführung analytischer Parameter in topologischen Gruppen *Ann. of Math.* **34** (1933) 170–190

[44] S. Packman: *e-mail letter to the author* (4 August, 1998)

[45] R. S. Palais: On the existence of slices for actions of non-compact Lie groups *Ann. of Math.* **73** (1961) 295–323

[46] R. S. Palais: C^1 actions of compact Lie groups on compact manifolds are C^1-equivalent to C^∞ actions *Amer. J. Math.* **92** (1970) 748–760

[47] R. S. Palais: When proper maps are closed *Proc. Amer. Math. Soc.* **24** (1970) 835–836

[48] A. du Plessis and T. Wall: *The Geometry of Topological Stability* (Oxford: Oxford University Press, 1995)

[49] L. S. Pontryagin: *Topological Groups, Second Edition* (New York: Gordon and Breach, 1966)

[50] W. Schmid: Poincaré and Lie groups *Bull. Amer. Math. Soc.* **6** (1982) 175–186

[51] F. Schur: Ueber den analytischen Charakter der eine endliche continuirliche Transformationsgruppe darstellenden Functionen *Math. Ann.* **41** (1893) 509–538

[52] N. Steenrod: A convenient category of topological spaces *Michigan Math. J.* **14** (1967) 133–152

[53] A. G. Wasserman: Equivariant differential topology *Topology* **8** (1969) 127–150

[54] H. Whitney: Analytic extensions of differentiable functions defined in closed sets *Trans. Amer. Math. Soc.* **36** (1934) 63–89

[55] H. Whitney: Differentiable manifolds *Ann. of Math.* **37** (1936) 645–680

[56] H. Yamabe: On the conjecture of Iwasava and Gleason *Ann. of Math.* **58** (1953) 48–54

[57] H. Yamabe: Generalization of a theorem of Gleason *Ann. of Math.* **58** (1953) 351–365

[58] C. T. Yang: Hilbert's fifth problem and related problems on transformation groups *In: Mathematical developments arising from Hilbert problems* (Proc. Sympos. Pure Math. **28** Amer. Math. Soc., 1976) 142–146

Sören Illman

Department of Mathematics and Statistics, P.O. Box 68 (Gustaf Hällströmin katu 2b), FIN-00014 University of Helsinki, Finland

E-mail: `soren.illman@helsinki.fi`

This page intentionally left blank

Handbook of Global Analysis
Demeter Krupka and David Saunders

Exterior differential systems

Niky Kamran

Contents

1 Introduction

The modern theory of exterior differential systems was founded by Elie Cartan, who gave a masterful account of the subject in his treatise "*Les systèmes différentiels extérieurs et leurs applications géométriques*", [20], published in 1945. The monograph by Kähler "*Einführung in die Theorie der Systeme von Differentialgleichungen*", [33], published in 1937, was also a major milestone in the development of the subject. More recently, the subject has undergone a remarkable revival, fuelled in part by the publication in 1991 of a major research treatise on the subject of exterior differential systems by Bryant, Chern, Gardner, Goldschmidt and Griffiths, [7]. Since its publication, this book has rightly become the standard reference on the subject.

The theory of exterior differential systems gives a geometric and coordinate-free approach to the formulation and solution of differential equations. In the framework of exterior differential systems, differential equations are replaced by differential ideals in the exterior algebra of differential forms on a manifold, and the solutions of differential equations correspond to integral manifolds of these ideals. Exterior differential systems are thus very well suited to the study of the differential systems that arise in differential geometry and in mechanics, particularly in geometric control theory. Important classical examples of problems that have been treated with great success using exterior differential systems include the local isometric embedding problems in Riemannian geometry, nearly all the classical deformation and classification problems for submanifolds, the local equivalence problem for G-structures (also known as the Cartan equivalence problem), and the study of sub-Riemannian structures and their invariants. The theory of exterior differential systems has a rich algebraic content, stemming from the systematic use of exterior algebra and representation theory in the geometric formulation of systems of differential equations as differential ideals. The analytic cornerstones of the subject are the Frobenius Theorem and the Cartan-Kähler Theorem, which are themselves consequences of the fundamental existence and uniqueness theorems for systems of ordinary differential equations on the one hand and the Cauchy-Kovalevskaia Theorem on the other. This means in particular that the Cartan-Kähler Theorem, which is the most general existence theorem that is currently available for exterior differential systems will in general only hold for systems of class C^ω. It is fair to say that in spite of the remarkable successes which have been achieved on some specific problems in the smooth category, in particular on the isometric embedding problem, [13], on systems of hyperbolic type, [48], the analytic aspects of the theory are still in need of development, so that a completely general version of the Cartan-Kähler Theorem holding for systems of class C^∞ is still waiting to be discovered.

It is of course impossible to cover all the different aspects the vast and beautiful subject of exterior differential systems in the context of the present article. What we have attempted to do is to present as many of the main concepts and results as possible in a self-contained manner. No proofs are given, but the most important existence theorems are illustrated by means of examples ranging from the elementary to the more sophisticated, so as to enable the reader to see how these theorems are to be applied in concrete situations. Bibliographical indications are also given to the main recent developments on each of the themes covered in the text.

Our paper is organized as follows. In Section 2, we introduce the basic notions of exterior differential systems, Pfaffian systems and integral manifolds, that we illustrate on a few very simple examples. Section 3 is devoted to the basic existence theorems for integral manifolds of systems of class C^∞, including the Frobenius and Darboux Theorems. These are essentially normal form theorems in which the expression and degree of generality of the integral manifolds of the given system are manifest. In Section 4, we consider the general problem of the existence of integral manifolds of exterior differential systems of class C^ω, and we define the class of involutive systems, for which the Cartan-Kähler Theorem is shown to imply the existence of integral manifolds tangent to sufficiently generic integral elements, called ordinary integral elements. Section 5 is devoted to the notion of prolongation and to the Cartan-Kuranishi Theorem. The prolongation of an exterior differential system is the system obtained by adding to the system its differential consequences, and the Cartan-Kuranishi Theorem asserts that after finitely many prolongations, a sufficiently generic system of class C^ω becomes either involutive, in which case it admits integral manifolds by virtue of the Cartan-Kähler Theorem, or has no integral manifolds. In Section 6, we present a version of the Cartan-Kähler Theorem due to Yang, [48], that applies to a class of involutive systems of class C^∞ whose characteristic variety satisfies a certain hyperbolicity condition. Section 7 gives an introduction to the characteristic cohomology theory for exterior differential systems developed by Bryant and Griffiths, [8], [9]. The caracteristic cohomology classes of a differential system correspond to its conservation laws and contain a lot of information about its integrability properties. In Section 8, we give a brief introduction to the classical topological obstructions in terms of characteristic classes to the global existence on a compact manifold of Pfaffian systems which are either completely integrable or of Goursat type. Section 9 illustrates how the existence Theorems of Section 3 can be applied to the study of normal forms of second-order scalar hyperbolic partial differential equations in the plane. In Section 10, we present a couple of geometric applications of the Cartan-Kähler Theorem and the involutivity test to two problems of classical local differential geometry, one having to do with submanifolds of projective space whose pencil of second fundamental forms can be simultaneously diagonalized, and the other with the existence of orthogonal local coordinates for 3-dimensional Riemannian metrics.

We conclude this section by remarking that there is a rich and very well developed theory of over-determined systems of partial differential equations, due to Spencer, [43], and developed by his students and collaborators, including Goldschmidt and Kumpera, which we will not consider in this article. The monograph [7] gives an account of the main results of the Spencer school. Likewise, Lie and Vessiot developed a geometric theory of differential equations based on the study of vector field systems. We refer the reader to the treatise by Stormark, [44], for an excellent account of Lie's theory and its developments.

2 Exterior differential systems

In this section, we introduce the concept of an exterior differential system, which we illustrate on a few examples. Unless otherwise specified, all the manifolds, maps and associated geometric structures such as bundles, differential forms, vector fields, etc... are assumed to be of class C^∞ and of constant rank.

Definition 2.1 An exterior differential system on an n-dimensional manifold M_n is a differentially closed ideal $\mathcal{I}$ in the ring $\Omega^*(M_n)$ of exterior differential forms on M_n.

The exterior differential systems considered in this paper will always be *finitely generated*. We shall use the notation

$$\mathcal{I} = \{\omega^1, \dots, \omega^q\} ,$$

to denote the algebraic ideal generated by q linearly independent differential forms $\omega^1, \dots, \omega^q$ in $\Omega^*(M_n)$.

Definition 2.2 A Pfaffian system on an n-dimensional manifold M_n is an exterior differential system $\mathcal{I}$ whose generators as a differential ideal are all 1-forms, that is a system $\mathcal{I}$ of the form

$$\mathcal{I} = \{\theta^1, \dots, \theta^s, d\theta^1, \dots, d\theta^s\} ,$$

where $\theta^1, \dots, \theta^s$ are linearly independent 1-forms on M_n. The integer s is called the rank of $\mathcal{I}$.

The sub-bundle of T^*M_n generated by $\theta^1, \dots, \theta^s$ will be denoted by I, and we will allow for a slight abuse of notation by using the same letter I to denote the $C^\infty(M_n;\mathbb{R})$-module of sections of I.

Definition 2.3 A p-dimensional integral manifold of an exterior differential system $\mathcal{I}$ is an immersion $f : W_p \to M_n$ such that

$$f^*\omega = 0 \quad \text{for all} \quad \omega \in \mathcal{I}.$$

We now illustrate the notion of integral manifold by means of a few examples.

Example 2.4 On $\mathbb{R}^3$ with coordinates (x, y, z) consider the Pfaffian system

$$\mathcal{I} = \{ dx + dy, dz + 2ydy \} .$$

The curves $(x, -x + c_1, -x^2 + c_2)$, where c_1, c_2 are arbitrary real constants are 1-dimensional integral manifolds of $\mathcal{I}$. The integral manifolds of $\mathcal{I}$ thus depend on two arbitrary real constants.

Example 2.5 We work on $\mathbb{R}^{2n+1}$ with coordinates $(x, u^1, \ldots, u^n, p_1, \ldots, p_n)$, and consider the Pfaffian system

$$\mathcal{I} = \{ du^i - p^i dx, dp^i \wedge dx, 1 \leq i \leq n\} .$$

This Pfaffian system admits all the curves $(x(t), u^1(t), \ldots,\ u^n(t), p_1(t), \ldots, p_n(t))$ such that

$$u^i_t - p^i x_t = 0 ,$$

as 1-dimensional integral manifolds. The integral manifolds of $\mathcal{I}$ thus depend on n arbitrary C^∞ functions of one variable.

Example 2.6 On M_4 defined as $\mathbb{R}^4$ with coordinates (x, y, z, u) minus the locus $y(x + y^2) = 0$, consider the Pfaffian system $\mathcal{I} = \{\omega, d\omega\}$, where

$$\omega = (x + y^2)y^2 dz - y(yz + u^2(x + y^2)^2)dx + (u^2 x(x + y^2)^2 - 2y^3 z)dy.$$

The Pfaffian system $\mathcal{I}$ admits all the submanifolds of M_4 of the form

$$\frac{z}{x + y^2} = f(\frac{x}{y}),\ u^2 = f'(\frac{x}{y}),$$

as integral manifolds. The integral manifolds of $\mathcal{I}$ thus depend on one arbitrary C^∞ functions of one variable.

Example 2.7 We consider the non-Pfaffian system

$$\mathcal{I} \ = \ \{ \sum_{i=1}^{n} dp_i \wedge dx^i \} ,$$

on $\mathbb{R}^{2n}$, with coordinates $(\ x^1, \ldots, x^n, p_1, \ldots, p_n\ )$. The submanifolds $(\ x^1, \ldots, x^n, p_1 = \frac{\partial f}{\partial x^1}, \ldots, p_n = \frac{\partial f}{\partial x^n}\ )$ are integral manifolds of $\mathcal{I}$ of dimension n for any choice of a C^∞ function f of n variables.

Example 2.8 On $\mathbb{R}^8$ with coordinates (x, y, u, p, q, r, s, t), we consider the Pfaffian system

$$\mathcal{I} = \{\theta^1, \theta^2, \theta^3, d\theta^1, d\theta^2, d\theta^3, dF\} ,$$

where

$$\theta^1 = du - pdx - qdy\ ,\ \theta^2 = dp - rdx - sdy\ ,\ \theta^3 = dq - sdx - tdy\ ,$$

and $F : \mathbb{R}^8 \to \mathbb{R}$ is a smooth function such that $(F_r, F_s, F_t) \neq (0, 0, 0)$. The surfaces

$$(\ x(w, z), y(w, z), u(w, z), p(w, z), q(w, z), r(w, z), s(w, z), t(w, z)\ ) ,$$

such that

$$F(\, x(w,z), y(w,z), u(w,z), p(w,z), q(w,z), r(w,z), s(w,z), t(w,z)\,) = 0\,,$$

and

$$u_w - px_w - qy_w = 0\,,\; u_z - px_z - qy_z = 0\,,$$
$$u_w - px_w - qy_w = 0\,,\; u_z - px_z - qy_z = 0\,,$$
$$u_w - px_w - qy_w = 0\,,\; u_z - px_z - qy_z = 0\,.$$

Note that if

$$\left|\frac{\partial(x,y)}{\partial(w,z)}\right| \neq 0\,,$$

then the integral surfaces of $\mathcal{I}$ can be locally parametrized as graphs of the form

$$(\, x, y, u(x,y), p(x,y), q(x,y), r(x,y), s(x,y), t(x,y)\,)\,, \tag{1}$$

with

$$p = u_x\,,\; q = u_y\,,\; r = u_{xx}\,,\; s = u_{xy}\,,\; t = u_{yy}\,,$$

and $u(x,y)$ will be a solution of the second-order partial differential equation

$$F(x, y, u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}) = 0\,. \tag{2}$$

We remark that one could have equivalently considered on the hypersurface M_7 of $\mathbb{R}^8$ the Pfaffian system

$$\mathcal{I}_F = \{i^*\theta^1, i^*\theta^2, i^*\theta^3, i^*d\theta^1, i^*d\theta^2, i^*d\theta^3\}\,,$$

obtained by pulling-back under the inclusion map $i : M_7 \to \mathbb{R}^8$ the generators of $\mathcal{I}$ to M_7. The integral manifolds of the form (1) will also correspond to solutions of the second-order partial differential equation (2).

3 Basic existence theorems for integral manifolds of C^∞ systems

The simple examples presented at the end of the preceding section show that an exterior differential system may admit integral manifolds depending on arbitrary constants or arbitrary functions, or even both. It is therefore natural to ask if there exist theorems that make it possible to determine a-priori whether an exterior differential system will admit any integral manifolds at all, and if so to predict their degree of generality. Our purpose in this section and and the next section is to review a number of classical results which aim to answer these questions.

The analytical cornerstones of the classical theory of exterior differential systems are the existence, uniqueness and smooth dependence on the initial data theorems for C^∞ systems of first-order ordinary differential equations, and the Cauchy-Kowalewskaia existence and uniqueness theorem for the non-characteristic Cauchy problem for C^ω systems of partial differential equations. Accordingly, we will consider the C^∞ and C^ω cases separately, and devote the present section to an overview of the classical existence theorems for exterior differential systems of class C^∞, leaving the C^ω case for the next section. So we will assume throughout this section that all the manifolds, maps and associated geometric structures such as bundles, differential forms, vector fields, etc... are of class C^∞.

It is important to remark that the differential systems to which the C^∞ theorems are applicable are rather special, even though they are often of great interest in geometric applications. Most of these theorems are actually normal form results depending only numerical invariants, in which the expression of the local integral manifolds and their degree of generality is manifest. As one would expect, these types normal form results apply only to systems which are "flat" in a well-defined geometric sense.

We begin with Frobenius theorem, which is the simplest, but nevertheless the most important existence theorem for integral manifolds of Pfaffian systems.

Theorem 3.1 *Let*

$$\mathcal{I} = \{\theta^1, \dots, \theta^s, d\theta^1, \dots, d\theta^s\}\,,$$

be a Pfaffian system of rank s on an n-manifold M_n and suppose that

$$d\theta^a \wedge \theta^1 \wedge \dots \wedge \theta^s = 0, \quad 1 \le a \le s\,. \tag{3}$$

Then for every $x \in M_n$, there exists a coordinate neighborhood U with $x \in U$ and local coordinates $(u^1, \dots, u^n)$ in which $\mathcal{I}$ is given by

$$\mathcal{I} = \{du^1, \dots, du^s\}\,.$$

It follows from this theorem that the local integral manifolds of $\mathcal{I}$ are given by the joint level sets $u^1 = c_1, \dots, u^s = c_s$, where $c_1, \dots c_s$ are real constants. Pfaffian systems satisfying the Frobenius condition (3) are often referred to as being *completely integrable*, and the functions $u^1, \dots, u^s$ are referred to as *first integrals* of $\mathcal{I}$. Example 2.4 is an example of a completely integrable Pfaffian system.

Even though the definition we have just given of complete integrability is a global one, Theorem 3.1 is local in nature. This is due in part to the fact that the existence theorems for solutions of ordinary and partial differential equations that underlie the proof of the Frobenius theorem are themselves local results. It is therefore natural to ask whether the local integral manifolds defined in Theorem 3.1 can be "glued" together to give rise to maximal integral submanifolds. The Frobenius theorem has a global version that shows this to be true. We now briefly recall the statement of this result, [22], [47].

Definition 3.2 A maximal integral manifold of a Pfaffian system $\mathcal{I}$ is a connected integral manifold whose image in not a proper subset of any other connected integral manifold of $\mathcal{I}$.

The global Frobenius Theorem is now as follows:

Theorem 3.3 *Let $\mathcal{I}$ be a completely integrable Pfaffian system of rank s on an n-manifold M_n. Then through any $x \in M_n$ there passes a unique maximal connected integral manifold W_s of M_n, and every other connected integral manifold of $\mathcal{I}$ passing through x is contained in W_s.*

There is a simple but nevertheless very important invariant structure that gives a coarse measure of the degree to which a Pfaffian system may fail to be completely integrable. This invariant is called the *derived flag*, and is defined as follows. Let

$$\mathcal{I} = \{\theta^1, \ldots, \theta^s, d\theta^1, \ldots, d\theta^s\}\,,$$

be a Pfaffian system and let

$$I = \{\theta^1, \ldots, \theta^s\} \subset \Omega^1(M_n)\,,$$

denote the corresponding $C^\infty(M_n; \mathbb{R})$-module of sections of T^*M_n. Let $\{I\}$ denote the ideal generated by I in the algebra $\Omega^*(M_n)$ of C^∞ differential forms on M_n. The exterior differential

$$d : I \to \Omega^2(M_n)\,,$$

induces a map

$$\delta : I \to \Omega^2(M_n)/(\{I\} \cap \Omega^2(M_n))\,.$$

We define the *first derived system* $I^{(1)}$ of I to be the kernel of δ,

$$I^{(1)} = \ker \delta\,.$$

So, $\mathcal{I}$ will satisfy the Frobenius condition (3) if and only if

$$I^{(1)} = I\,.$$

The derived flag of I is the flag of $C^\infty(M_n; \mathbb{R})$-modules corresponding to the higher derived systems of I, that is,

$$\cdots \subset I^{(k)} \subset I^{(k-1)} \subset \cdots \subset I^{(1)} \subset I\,,$$

where

$$I^{(k)} = (I^{(k-1)})^{(1)}\,.$$

The string of codimensions of the elements of the derived flag of a Pfaffian system I is a numerical invariant, called the *type* of a Pfaffian system. We let N be the smallest integer such that

$$I^{(N+1)} = I^{(N)}\,,$$

and define the integers $P_a, 0 \le a \le N+1$, by

$$P_0 = \dim I^{(N)}, \quad P_{N-i} = \dim I^{(i)}/I^{(i+1)}, \quad P_{N+1} = n - \dim I\,.$$

The type of I is then defined as the $(N+2)$-tuple $(P_0, \ldots, P_{N+1})$. The type of a Pfaffian system is not arbitrary. We have, [25]:

Proposition 3.4 *Let I be Pfaffian system of type $(P_0, \ldots, P_{N+1})$ on an n-manifold M_n. Then for all $-1 \leq i \leq N-1$, we have*

$$P_{N-i-1} \leq P_{N-i}(P_{N-i+1} + \cdots + P_{N+1}) + \binom{P_{N-i}}{2}.$$

The local Frobenius theorem (Theorem 3.1) says that the generators of a completely integrable Pfaffian system of rank s can be locally chosen as the differentials of s functionally independent local coordinate functions. For exterior differential systems which are not completely integrable Pfaffian systems, the question remains of knowing if generators can be expressed in terms of a minimal set of local coordinates and their differentials. There is a classical construction, which we now review, that provides such a set of coordinates. We let Char$(\mathcal{I})$ denote the $C^\infty(M_n;\mathbb{R})$-module of *Cauchy characteristic vector fields* of $\mathcal{I}$, defined by

$$\mathrm{Char}(\mathcal{I}) = \{X \in \Gamma(TM_n) \,|\, X \in \mathcal{I}^\perp,\ X \rfloor d\mathcal{I} \subset \mathcal{I}\}.$$

Definition 3.5 The Cartan system of $\mathcal{I}$ is defined as the Pfaffian system generated as a differential ideal by the 1-forms that annihilate all Cauchy characteristic vector fields. The class of an exterior differential system is by definition the rank of its Cartan system.

We have:

Proposition 3.6 *The Cartan system $C(\mathcal{I})$ of any exterior differential system $\mathcal{I}$ is a completely integrable Pfaffian system.*

The following retraction theorem shows that the first integrals of the Cartan system $C(\mathcal{I})$ provide a minimal set of local coordinates with which one can express the generators of $\mathcal{I}$.

Theorem 3.7 *Let $\mathcal{I}$ be an exterior differential system of class r and let $\{w^1, \ldots, w^r\}$ denote a set of local first integrals of the Cartan system $C(\mathcal{I})$. Then this set can be completed to a local coordinate chart $(w^1, \ldots, w^r, y^{r+1}, \ldots, y^n)$ in which $\mathcal{I}$ is generated by 1-forms expressible in terms of $w^1, \ldots, w^r$ and their differentials.*

Going back to the examples of Section 2, we see that Example 2.4 is a completely integrable Pfaffian system of rank two and therefore of class two, whereas Example 2.6 is a Pfaffian system which is not completely integrable, of class three.

The simplest class of Pfaffian systems which are not completely integrable are the Darboux systems, which are at the source of contact geometry:

Theorem 3.8 *Let*

$$\mathcal{I} = \{\omega, d\omega\},$$

be a Pfaffian system on M_n. Suppose that for some integer $r \geq 0$ we have

$$(d\omega)^r \wedge \omega \neq 0, \quad (d\omega)^{r+1} \wedge \omega = 0, \tag{4}$$

on M_n. Then for every $x \in M_n$, there exists a coordinate neighborhood U with $x \in U$ and local coordinates $(x^1, \ldots, x^r, z, p^1, \ldots, p^r, u^{2r+2}, \ldots, u^n)$ such that

$$\mathcal{I} = \{dz - \sum_{i=1}^{r} p_i dx^i,\ \sum_{i=1}^{r} dp_i \wedge dx^i\}.$$

The integral manifolds of $\mathcal{I}$ are locally parametrized by one arbitrary C^∞ function of one variable. If $\mathcal{I}$ satisfies (4), then the derived flag of the corresponding $C^\infty(M_n;\mathbb{R})$-module I is given by

$$I = \{\omega\}, \quad I^{(1)} = \{0\}.$$

Example 2.6 is a Darboux system.

We now state the Goursat normal form theorem, which is an extension of the preceding result:

Theorem 3.9 *Let*

$$\mathcal{I} = \{\omega^1, \ldots, \omega^r, d\omega^1, \ldots d\omega^r\},$$

be a Pfaffian system on M_n. Suppose that there exist 1-forms α and π, where α and π are not congruent to zero modulo $\mathcal{I}$, such that,

$$\begin{aligned} d\omega^1 &\equiv \omega^2 \wedge \pi, \quad \mod \{\omega^1\}, \\ d\omega^2 &\equiv \omega^3 \wedge \pi, \quad \mod \{\omega^1, \omega^2\}, \\ &\vdots \\ d\omega^{r-1} &\equiv \omega^r \wedge \pi, \quad \mod \{\omega^1, \ldots, \omega^{r-1}\}, \\ d\omega^r &\equiv \alpha \wedge \pi, \quad \mod \{\omega^1, \ldots \omega^r\}. \end{aligned}$$

Then there exist local coordinates $(x, y, y^1, \ldots, y^r)$ such that

$$\mathcal{I} = \{dy - y^1 dx, \ \ldots, dy^{r-1} - y^r dx, dy^1 \wedge dx, \ldots, dy^r \wedge dx\}.$$

The integral manifolds of a Goursat system are thus locally parametrized by one arbitrary C^∞ function of one variable.

An interesting application of the Goursat normal form arises in Cartan's proof, [18], of Hilbert's theorem [30] on the non-existence of parametric solutions of finite rank for the Hilbert-Cartan equation. Recall [35] that an under-determined ordinary differential equation

$$v' = F(x, u, v, u', u''), \tag{5}$$

is said to admit parametric solutions of finite rank if its solutions are of the form

$$\begin{aligned} x &= X(t, w(t), w'(t), \ldots, w^{(r)}(t)), \\ u &= U(t, w(t), w'(t), \ldots, w^{(r)}(t)), \\ v &= V(t, w(t), w'(t), \ldots, w^{(r)}(t)), \end{aligned}$$

where $w(t)$ is an arbitrary C^∞ function. The Hilbert-Cartan equation is given by

$$v' = (u'')^2, \tag{6}$$

and Hilbert's theorem asserts that (6) does not admit parametric solutions of finite rank. Cartan's proof of Hilbert's result proceeds as follows. For a Goursat system $\mathcal{I}$, we have

$$I = \{\omega^1, \ldots, \omega^r\},$$

$$
\begin{aligned}
I^{(1)} &= \{\omega^1, \dots, \omega^{r-1}\}\,,\\
&\vdots\\
I^{(r-1)} &= \{\omega^1\}\,,\\
I^{(r)} &= \{0\}\,,
\end{aligned}
$$

so that the sequence of dimensions of the elements of the derived flag of I is given by

$$\dim I^{(k)} = r - k, \quad 0 \le k \le r\,.$$

Now, if the equation (5) is to admit parametric solutions of finite rank, then we must have $r = 3$ and the Pfaffian system $\mathcal{I}$ corresponding to

$$I = \{du - u'dx,\, du' - u''dx,\, dv - Fdx\}\,,$$

must be a Goursat system with $r = 3$. It is easy to calculate the derived flag of I: we have

$$I^{(1)} = \{du - u'dx,\, dv - \frac{\partial F}{\partial u''}du' - (F - u''\frac{\partial F}{\partial u''})dx\}\,,$$

and $\dim I^{(2)} = 1$ if and only if

$$\frac{\partial^2 F}{(\partial u'')^2} = 0\,.$$

The Hilbert-Cartan equation does not satisfy this condition, it follows that it does not admit parametric solutions of finite rank. On the other hand, the ordinary differential equation

$$v' = u^m u''\,,$$

admits the parametric solutions of finite rank given by, [16],

$$
\begin{aligned}
x(t) &= -2tf''(t) - f'(t)\,,\\
y(t) &= [(m+1)^2 t^3 (f''(t))^2]^{\frac{1}{m+1}}\,,\\
z(t) &= (m-1)t^2 f''(t) - mtf'(t) + mf(t)\,.
\end{aligned}
$$

4 Involutive analytic systems and the Cartan-Kähler Theorem

The theorems reviewed in Section 3 can be thought of as normal form results for exterior differential systems, in which the expression of the local integral manifolds is manifest. These theorems therefore apply only to very special systems, and it is natural to expect that a general existence theorem should exist which will be applicable to a much wider class of systems, for which one would not expect that a normal form would be readily available on the basis of the behavior of a few simple invariants. Such a general theorem does indeed exist, and is known as the Cartan-Kähler Theorem. This existence theorem applies to the general class of exterior differential systems which are *in involution*. The proof of the Cartan-Kähler Theorem involves successive applications of the Cauchy-Kovalevskaia theorem, and consequently all the manifolds and differential systems considered in this

Section will be assumed to be of class C^ω. Our objective here is to present the statement of this very important theorem, and to show how it can be applied in a few simple concrete situations. Section 10 will contain more examples. In Section 6, we will present a version of the Cartan-Kähler Theorem which applies to a special class exterior differential systems of class C^∞.

We begin with some elementary algebraic preliminaries. Let V be an n-dimensional real vector space. The pairing $< .,. >$ between V and V^* induces a pairing between the exterior algebras $\bigwedge(V)$ and $\bigwedge(V^*)$; namely, if $\{e_1, \ldots, e_n\}$ denotes a basis of V and $\{e^{*1}, \ldots, e^{*n}\}$ denotes the dual basis of V^*, then the pairing between the exterior algebras is defined by

$$< v, \phi >= \frac{1}{p!} \sum_{1 \leq i_1 \ldots i_p \leq n} v^{i_1 \ldots i_p} \phi_{i_1 \ldots i_p} \,,$$

for

$$v = \frac{1}{p!} \sum_{1 \leq i_1 \ldots i_p \leq n} v^{i_1 \ldots i_p} e_{i_1} \wedge \cdots \wedge e_{i_p}, \; \phi = \frac{1}{p!} \sum_{1 \leq i_1 \ldots i_p \leq n} \phi_{i_1 \ldots i_p} e^{*i_1} \wedge \cdots \wedge e^{*i_p} \,,$$

and

$$< v, \phi >= 0 \,,$$

for $v \in \bigwedge^p(V)$ and $\phi \in \bigwedge^q(V^*)$, with $p \neq q$. To a k-dimensional subspace $W \subset V$ with basis $\{v_1, \ldots, v_k\}$, we associate the decomposable k-vector $v_W = v_1 \wedge \cdots \wedge v_k \in \bigwedge^k(V)$. We have $w \in W$ if and only if $w \wedge v_W = 0$, so that v_W can be thought of as representing the subspace W and it is easily verified that any two k-vectors representing the same subspace are non-zero multiples of each other. We denote by $[W] \in \mathbb{P}(\bigwedge^k(V))$ the equivalence class of any k-vector representing W.

It follows from Definition 2.3 that an immersion $h : W_m \to M_n$ is an integral manifold of an exterior differential system $\mathcal{I}$ on an n-manifold M_n if and only if

$$< h_* T_u W_m, \omega >= 0 \,,$$

for all $u \in W_m$ and $\omega \in \mathcal{I}$. This suggests the following definition:

Definition 4.1 Let $x \in M_n$ and let $E^p \subset T_x M_n$ denote a p-dimensional subspace of the tangent space to M_n at x. The pair (x, E^p) is a called a p-dimensional integral element of $\mathcal{I}$ if $< E^p, \omega|_x >= 0$ for all $\omega \in \mathcal{I}$.

There is of course no reason for an integral element based at a given point to be tangent to an integral manifold passing through that point. The purpose of the Cartan-Kähler theorem is precisely to give necessary and sufficient conditions under which this will be the case.

The construction of integral manifolds of maximal dimension is done by building up the integral manifolds one dimension at a time, by successive applications of the Cauchy-Kovalevskaia Theorem. We thus introduce the notion of the polar space of an integral element, which describes all the higher dimensional integral elements containing a given integral element.

Definition 4.2 The polar space $H(E^p)$ of an integral element (x, E^p) of an exterior differential system $\mathcal{I}$ is the set of tangent vectors $v \in T_x M_n$ such that

$$< [E^p] \wedge v, \omega|_x >= 0 \,,$$

for all $\omega \in \mathcal{I}$.

The polar space of $H(E^p)$ is thus a subspace of $T_x M_n$, containing E^p,

$$E^p \subseteq H(E^p) \,,$$

and we can write

$$\dim H(E^p) = p + 1 + \sigma_{p+1} \,,$$

where $\sigma_{p+1} \geq -1$. Consider now the Grassmann bundle $\pi : G_p(M_n) \to M_n$, whose fiber at $x \in M_n$ is the Grassmann manifold of p-dimensional subspaces of $T_x M_n$, and let $\mathcal{V}_p(\mathcal{I}) \subset G_p(M_n)$ denote the variety of p-dimensional integral elements of $\mathcal{I}$. We will be interested in flags of integral elements which are in general position in the Grassmann bundle. These flags will be nested in those integral elements of maximal dimension, called *ordinary*, to which the Cartan-Kähler Theorem will apply.

Definition 4.3 An integral element (x, E^p) of an exterior differential system $\mathcal{I}$ is said to be Kähler-regular if there exist s independent p-forms $\beta^1, \ldots, \beta^s$ such that $\mathcal{V}_p(\mathcal{I})$ is given in a neighborhood of (x, E^p) by the equations

$$< [E^p], \beta^1 >= 0 \,, \; \ldots \,, < [E^p], \beta^1 >= 0 \,,$$

and the rank of the polar equations of a p-dimensional integral element is constant in a neighborhood of (x, E^p).

The concept of an exterior differential systems does not single out a set of independent variables. However, one is led in many applications to introduce a transversality condition which restricts the set of integral manifolds under consideration (see Section 10 for examples where such transversality conditions arise in geometrical applications). This leads to the following definition:

Definition 4.4 An exterior differential system with independence condition on a manifold M_n is an exterior differential system $\mathcal{I}$ endowed with a decomposable p-form Ω which is not congruent to zero modulo $\mathcal{I}$. The integral elements and manifolds of dimension p on which $\Omega \neq 0$ are said to be admissible.

We are now ready to define the notion of an ordinary integral element of an exterior differential system with independence condition.

Definition 4.5 An admissible integral element (x, E^p) of an exterior differential system with independence condition $(\mathcal{I}, \Omega)$ is said to be ordinary if there exists a flag

$$E^0 \subset E^1 \subset E^2 \subset \cdots \subset E^{p-1} \subset E^p \,,$$

such that each pair (x, E^i), $1 \leq i \leq p-1$, is Kähler-regular.

We can now define the concept of an exterior differential system in involution:

Definition 4.6 An exterior differential system with independence condition $(\mathcal{I}, \Omega)$ is said to be in involution if there exists an ordinary integral element (x, E^p) through each $x \in M_n$.

It is helpful to illustrate the concepts that we have just introduced by means of an elementary example.

Example 4.7 This example comes from [7]. We work in $\mathbb{R}^5$ and consider the differential ideal $\mathcal{I}$ generated by the 1-forms

$$\omega^1 = dx^1 + (x^3 - x^4x^5)dx^4\,, \quad \omega^2 = dx^2 + (x^3 + x^4x^5)dx^5\,,$$

so that

$$\mathcal{I} = \{\omega^1, \omega^2, d\omega^1 = \omega^3 \wedge dx^4, d\omega^2 = \omega^3 \wedge dx^5\}$$

with

$$\omega^3 = dx^3 + x^5dx^4 - x^4dx^5\,.$$

The 1-dimensional integral elements (x, E^1) of $\mathcal{I}$ are all the 1-dimensional subspaces $E^1 \subset K_x$ of the 3-plane field K defined by

$$K_x = \{X \in T_x\mathbb{R}^5 \,|\, \omega^1(X) = \omega^2(X) = 0\}\,.$$

In other words, $V_1(\mathcal{I}) = \mathbb{P}(K)$, the projectivization of K. Furthermore, we have a unique 2-dimensional integral element (x, E^2) through every $x \in \mathbb{R}^5$, defined by

$$E^2 = \{X \in T_x\mathbb{R}^5 \,|\, \omega^1(X) = \omega^2(X) = \omega^3(X) = 0\}\,,$$

which implies that the integral elements (x, E^2) are of maximal dimension and that $V_2(\mathcal{I}) \simeq \mathbb{R}^5$. We now endow $\mathcal{I}$ with the independence condition $\Omega = dx^4 \wedge dx^5$, and determine the regular integral elements of $(\mathcal{I}, \Omega)$. Every 2-plane in $G_2(\mathbb{R}^5, \Omega)$ has a basis $\{X_4, X_5\}$ of the form

$$\begin{aligned} X_4 &= \frac{\partial}{\partial x^4} + p_4^1\frac{\partial}{\partial x^1} + p_4^2\frac{\partial}{\partial x^2} + p_4^3\frac{\partial}{\partial x^3}\,, \\ X_5 &= \frac{\partial}{\partial x^5} + p_5^1\frac{\partial}{\partial x^1} + p_5^2\frac{\partial}{\partial x^2} + p_5^3\frac{\partial}{\partial x^3}\,. \end{aligned}$$

An element $(x, [X_4 \wedge X_5]) \in G_2(\mathbb{R}^5, \Omega)$ will be an admissible integral element if and only if

$$p_5^1 = p_4^2 = p_4^1 - x^3 + x^4x^5 = p_5^2 + x^3 + x^4x^5 = p_5^3 - x^4 = p_4^3 - x^5 = 0\,. \tag{7}$$

This shows that all the integral elements defined by (7) are Kähler-regular. It is easily verified that they are also ordinary.

We are now ready to state the Cartan-Kähler Theorem, which is the fundamental existence theorem for integral manifolds of analytic exterior differential systems.

Theorem 4.8 *If $(\mathcal{I}, \Omega)$ is in involution and (x, E^p) is an ordinary integral element, then there exists an admissible integral manifold $W_p \subset M_n$ through x, such that $T_xW_p = E^p$.*

It is thus very important to have a criterion to determine if a given exterior differential system with independence condition is in involution. Such a criterion is provided by E. Cartan's involutivity test.

Theorem 4.9 *Let $\mathcal{I}$ be an exterior differential system on M_n, and let*

$$E^0 \subset E^1 \subset E^2 \subset \cdots \subset E^{p-1} \subset E^p\,,$$

be a flag of integral elements of $\mathcal{I}$, based at $x \in M$. Let c_k, $1 \leq k \leq n-1$ denote the codimension of $H(E^k)$ in T_xM. Then the codimension c_{E^p} of $\mathcal{V}_p(\mathcal{I})$ in $G_p(M_n)$ at E_p satisfies the following bound,

$$c_{E^p} \leq \sum_{i=0}^{p-1} c_i\,. \tag{8}$$

Moreover, (x, E_p) is ordinary if and only if it has a neighborhood U in $G_p(M_n)$ such that $\mathcal{V}_p(\mathcal{I})$ is a smooth manifold of codimension

$$c = \sum_{i=0}^{p-1} c_i\,, \tag{9}$$

in U

In practice, it may be quite difficult to check whether an exterior differential system is in involution. There is however an important class of Pfaffian systems, called *quasi-linear*, for which the involutivity criterion of Theorem 4.9 can be verified by means of a linear-algebraic test. Quasi-linear systems are defined as follows. Consider a Pfaffian system with independence condition $(\mathcal{I}, \Omega)$, where

$$\mathcal{I} = \{\theta^1, \ldots, \theta^s, d\theta^1, \ldots, d\theta^s\},$$
$$\Omega = \omega^1 \wedge \cdots \wedge \omega^p\,,$$

and let $\pi^1, \ldots, \pi^l$, $l = n - s - p$, be 1-forms such that $\theta^1 \wedge \cdots \wedge \theta^s \wedge \omega^1 \wedge \cdots \wedge \omega^p \wedge \cdots \wedge \pi^1 \wedge \cdots \wedge \pi^l \neq 0$. We have, for $1 \leq i \leq s$,

$$d\theta^i \equiv \sum_{\alpha=1}^{l}\sum_{b=1}^{p} A^i_{\alpha b}\pi^\alpha \wedge \omega^b + \frac{1}{2}\sum_{a,b=1}^{p} B^i_{ab}\omega^a \wedge \omega^b + \frac{1}{2}\sum_{\alpha,\beta=1}^{l} C^i_{\alpha\beta}\pi^\alpha \wedge \pi^\beta \mod \theta^1, \ldots, \theta^s\,.$$

Definition 4.10 A Pfaffian system with independence condition $(\mathcal{I}, \Omega)$ is said to be quasi-linear if

$$C^i_{\alpha\beta} = 0, \quad 1 \leq i \leq s, \quad 1 \leq \alpha, \beta \leq l.$$

It is important to observe that if $(\mathcal{I}, \Omega)$ is quasi-linear, then $\mathcal{V}_p(\mathcal{I}, \Omega)$ is an affine bundle over M_n. As a consequence of this, the determination of the involutivity conditions is much easier than in the general case. The admissible integral elements (x, E^p) are of the form

$$(\pi^\alpha - \sum_{b=1}^{p} t^\alpha_b \omega^b)|_{E^p} = 0\,,$$

where $1 \leq \alpha \leq l$, and the polar equations of E^p are given by

$$\sum_{\alpha=1}^{l} (A^i_{\alpha b} t^\alpha_c - A^i_{\alpha c} t^\alpha_b) + B^i_{bc} = 0\,, \tag{10}$$

where $1 \leq b, c \leq p$ and $1 \leq i \leq s$. We denote the dimension of the set of solutions of the linear system (10) by d. Define the *reduced characters* $s'_1, \ldots, s'_r$, $r \leq p$ of $(\mathcal{I}, \Omega)$ by

$$s'_1 + \cdots + s'_r = \max_{v_1 \ldots v_r \in \mathbb{R}^l} \mathrm{rk} \begin{pmatrix} \sum_{\alpha=1}^{l} v_1^\alpha A^i_{\alpha b} \\ \vdots \\ \sum_{\alpha=1}^{l} v_r^\alpha A^i_{\alpha b} \end{pmatrix} \tag{11}$$

The involutivity test and criterion given in Theorems 8 and 9 take the following form:

Theorem 4.11 *We have*

$$d \leq \sum_{i=1}^{p} i s'_i ,$$

with equality if and only if the system $(\mathcal{I}, \Omega)$ is in involution. If $s'_q = k \neq 0$ with q maximal, then the admissible local integral manifolds are parametrized by k C^ω functions of q variables.

We now illustrate Theorem 4.11 on the following example, taken from [6].

Example 4.12 We consider a scalar partial differential equation

$$F(x^i, u, \frac{\partial u}{\partial x^i}, \frac{\partial^2 u}{\partial x^i \partial x^j}) = 0, \quad 1 \leq i, j \leq p, \tag{12}$$

where F is assumed to be C^ω in all its arguments. We apply the Cartan-Kähler theorem to show that the local solutions of the partial differential equation (12) are parametrized by two analytic functions of $p-1$ variables. To the partial differential equation (12), we associate on $\mathbb{R}^{\frac{p(p+5)}{2}+1}$, with local coordinates $(x^i, u, u_i, u_{ij} = u_{ji})$, $1 \leq i, j \leq p$, an exterior differential system with independence condition $(\mathcal{I}, \Omega)$, by letting $\mathcal{I}$ be the differential ideal generated by

$$F(x^i, u, u_i, u_{ij}) = 0, \quad \theta_0 = du - \sum_{i=1}^{p} dx^i,$$

$$\theta_i = du_i - \sum_{j=1}^{n} u_{ij} dx^j, \quad 1 \leq i \leq p,$$

where we assume that

$$\det(\frac{\partial F}{\partial u_{ij}})|_{F=0} \neq 0, \tag{13}$$

and the independence condition Ω be defined by

$$\Omega = dx^1 \wedge \cdots \wedge dx^p .$$

The structure equations of $(\mathcal{I}, \Omega)$ are given by

$$dF = 0, \quad d\theta_0 \equiv 0, \quad d\theta_i \equiv \sum_{j=1}^{n} \pi_{ij} \wedge dx^j, \qquad \mathrm{mod}\ \theta_0 \ldots \theta_p , \tag{14}$$

where $\pi_{ij} = -du_{ij}$. In order to compute the reduced characters of $(\mathcal{I}, \Omega)$, it is convenient to exploit the non-degeneracy condition (13) to put the above structure equations in normal form. If we perform a change of coframe according to

$$\bar{\omega}^i = \sum_{j=1}^{p} a^i{}_j \, dx^j, \quad \bar{\theta}_i = \sum_{j=1}^{p} (a^{-1})_i{}^j \, \theta_j \,,$$

we obtain the following transformation law for the 1-forms π_{ij},

$$\bar{\pi}_{ij} = -\sum_{k,l=1}^{p} du_{kl} \, a^k{}_i a^l{}_j \,.$$

Using the rank condition (13), we can therefore choose $(a^i{}_j)$ in such a way as to have

$$\sum_{i,j=1}^{p} \frac{\partial F}{\partial u_{ij}} \, (a^{-1})_i{}^k (a^{-1})_j^{\,l} = \delta_{ij} \varepsilon_i \,,$$

where $\varepsilon_i^2 = 1$. The structure equation $dF = 0$ now becomes

$$\sum_{i=1}^{p} \varepsilon_i \bar{\pi}_{ii} + \sum_{k=1}^{p} b_k \bar{\omega}^k \equiv 0, \quad \bmod \bar{\theta}_0 \ldots \bar{\theta}_p \,,$$

for some functions b_k, $1 \leq k \leq p$. We now let

$$\bar{\bar{\pi}}_{ii} = \bar{\pi}_{ii} + \varepsilon_i b_i \bar{\omega}_i, \quad \bar{\bar{\pi}}_{ij} = \bar{\pi}_{ij}, \quad 1 \leq i \neq j \leq p \,.$$

Dropping bars, the structure equations (14) become

$$d\theta_0 \equiv 0, \quad d\theta_i \equiv \sum_{j=1}^{p} \pi_{ij} \wedge \omega^j, \quad \bmod \theta_0, \ldots, \theta_p \,,$$

where

$$\sum_{i=1}^{p} \varepsilon_i \pi_{ii} \equiv 0, \; \pi_{ij} \equiv \pi_{ji}, \quad \bmod \theta_0, \ldots, \theta_p \,, \tag{15}$$

where $1 \leq i, j \leq p$. We are now ready to compute the reduced characters of $(\mathcal{I}, \Omega)$ using (11). We have

$$s_1' = p, \; s_2' = p - 1, \ldots, s_{p-1}' = 2, \; s_p' = 0 \,,$$

where the final drop from 2 to 0 is due to the trace condition (15). We thus have

$$\sum_{i=1}^{p} i s_i' = \frac{p(p+1)(p+2)}{6} - p \,.$$

On the other hand, an element (x, E^p) of the Grassmann bundle will be an admissible integral element of $(\mathcal{I}, \Omega)$ if and only if

$$\theta_0|_{E^p} = 0\,,\; \theta_i|_{E^p} = 0\,,\; (\pi_{ij} - \sum_{k=1}^{p} l_{ijk}\omega^k)|_{E^p} = 0\,,$$

where for $1 \leq i, j, k \leq p$, and where we have

$$l_{ijk} = l_{jik} = l_{ikj}\,,\; \sum_{i=1}^{p} \varepsilon_i l_{iik} = 0\,.$$

The dimension of the solution space of the polar equations of (x, E^p) is thus given by

$$\binom{p+2}{p-1} - p = \frac{p(p+1)(p+2)}{6} - p\,,$$

and the system is in involution, with top character $s'_{p-1} = 2$. The local C^ω solutions are thus parametrized by 2 arbitrary functions of $p-1$ variables, as claimed.

5 Prolongation and the Cartan-Kuranishi Theorem

Roughly speaking, the prolongations of a differential system are the differential systems obtained by adjoining to the original differential system its differential consequences. The concept of *prolongation tower*, which will be defined below, gives an abstract formulation of the operation of prolongation. A general conjecture of Elie Cartan, [20], proved by Kuranishi, [36], for a wide class of differential systems, states that an analytic differential system either becomes involutive after finitely many prolongations, or has no solutions. This result is known as the Cartan-Kuranishi Theorem. The proof of Cartan's conjecture has been given under a different set of hypotheses in the treatise [7]. Our purpose in this section is to review some of the basic aspects of the prolongation theorem. We assume that all the manifolds and differential systems under consideration are of class C^ω.

The prolongation tower of an exterior differential system with independence condition $(\mathcal{I}, \Omega)$ on an n-dimensional manifold M is defined as follows. Let $f : W_p \to M$ be an immersion and let $f_* : W_p \to G_p(M)$ denote the map into the Grassmann bundle of p-planes in TM determined by f. The Grassmann bundle $G_p(M)$ is endowed with a canonical exterior differential system $\mathcal{C}^{(1)}$ defined by the property that $f_*{}^*\mathcal{C}^{(1)} = 0$ for any immersion $f : W_p \to M$. Using affine fiber coordinates $(x^i, u^\alpha, u^\alpha_i)$, $1 \leq i \leq p, 1 \leq \alpha \leq n$, on the Grassmann bundle $G_p(M)$, the system $\mathcal{C}^{(1)}$ is defined as the differential ideal generated the 1-forms

$$\theta^\alpha = du^\alpha - \sum_{i=1}^{p} u^\alpha_i dx^i\,. \tag{16}$$

We choose a component $V_p(\mathcal{I})$ of the sub-variety of $G_p(M)$ defined by the p-dimensional admissible integral elements of $\mathcal{I}$ and assume $V_p(\mathcal{I})$ to be C^ω manifold.

Definition 5.1 The *first prolongation* of $\mathcal{I}$ is the exterior differential system $\mathcal{I}^{(1)}$ defined by

$$\mathcal{I}^{(1)} = \mathcal{C}^{(1)}|_{V_p(\mathcal{I})}\,.$$

For notational simplicity, we use the notation M^1 to denote the $V_p(\mathcal{I})$. We also assume that the map $\pi^{1,0} : (M^{(1)}, \mathcal{I}^{(1)}) \to (M, \mathcal{I})$ is a C^ω submersion. The prolongation tower of $\mathcal{I}$ is then defined by induction,

$$\cdots \xrightarrow{\pi^{k+1,k}} (M^{(k)}, \mathcal{I}^{(k)}) \xrightarrow{\pi^{k,k-1}} \cdots \xrightarrow{\pi^{2,1}} (M^{(1)}, \mathcal{I}^{(1)}) \xrightarrow{\pi^{1,0}} (M, \mathcal{I})\,.$$

The *infinite prolongation* $(M^{(\infty)}, \mathcal{I}^{\infty})$ of $(M, \mathcal{I})$ is then defined as the inverse limit of this tower.

$$M^{(\infty)} := \varprojlim M^{(k)},\; \mathcal{I}^{(\infty)} = \bigcup_{k \geq 0} \mathcal{I}^{(k)}\,.$$

We now present a statement of the prolongation theorem as given in [7]:

Theorem 5.2 *There exists an integer k such that for all $l \geq k$, each of the systems $(\mathcal{I}^{(l)}, \Omega^{(l)})$ is involutive. Furthermore, if $M^{(k)}$ is empty for some $k \geq 1$, then $(\mathcal{I}, \Omega)$ has no n-dimensional integral manifolds.*

6 A Cartan-Kähler Theorem for C^∞ Pfaffian systems

Our objective in this Section is to present the C^∞ Cartan-Kähler Theorem which was obtained by Yang, [48]. A geometric application of this theorem will be presented in Section 10.

We begin with some algebraic preliminaries, the purpose of which is to define the concept of an *hyperbolic determined involutive symbol*. Let B, V and W be real vector spaces of dimensions t, n and s respectively.

Definition 6.1 A symbol is a surjective homomorphism $\sigma : W \otimes V^* \to B^*$.

We consider now the flag variety $\mathcal{F}(V)$, whose elements are the complete flags

$$V_1 \subset \ldots \subset V_k \subset \ldots \subset V_n \simeq V\,, \quad \dim V_k = k\,, 1 \leq k \leq n\,.$$

By using for $1 \leq k \leq n$ the short exact sequence

$$0 \longrightarrow V_k^\perp \longrightarrow V^* \xrightarrow{\pi_k} V_k^* \longrightarrow 0\,,$$

we can define for $1 \leq k \leq n$ the symbols $\overline{\sigma}_k$ by requiring the following sequence

$$0 \longrightarrow 1 \otimes \pi_k(A) \longrightarrow W \otimes V_k^* \xrightarrow{\overline{\sigma}_k} B^*/\sigma(W \otimes V_k^\perp) \longrightarrow 0\,,$$

where

$$A = \ker \sigma\,,$$

to be exact.

For any symbol σ, we define a string of *reduced Cartan characters* $s'_1, \ldots, s'_n$ by

$$s'_k = S'_k - S'_{k-1}\,, \tag{17}$$

where

$$S_k = \dim \pi_k(A)\,, \quad \text{for} \quad 1 \leq k \leq n-1\,, \quad S_n = \dim A\,, \tag{18}$$

and

$$S'_k = \max_{\mathcal{F}} S_k \,. \tag{19}$$

The *first prolongation* $\sigma^{(1)}$ of a symbol σ is then defined as the map $\sigma^{(1)} : W \otimes S^2V^* \to B \otimes V^*$ given by

$$\sigma^{(1)} = \sigma \otimes 1_{W \otimes S^2 V^*} \,.$$

If we denote by $A^{(1)}$ the kernel of $\sigma^{(1)}$, then it can be shown [48] that

$$\dim A = \sum_{i=1}^{n} s'_i \,, \quad \dim A^{(1)} \leq \sum_{i=1}^{n} i s'_i \,.$$

This leads to define the notion of an involutive symbol.

Definition 6.2 A symbol σ is said to be involutive if

$$\dim A^{(1)} = \sum_{i=1}^{n} i s'_i \,.$$

In order to define the concept of hyperbolicity of a symbol, we introduce the very important notion of characteristic variety of a symbol. We let $V_{\mathbb{C}}, W_{\mathbb{C}}$ and $B_{\mathbb{C}}$ denote the complexifications of V, W and B respcetively, and $\mathbb{P}(V_{\mathbb{C}})$ denote the complex projective space associated to $V_{\mathbb{C}}$. For fixed $\xi \in V^* \setminus \{0\}$, we define σ_ξ to be the map from W to B^* which maps $w \in W$ to $\sigma(w \otimes \xi)$. The *characteristic variety* Θ_σ of a symbol σ is the subvariety of $\mathbb{P}(V^*_{\mathbb{C}})$ defined by the vanishing of all s by s minors of the matrix representing σ_ξ in any choice of bases for $V_{\mathbb{C}}, W_{\mathbb{C}}$ and $B_{\mathbb{C}}$. Hyperbolicity will first be defined for determined symbols. A symbol $\sigma : W \otimes V^* \to B^*$ is said to be *determined* if $\dim W = \dim B$ and $\Theta_\sigma \neq \mathbb{P}(V^*_{\mathbb{C}})$.

Definition 6.3 A determined symbol σ is said to be hyperbolic if there exists a hyperplane $H \subset V$ and a smooth map $r : \mathbb{P}H^* \to GL(W)$ such that

$$\mathbb{P}H^{\perp}_{\mathbb{C}} \cap \Theta_\sigma = \emptyset \,, \tag{20}$$

and the map

$$r(\pi[\tau])\sigma_\tau^{-1}\sigma_\xi r(\pi\xi)^{-1} \,, \tag{21}$$

is diagonal, where $\pi : \mathbb{P}V^*_{\mathbb{C}}/\mathbb{P}H^{\perp} \to \mathbb{P}H^*$ is the canonical projection, where $[\tau] \in \mathbb{P}H^{\perp}$ and where $\xi \in \mathbb{P}V^*_{\mathbb{C}}/\mathbb{P}H^{\perp}$.

A hyperplane H satisfying the conditions (20) and (21) will be called *space-like*. Finallly, we are ready to define what it means for a symbol to be involutive hyperbolic.

Definition 6.4 A symbol $\sigma : W \otimes V^* \to B^*$ is said to be involutive hyperbolic if it is involutive in the sense of Definition 6.2, and if there exists a flag

$$V_1 \subset \ldots \subset V_k \subset \ldots \subset V_n \simeq V \,, \quad \dim V_k = k \,, \quad 1 \leq k \leq n \,,$$

and decompositions

$$\sigma = \oplus_{i=1}^{n} \sigma_i \,, \quad B^* = \oplus_{i=1}^{n} W_i \,,$$

such that each map $\pi_j \circ \sigma_k \circ \pi_j : W_j \otimes V^*_k \to W_j$ is a determined hyperbolic symbol for which V_k is a space-like hyperplane.

This concludes the algebraic content of this section. We now turn our attention to differential systems of class C^∞. (Throughout this section, all the manifolds, bundles, maps and other geometric structures are assumed to be of class C^∞.) Let $\pi : E \to M$ be a fibered manifold with fibers diffeomorphic to an s-dimensional manifold F, and let B denote a vector bundle over an n-manifold M, of rank t. A *quasi-linear differential operator* is a map

$$P : J^1E \to \pi_1^* B^* ,$$

which is a bundle map over E. The *first prolongation* of P is an affine bundle map

$$p^1P : J^2E \to \pi_2^* J^1 B^* ,$$

such that

$$p^1P(x,(j^2u)(x)) = j^1[P(x,(j^1u)(x))] .$$

The local coordinate expressions of P and p^1P are easily determined. We have,

$$p^1P_{(x,u)} = P_{(x,u)} + P^{(1)}_{(x,u)}$$

where

$$P_{(x,u)} = \sum_{i=1}^{n} a^i(x,u)p^i + b(x,u) ,$$

and

$$P^{(1)}_{(x,u)} = \sum_{i,j=1}^{n}\sum_{\alpha=1}^{s}[a^i p_{ij} + \frac{\partial a^i}{\partial u^\alpha}p^\alpha_j + (\frac{\partial a^i}{\partial x^j}p_i + \frac{\partial b}{\partial x^j})]dx^j ,$$

and where we have used the identifications

$$\begin{aligned} J^1_xB^* &\simeq B^*_x \oplus B^*_x \otimes T^*_xM , \\ J^2_{(x,u)}E &\simeq T_uE_x \otimes T^*_xM \oplus T_uE_x \otimes S^2T^*_xM . \end{aligned}$$

Given $(f, \sum_{i=1}^n f_i dx^i) \in \operatorname{im} p^1P$, we define

$$\begin{aligned} S_{(x,u)}(f) &= P^{-1}_{(x,u)}(f) , \\ p^1S_{(x,u)}(f, \sum_{i=1}^n f_i dx^i) &= (p^1P)^{-1}(f, \sum_{i=1}^n f_i dx^i) . \end{aligned}$$

One shows [48] that

$$\dim S_{(x,u)}(f) = \sum_{i=1}^n s'_i(x,u) ,$$

and that the projection $\pi_1^2 : J^2E \to J^1E$ induces a map $g : p^1S_{(x,u)}(f, \sum_{i=1}^n f_i dx^i) \to S_{(x,u)}(f)$ whose fibers are affine subspaces F such that

$$\dim F \leq \sum_{i=1}^n i s'_i(x,u) .$$

Definition 6.5 A quasi-linear differential operator $P : J^1E \to \pi_1^* B^*$ is said to be involutive if the characters $s_i'(x,u)$ are constant in E and the map g is surjective with fibers of maximal dimension,

$$\dim F = \sum_{i=1}^{n} i s_i'(x,u)\,.$$

Quasi-linear differential operators which are involutive hyperbolic are now defined as follows:

Definition 6.6 A quasi-linear differential operator $P : J^1E \to \pi_1^* B^*$ is said to be involutive hyperbolic if its symbol $\sigma_{(x,u)} : T_uE_x \otimes T_x^*M \to B_x^*$ is an involutive hyperbolic symbol for all $(x,u) \in E$, and if the corresponding splittings and diagonalizing maps can be chosen to vary smoothly with $(x,u) \in E$.

We now have the following local existence theorem on the local solvability of involutive hyperbolic quasi-linear differential operators.

Theorem 6.7 *Let $P : J^1E \to \pi_1^* B^*$ be an involutive hyperbolic quasi-linear differential operator, and let f be a C^∞ section of B^* such that for all $(x,u) \in E$, we have*

$$(j^1 f)(x) \in p^1 P(J^1_{(x,u)}E)\,.$$

Then there exist C^∞ solutions of the partial differential equation

$$Pu = f\,,$$

in a sufficiently small neighborhood of any $x_0 \in M$.

We now show how Theorem 6.7 can be used to obtain a Cartan-Kähler for a certain class of C^∞ quasi-linear Pfaffian systems. Recall that a Pfaffian system with independence condition $(\mathcal{I}, \Omega)$, where

$$\mathcal{I} = \{\theta^1, \dots, \theta^s, d\theta^1, \dots, d\theta^s\},$$
$$\Omega = \omega^1 \wedge \dots \wedge \omega^p,$$

is quasi-linear if

$$d\theta^i \equiv \sum_{\alpha=1}^{l} \sum_{b=1}^{p} A^i_{\alpha b} \pi^\alpha \wedge \omega^b + \frac{1}{2} \sum_{a,b=1}^{p} B^i_{ab} \omega^a \wedge \omega^b \mod \theta^1, \dots, \theta^s$$

where $\pi^1, \dots, \pi^l$, $l = n - s - p$, are 1-forms such that $\theta^1 \wedge \dots \wedge \theta^s \wedge \omega^1 \wedge \dots \wedge \omega^p \wedge \dots \wedge \pi^1 \wedge \dots \wedge \pi^l \neq 0$. The symbol of a quasi-linear Pfaffian system is now defined as follows. Let W^* denote the sub-bundle of T^*M spanned by $\theta^1, \dots, \theta^s$, $V*$ denote the sub-bundle of T^*M spanned by $\omega^1, \dots, \omega^p$ and let B^* be the quotient bundle defined by

$$B^* = W \otimes V^*/A\,,$$

where A denotes the image in $W \otimes V^*$ of the homomorphism defined by $(A^i_{\alpha b})$. (All maps are assumed to be of constant rank.) The symbol σ of the Pfaffian system $(\mathcal{I}, \Omega)$ is defined as the quotient map $\sigma : W \otimes V^* \to B^*$. It is easily verified that the reduced characters of the Pfaffian system $(\mathcal{I}, \Omega)$ are equal to the reduced characters of its symbol, defined by (17), (18) and (19). A quasi-linear Pfaffian systems with independence condition $(\mathcal{I}, \Omega)$ is now said to be involutive hyperbolic if its symbol is involutive hyperbolic. The Cartan-Kähler Theorem for quasi-linear Pfaffian systems with independence condition is as follows [48]:

Theorem 6.8 *A C^∞ quasi-linear Pfaffian systems with independence condition $(\mathcal{I}, \Omega)$ on a C^∞ manifold has admissible integral manifolds if it has an involutive hyperbolic prolongation.*

7 Characteristic cohomology

The characteristic cohomology theory for exterior differential systems is a powerful generalization of the classical theory of conservation laws for partial differential equations. It was defined in a foundational paper by Bryant and Griffiths, [8], and studied in detail in a number of other papers, including [9], [10], [11], [23], [46]. The work of Vinogradov [45] on the C-spectral sequence can be thought of as a precursor of this theory, but it is Bryant and Griffiths who discovered the general formulation of this cohomology theory in the context of exterior differential systems, making full use of the local invariants provided by the method of equivalence of Cartan, [26], to perform a refined analysis of this cohomology for various classes of differential systems. Our purpose in this section is to present some of the basic definitions of the theory, and some general vanishing theorems for characteristic cohomology of certain general classes of differential systems. More precise theorems on the structure of the characteristic cohomology for specific classes of differential systems will be presented in Section 9.

Recall from Section 5 that to any given exterior differential system $\mathcal{I}$ on a manifold M, one can associate a prolongation tower and an infinite prolongation $(M^{(\infty)}, \mathcal{I}^{(\infty)})$. On $M^{(\infty)}$, we define a differential complex $(\bar{\mathcal{I}}^{(\infty)*}, \bar{d})$, where

$$\bar{\mathcal{I}}^{(\infty)*} := \Omega^*(M^{(\infty)})/\mathcal{I}^{(\infty)}, \, \bar{d} := d \mod \mathcal{I}^{(\infty)} \, .$$

Definition 7.1 The characteristic cohomology $\bar{H}^*(\mathcal{I})$ of $\mathcal{I}$ is defined as the cohomology of the differential complex $(\bar{\mathcal{I}}^{(\infty)*}, \bar{d})$.

Consider first the case where $\mathcal{I}$ is the empty exterior differential system on M whose integral manifolds are all the immersions from a manifold W into M. The manifolds $M^{(k)}$ appearing in the prolongation tower are the sets of all k-jets of immersions of W into M and the exterior differential system $\mathcal{I}^{(\infty)}$ endowed with $\bar{d}$ can be roughly thought of as a generalization to the context of exterior differential systems of the filtered complex associated by vertical grading to the tautological variational bi-complex of a fibered manifold, [45]. Similarly, the differential complex $(\bar{\mathcal{I}}^{(\infty)*}, \bar{d})$ associated to an exterior diffferential system can be viewed generalization of the filtered complex associated to the constrained variational bi-complex of a partial differential equation $\mathcal{R}^{(\infty)} \subset J^\infty(E)$. One should therefore expect that there will be some general vanishing theorems for the characteristic cohomology which will correspond to the generalized Poincaré Lemmas which hold true for the variational bi-complex. This is indeed the case, [8]:

Theorem 7.2 *For the empty exterior differential system on M, we have*

$$\bar{H}^q = 0 \, , \quad 0 < q < n \, .$$

We also have vanishing theorems which depend on the fairly ”coarse” data provided by the characters of a general involutive Pfaffian system with independence condition [8]:

Theorem 7.3 *Let $(\mathcal{I}, \Omega)$ be an involutive Pfaffian system with independence condition exterior differential system, with reduced characters s'_i satisfying*

$$s'_0 = s'_1 = \cdots = s'_{p-l-1} = 0,$$

where p denotes the dimension of the admissible integral manifolds of maximal dimension. Then we have

$$\bar{H}^q(\mathcal{I}^{(\infty)}) = 0, \quad 0 < q < p - l.$$

For $l = 1$, one recovers Vinogradov’s ”two line” theorem, [45]. We should point out that V. Itskov, [31], has developed an equivariant version of the characteristic cohomology theory for exterior differential systems with symmetry.

Beyond these general vanishing theorems, there are a number of more specific theorems dealing with the characteristic cohomology of exterior differential systems associated to differential equations, [9], [10], [11], [12], [23], [46]. A very attractive theme in these works is that properties such as the existence of equations with infinitely many non-trivial conservation laws (corresponding to non-zero cohomology classes in their characteristic cohomology) can be detected through the behavior of local invariants computed by means of Cartan’s method of equivalence [26].

8 Topological obstructions

Even though the results reviewed in the preceding sections were stated for differential systems defined on manifolds, they are still largely local in nature. It is therefore natural to ask whether there are topological obstructions the global existence of exterior differential systems of a given kind. This is a broad and difficult question, which touches on several important research areas, notably the classification of foliated, contact and symplectic manifolds. Our goal in this section is to review some of the elementary classical results concerning the topological obstructions the existence of completely integrable Pfaffian systems, or certain special Pfaffian systems of codimension two. These obstructions are obtained through the classical tool provided by Chern-Weil theory, [22].

The topological implications of the global existence of completely integrable systems were brought to light by Bott, [3], in the context of his work on foliations. It is indeed a remarkable fact that the existence of a globally defined completely integrable system on a compact manifold puts topological restrictions expressible in terms of the vanishing of certain Pontrjagin classes in the cohomology ring of M_n. An n-dimensional manifold M_n admitting an $(n-r)$-dimensional foliation $\mathcal{F}$ will by definition be endowed with a completely integrable Pfaffian system I of dimension r. The tangent space to the leaf L_{n-r} of $\mathcal{F}$ through a point x of M_n is given by

$$T_x L_{n-r} = I_x^{\perp} \,.$$

Bott’s vanishing theorem, [3], reads as follows:

Theorem 8.1 *If I is an r-dimensional completely integrable Pfaffian system on a compact n-manifold M_n, then*

$$p_k(I) = 0, \quad \textit{for all } k > 2r \,. \tag{22}$$

Equivalently, the vanishing condition (22) can be stated in terms of the quotient bundle $TM_n/I^\perp$ as

$$p_k(TM_n/I^\perp) = 0, \quad \text{for all } k > 2r\,.$$

A proof of Theorem 8.1 is given by Chern in [21]. One first observes that the existence of a globally defined sub-bundle I of T^*M_n gives rise to a reduction of the coframe bundle of M_n to a sub-bundle whose structure group G consists of block-triangular matrices. One then defines an adapted connection taking values in the differential ideal $\mathcal{I}$. Its curvature 2-form Ω will also take values in $\mathcal{I}$ and therefore any polynomial of degree $l > n-k$ in Ω will be identically zero.

It is natural to ask if obstructions similar to the ones given in Bott's Theorem exist for Pfaffian systems which are not completely integrable. A result of this type was obtained by Buemi [15].

Theorem 8.2 *If I is an r-dimensional Pfaffian system of type $(1,\ldots,1,2)$ on a compact $(r+2)$-manifold M_{r+2}, then*

$$p_k(M_{r+2}) = 0, \quad \textit{for all } k \geq 0\,.$$

The method of proof of Theorem 8.2 is similar to the one we sketched for Theorem 8.1, the point being that in the case of Theorem 8.2, the structure group reduces to a discrete group of diagonal matrices with diagonal entries being equal to ± 1.

Example 8.3 It is important to observe that the vanishing Theorem 8.2 is not merely a consequence of the fact that M_{n+2} admits a 2-plane field whose annihilator splits into a direct sum of line bundles. This is nicely illustrated by the following example, [15]. Consider the complex surface $M_4 = S^2 \times M_2$, where M_2 is a Riemann surface of genus 2. The Euler characteristic $\chi(M)$ is equal to -4 and the index $I(M_4)$ of the intersection form of M equals 0. If we blow up M at four distinct points, we obtain a complex surface $\tilde{M}_4$ with $\chi(\tilde{M}_4) = 0$ and $I(\tilde{M}_4) \neq 0$. By the Hirzebruch signature formula, it follows that $p_1(\tilde{M}_4) \neq 0$. Since $\chi(\tilde{M}_4) = 0$, we know that $\tilde{M}$ will admit a non-vanishing vector field X. Using the complex structure J of $\tilde{M}_4$, we obtain another vector field $Y = J(X)$, such that $X \wedge Y \neq 0$. Choose now a Riemannian metric on $\tilde{M}_4$ and let $\mathcal{E}$ be the 2-plane field orthogonal to $\mathcal{D}$ relative to this metric. The annihilator of the distribution $\mathcal{E}$ in $T^*\tilde{M}_4$ is a sub-bundle which is isomorphic to $\mathcal{D}$ and therefore splits into a direct sum of two line bundles. Yet, we have $p_1(\tilde{M}_4) \neq 0$.

It is a remarkable fact that similar topological obstructions can be obtained for any Pfaffian system, just based on its class. Martinet [39] proved the following:

Theorem 8.4 *If I is a Pfaffian system of class less or equal than r on a compact n-manifold M_n, then*

$$p_k(M_n) = 0, \quad \textit{for all } k > 2r\,.$$

9 Applications to second-order scalar hyperbolic partial differential equations in the plane

In this section, we present some more applications of the existence and normal form results of Section 3 to local normal form results, specifically to the case of the contact geometry of hyperbolic second-order partial differential equations in the plane. These results are elementary in nature since they do not call on a complete knowledge of the local invariants of this class of systems. The results of this section are taken from [27]. We refer the reader to [10], [11], [1], [28] for further results on hyperbolic systems, including the computation of the characteristic cohomology and a study of Darboux integrability.

Recall from Example 2.8 that to any second-order partial differential equation

$$F(x, y, u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}) = 0\,, \tag{23}$$

one associates the Pfaffian system

$$\mathcal{I} = \{\, \theta^1, \theta^2, \theta^3, d\theta^1, d\theta^2, d\theta^3, dF\}\,,$$

where

$$\theta^1 = du - pdx - qdy\;,\; \theta^2 = dp - rdx - sdy\;,\; \theta^3 = dq - sdx - tdy\;,$$

We assume the partial differential equation (23) to be hyperbolic, meaning that

$$F_{u_{xx}} F_{u_{yy}} - \frac{1}{4} F^2_{u_{xy}} < 0\,, \tag{24}$$

on the locus $F = 0$ of $\mathbb{R}^8$. The existence theorems of Section 3 can be applied to recover some classical normal form results for equations of the form (23) under the pseudo-group of contact transformations. Instead of working on $\mathbb{R}^8$, we work instead on the subset M_7 of $\mathbb{R}^8$ defined by (23), which we assume to be a C^∞ hypersurface, and consider the Pfaffian system generated by the pull-backs of the 1-forms $\theta^1, \theta^2, \theta^3$ to M_7. We will commit a slight abuse of notation and also denote this Pfaffian system by $\mathcal{I}$. We first recall from [27] that a local coframe can be chosen on M_7 so that the structure equations of $\mathcal{I}$ take a simple form:

Proposition 9.1 *Locally, there exists a coframe* $(\omega^1, \pi^2, \pi^3, \omega^6, \omega^7)$ *on* M_7 *such that* $\mathcal{I} = \{\omega^1, \pi^2, \pi^3\}$ *and*

$$d\omega^1 \equiv 0\,, \quad d\pi^2 \equiv \omega^4 \wedge \omega^5\,, \quad d\pi^3 \equiv \omega^6 \wedge \omega^7\,,$$

where the congruences are modulo $\{\omega^1, \pi^2, \pi^3\}$.

The Pfaffian subsystems $M_1 = \{\omega^1, \pi^2\}$ and $M_2 = \{\omega^1, \pi^3\}$ are invariants of the contact geometry of (23) arising from the hyperbolicity condition (24). The following result gives a characterization of Monge-Ampère equations, [27]:

Proposition 9.2 *A hyperbolic equation*

$$F(x, y, u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}) = 0\,,$$

is equivalent to a hyperbolic equation of Monge-Ampère form

$$a(u_{xx}u_{yy} - u^2_{xy}) + bu_{xx} + 2cu_{xy} + du_{yy} = 0\,,$$

where a, b, c, d *are functions of* (x, y, u, u_x, u_y)*, if and only if* M_1 *and* M_2 *are both of class six.*

We define the M_i-characteristic systems $C(I, dM_i)$ of $\mathcal{I}$ by

$$C(I, dM_i) = \text{Char}(I, dM_i)^{\perp}\,, i = 1, 2\,,$$

where

$$\text{Char}(I, dM_i) = \{X \in \Gamma(TM_7)\,|\, X \in I^{\perp}\,,\, X\rfloor dM_i \subset I\}\,.$$

One can obtain some interesting normal form results for hyperbolic Monge-Ampère equations based on the derived flag structure of the characteristic systems $C(I, dM_i)\,, i = 1, 2$. The following result, due in its original form to Lie, [38], gives a characterization of the contact orbit of the wave equation $u_{xy} = 0$.

Theorem 9.3 *A hyperbolic equation*

$$F(x, y, u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}) = 0\,,$$

is contact equivalent to the wave equation

$$u_{xy} = 0\,,$$

if and only if each of the M_i-characteristic systems is of class six and has a completely integrable second derived system, that is

$$C(I, dM_1)^{(2)} = C(I, dM_1)^{(3)}\,, \quad C(I, dM_2)^{(2)} = C(I, dM_2)^{(3)}\,.$$

A proof of Lie's theorem based on the constructive proof of Theorem 3.8 is given in [27].

Corollary 9.4 *A hyperbolic equation*

$$u_{xy} = f(x, y, u, u_x, u_y) = 0\,,$$

is contact equivalent to the wave equation $u_{xy} = 0$ if and only if f satisfies

$$f_{u_x u_x} = 0\,, \quad -f_{u_x} f_{u_y} + f_{u_x u_y} - f_u + u_x f_{u_x u} + f_{u_x y} = 0\,,$$

and

$$f_{u_y u_y} = 0\,, \quad -f_{u_x} f_{u_y} + f_{u_x u_y} - f_u + u_y f_{u_y u} + f_{u_y y} = 0\,.$$

10 Some applications to differential geometry

The theory of exterior differential systems has numerous applications to classical differential geometry, many of which were beautifully presented by Cartan himself in [20]. The book [32] also contains a number of very nice elementary geometrical examples. There are of course very sophisticated and important applications of exterior differential systems in differential geometry, most notably to the rigidity of isometric embeddings, exceptional holonomy, [4], Bochner-Kähler metrics, [5], minimal submanifolds, [37], calibrated submanifolds, [41] and pseudoholomorphic curves, [40]. We will not consider these topics here, and focus instead on a few classical examples, which will serve to illustrate how the Cartan-Kähler theorem and the involutivity test can be applied in concrete situations.

10.1 Cartan submanifolds

The first class of examples we will consider are the class of Cartan submanifolds:

Definition 10.1 A Riemannian manifold (M_n, g) isometrically immersed in $\mathbb{P}^{2n}$ is said to be a Cartan submanifold if the second-order osculating space of M_n is everywhere 2n-dimensional, and if near every point $x \in M_n$ there exist local coordinates $(x^1, \ldots, x^n)$ such that the net of coordinate curves is conjugate.

This means that if $\mathbf{X} : U \subset \mathbb{R}^n \to \mathbb{R}^{2n}$ is the local coordinate expression of such an immersion, then

$$\mathbf{X}_{ij} = \sum_{k=1}^{n} \Gamma^k_{ij} \mathbf{X}_k + \sum_{\alpha=1}^{n} \Omega^{\alpha}_{ij} \mathbf{N}_{\alpha}, \quad 1 \leq i, j \leq n\,,$$

where the $\mathbf{N}_\alpha$ are normal vector fields, and where

$$\Omega^{\alpha}_{ij} = 0, \quad 1 \leq i \neq j \leq n, \quad 1 \leq \alpha \leq n\,.$$

Cartan submanifolds are at the basis of the higher-dimensional generalization of the transformation theory of Laplace for linear hyperbolic second-order equations, [34]. We now give two examples of Cartan submanifolds.

Example 10.2 The Clifford torus $\mathbb{T}^n \subset \mathbb{R}^{2n}$, given by

$$X(x^1, \ldots, x^n) = (\cos x^1, \sin x^1, \ldots, \cos x^n, \sin x^n)\,,$$

is a Cartan submanifold.

Example 10.3 Let $c_3, \ldots, c_n$ be real numbers such that

$$\sum_{j=3}^{n} c_j^2 < 1, \quad c_j \neq 0\,,$$

and let

$$\lambda := (1 - \sum_{j=3}^{n} c_j^2)^{\frac{1}{2}}\,,$$

and let r and μ be non-zero real numbers such that $r^2 = \mu^2 + 1$. The toroidal submanifold of $S^{2n-1} \subset \mathbb{R}^{2n}$ given by, [42],

$$\mathbf{X} = (\lambda f_0, \lambda f_1, \lambda f_2 \cos x^2, \lambda f_2 \sin x^2, c_3 \cos x^3, c_3 \sin x^3, \ldots, c_n \cos x^n, c_n \sin x^n)\,,$$

where

$$0 < x^1 < \frac{\pi}{2\mu}, \quad 0 < x^j < 2\pi, \quad 2 \leq j \leq n\,,$$

and

$$f_0 = \frac{\mu}{r} \sin r x^1 \cos \mu x^1 - \cos r x^1 \sin \mu x^1\,,$$

$$f_1 = \frac{\mu}{r}\cos rx^1 \cos \mu x^1 - \sin rx^1 \sin \mu x^1 \,,$$
$$f_2 = -\frac{1}{r}\cos \mu x^1 \,,$$

is a Cartan submanifold.

A very natural question is to determine the degree of generality of the set of Cartan submanifolds, given the dimension of the ambient space. This question was studied by Elie Cartan, [19] by means of the Cartan-Kähler theorem and the involutivity test for exterior differential systems. We now state Cartan's theorem for submanifolds of $\mathbb{P}^{2n}(\mathbb{R})$ and sketch its proof.

Theorem 10.4 *Locally, the set of Cartan submanifolds* $\mathbf{X} : U \subset \mathbb{R}^n \to \mathbb{P}^{2n}$ *of class* C^ω *is parametrized by* $n(n-1)$ *functions of 2 variables.*

Proof. We now sketch the proof. The starting point is given by the structure equations for the bundle of projective frames in $\mathbb{P}^{2n}$, which are given by

$$d\mathbf{A} = \omega_{00}\mathbf{A} + \omega_1\mathbf{A}_1 + \cdots + \omega_{2n}\mathbf{A}_{2n},$$
$$d\mathbf{A}_i = \omega_{i0}\mathbf{A} + \omega_{i1}\mathbf{A}_1 + \cdots + \omega_{i2n}\mathbf{A}_{2n},$$
$$d\omega_i = \omega_{00} \wedge \omega_i + \sum_{j=1}^{2n} \omega_j \wedge \omega_{ji},$$
$$d\omega_{ij} = \omega_{i0} \wedge \omega_j + \sum_{k=1}^{2n} \omega_{ik} \wedge \omega_{kj},$$

where $1 \le i, j \le 2n$. The projective frame $(\mathbf{A}, \mathbf{A}_1, \ldots, \mathbf{A}_{2n})$ to $\mathbf{X}$, can be chosen in such a way that at every point of the image of $\mathbf{X}$, we have

$$\omega_{n+1} = \cdots = \omega_{2n} = 0\,, \quad \omega_{i\alpha} = \sum_{j=1}^{n} \Omega_{ij}^{\alpha-n}\omega_j\,,$$

where $1 \le i \le n$, $n+1 \le \alpha \le 2n$, and

$$\Omega_{ij}^{\alpha-n} = \Omega_{ji}^{\alpha-n}.$$

We now consider the pencil of quadratic differential forms

$$\Omega = \sum_{\alpha=n+1}^{2n} \lambda_\alpha \Omega_\alpha\,,$$

where

$$\Omega_\alpha = \sum_{i=1}^{n} \omega_i \otimes \omega_{i\alpha} = \sum_{i,j=1}^{n} \Omega_{ij}^{\alpha-n}\omega_i \otimes \omega_j, \quad n+1 \le \alpha \le 2n\,.$$

We are going to construct an exterior differential system on the the bundle of projective frames, whose integral manifolds will correspond to the immersions $\mathbf{X}$ such that

$$\Omega = \sum_{i=1}^{n} \nu_i\, \omega_i^2\,. \tag{25}$$

The projective frame $(\mathbf{A}, \mathbf{A}_1, \ldots, \mathbf{A}_{2n})$ can be adapted further, in such a way that the condition (25) reduces to

$$\Omega_{n+i} = \omega_i^2,$$

and the problem thus reduces to constructing the integral manifolds of the Pfaffian system with independence condition on the bundle of projective frames generated by the 1-forms

$$\begin{aligned} &\omega_\alpha, \quad n+1 \leq \alpha \leq 2n, \\ &\omega_{i,n+i} - \omega_i, \quad 1 \leq i \leq n, \\ &\omega_{j,n+i}, \quad 1 \leq i \neq j \leq n, \end{aligned}$$

with independence condition given by $\omega_1 \wedge \cdots \wedge \omega_n$. One verifies that this system is in involution, with reduced Cartan characters given by

$$s_1' = 0, \quad s_2' = n(n-1), \quad s_{2+l}' = 0.$$

The conclusion follows by applying Cartan's involutivity test, [20], [7]. □

10.2 Orthogonal coordinates for Riemannian metrics

It is a classical theorem of Darboux that given a three-dimensional Riemannian metric of class C^ω, one can for any point find an open coordinate neighborhood in which the metric is diagonal. Cartan's treatise [20] gives a proof of this result using the Cartan-Kähler Theorem. We show how this is done, following the presentation of [7]. We let (M_n, g) be a Riemannian manifold consider on the orthonormal frame bundle $\mathbb{F}(M_n)$ a coframe $\omega_i\,, \omega_{ij}\,, 1 \leq i, j, \leq n$ satisfying the structure equations defining the Levi-Civitá connection and the Riemann curvature of g,

$$\begin{aligned} d\omega_i &= -\sum_{j=1}^{n} \omega_{ij} \wedge \omega_j\,, \\ d\omega_{ij} &= -\sum_{k=1}^{n} \omega_{ik} \wedge \omega_{kj} + \frac{1}{2} R_{ijkl} \omega_k \wedge \omega_l\,. \end{aligned}$$

On $\mathbb{F}(M_n)$, we consider the exterior differential system with independence condition $(\mathcal{I}, \Omega)$, where $\mathcal{I}$ is the differential ideal generated by the three 3-forms

$$\omega_i \wedge d\omega_i\,, \quad 1 \leq i \leq 3\,,$$

with independence condition given by

$$\Omega = \omega_1 \wedge \ldots \wedge \omega_n\,.$$

Every admissible integral manifold W of $(\mathcal{I}, \Omega)$ is locally a section s of $\pi : \mathbb{F}(M_n) \to M_n$ such that the 1-forms $\eta_i = s^* \omega_i$ satisfy

$$\eta_i \wedge d\eta_i = 0\,, \quad 1 \leq i \leq 3\,.$$

By the Frobenius theorem 3.1, any such a section will give rise in its domain of definition to local coordinates $(u^1, \dots, u^n)$ in which

$$ds^2 = \sum_{i=1}^{n} h_i(u^1, \dots, u^n)\,(du^i)^2\,,$$

that is an orthogonal local coordinate system. The following theorem asserts that such sections always exist when $n = 3$:

Theorem 10.5 *Let (M_3, g) be a three-dimensional Riemannian metric of class $C\omega$. For any $x \in M_3$, there exists an open coordinate neighborhood with local coordinates $(u^1, \dots, u^3)$ in which the metric is diagonal, that is*

$$ds^2 = \sum_{i=1}^{3} h_i(u^1, u^2, u^3)\,(du^i)^2\,.$$

Using the structure equations, it is easy to characterize admissible the n-dimensional integral elements of $(\mathcal{I}, \Omega)$. We have,

Lemma 10.6 *An n-dimensional subspace E of $T_z\mathbb{F}(M_n)$ defines an admissible n-dimensional integral element of $(\mathcal{I}, \Omega)$ based at $z \in \mathbb{F}(M_n)$ if an only if it its annihilator is given by equations of the form*

$$\omega_{ij} + l_{ij}\omega_i - l_{ji}\omega_j = 0 \quad 1 \le i \ne j \le n\,,$$

where $l_{ij}\,, 1 \le i \ne j \le n$, are $n^2 - n$ arbitrary real constants.

It follows that the set of admissible integral elements of $(\mathcal{I}, \Omega)$ based at $z \in \mathbb{F}(M_n)$ is an analytic manifold of dimension $n^2 - n$. Therefore $\mathcal{V}_n(\mathcal{I}, \Omega)$ is an analytic manifold whose dimension is given by

$$\dim \mathcal{V}_n(\mathcal{I}, \Omega) = \frac{n(n+1)}{2} + n^2 - n = \frac{1}{2}(3n^2 - n)\,.$$

The codimension c of $\mathcal{V}_n(\mathcal{I}, \Omega)$ in the Grassmann bundle $G_n(\mathbb{F}(M_n))$ is thus given by

$$c = \frac{1}{2}((n-2)(n-1)n)\,. \tag{26}$$

Lemma 10.7 *The system $(\mathcal{I}, \Omega)$ is in involution for $n = 3$.*

Proof. Again, we sketch the proof. By Theorem 4.8, it suffices to show that there exists a flag

$$\{0\} \subset E^1 \subset E^2 \subset E^3\,,$$

which is ordinary, and by Theorem 4.9 and (26), this is equivalent to showing that the sum of the codimensions (c_0, c_1, c_2) satisfies the equality stated in the involutivity criterion (9), that is,

$$c_0 + c_1 + c_2 = 3\,,$$

or simply that $c_2 = 3$, since $\mathcal{I}$ does not contain any forms of degree less than 3. This is easily done by choosing E_2 to be a 2-plane on which

$$\omega_2 \wedge \omega_3 \neq 0\,, \quad \omega_{21} \wedge \omega_3 \neq 0\,, \quad \omega_{12} \wedge \omega_2 \neq 0\,.$$

□

On the other hand, it is not too hard to see that $(\mathcal{I}, \Omega)$ does not contain any ordinary integral elements for $n \geq 4$. Indeed, if we consider a flag

$$E^0 \subset E^1 \subset E^2 \subset \cdots \subset E^{p-1} \subset E^p\,,$$

we have $c_0 = c_1 = 0$ and we obtain the bounds

$$c_2 \leq n \quad c_k \leq \frac{1}{2} n(n-1)\,,$$

on the codimensions, so that

$$\sum_{i=0}^{n-1} c_i \leq \frac{1}{2} n(n^2 - 4n + 5).$$

But we have

$$\operatorname{codim} \mathcal{V}(\mathcal{I}, \Omega) = \frac{1}{2} n(n^2 - 3n + 2)$$

so that the inequality (8) is always strict for $n \geq 4$, as claimed. We conclude by remarking that Theorem 10.5 is also valid for C^∞ metrics, [24].

Bibliography

[1] I. Anderson and N. Kamran: The variational bicomplex for hyperbolic second-order scalar partial differential equations in the plane *Duke Math. J.* **87(2)** (1997) 265–319

[2] E. Berger, R. Bryant and P. Griffiths: The Gauss equations and rigidity of isometric embeddings *Duke Math. J.* **50(3)** (1983) 803–892

[3] R. Bott: On a topological obstruction to integrability *In: Global Analysis* (Proc. Symp. in Pure Math. 16, American Mathematical Society, Providence, 1970) 127–131

[4] R. Bryant: Metrics with exceptional holonomy *Ann of Math. (2)* **126(3)** (1987) 525–576

[5] R. Bryant: Bochner-Kähler metrics *J. Amer. Math. Soc.* **14(3)** (2001) 623–715

[6] R. Bryant, S. S. Chern and P. Griffiths: Exterior differential systems *In: Proceedings of the 1980 Beijing Symposium on Differential Geometry and Differential Equations, Vol. 1, 2, 3* (Beijing 1980, Science Press, 1982) 219–338

[7] R. Bryant, S. S. Chern, R. Gardner, H. Goldschmidt and P. Griffiths: *Exterior differential systems* (MSRI Publications, 18, Springer-Verlag, New York, 1991)

[8] R. Bryant and P. Griffiths: Characteristic cohomology of differential systems. I. General theory *J. Amer. Math. Soc.* **8(3)** (1995) 507–596

[9] R. Bryant and P. Griffiths: Characteristic cohomology of differential systems. II. Conservation laws for a class of parabolic equations *Duke Math. J.* **78(3)** (1995) 531–676

[10] R. Bryant, P. Griffiths and L. Hsu: Hyperbolic exterior differential systems and their conservation laws. I *Selecta Math. (N.S.)* **1(1)** (1995) 21–112

[11] R. Bryant, P. Griffiths and L. Hsu: Hyperbolic exterior differential systems and their conservation laws. II *Selecta Math. (N.S.)* **1(2)** (1995) 265–323

[12] R. Bryant, P. Griffiths and L. Hsu: Toward a geometry of differential equations *In: Geometry, topology, and physics* (Conf. Proc. Lecture Notes Geom. Topology, IV, Internat. Press, Cambridge, MA, 1995) 1–76

[13] R. Bryant, P. Griffiths and D. Yang: Characteristics and existence of isometric embeddings *Duke Math. J.* **50(4)** (1983) 893–994

[14] R. Bryant and L. Hsu: Rigidity of integral curves of rank 2 distributions *Invent. Math.* **114(2)** (1993) 435–461

[15] R. Buemi: An obstruction to certain non-integrable 2-plane fields *Topology* **16** (1977) 173–176

[16] E. Cartan: Sur quelques quadratures dont l'élément différentiel contient des fonctions arbitraires *Bull. Soc. Math. France* **29** (1901) 118–130

[17] E. Cartan: Les systèmes de Pfaff à cinq variables et les équations aux dérivées partielles du second ordre *Ann. Ec. Normale* **27** (1910) 109–192

[18] E. Cartan: Sur l' intégration de certains systèmes indéterminés d' équations différentielles *J. für die reine und angew. Math.* **145** (1915) 86–91

[19] E. Cartan: Sur les variétés à courbure constante d'un espace euclidien ou non-euclidien *Bull. Soc. Math. France* **77** (1919) 125–160

[20] E. Cartan: *Les systèmes différentiels extérieurs et leurs applications géométriques* (Hermann, Paris, 1945)

[21] S. S. Chern: *Complex manifolds without potential theory* (Springer-Verlag, New York, 1979)

[22] C. Chevalley: *Theory of Lie groups I* (Princeton University Press, Princeton, 1946)

[23] J. Clelland: Geometry of conservation laws for a class of parabolic partial differential equations *Selecta Math. (N.S.)* **3(1)** (1997) 1–77

[24] D. de Turck and D. Yang: Existence of elastic deformations with prescribed principal strains and triply orthogonal systems *Duke Math. J.* **51(2)** (1984) 243–260

[25] R. B. Gardner: Invariants of Pfaffian systems *Trans. Amer. Math. Soc.* **126** (1967) 514–533

[26] R. B. Gardner: *The method of equivalence and its applications* (CBMS Conference Series in Applied Mathematics 58, SIAM, Philadelphia, 1989)

[27] R. B. Gardner and N. Kamran: Characteristics and the geometry of hyperbolic equations in the plane *J. Differential Equations* **104(1)** (1993) 60–116

[28] R. B. Gardner and N. Kamran: Normal forms and focal systems for determined systems of two first-order partial differential equations in the plane *Indiana Univ. Math. J.* **44(4)** (1995) 1127–1162

[29] P. Griffiths: *Exterior differential systems and the calculus of variations* (Progress in Mathematics 25, Birkhäuser Boston, 1983)

[30] D. Hilbert: Über den Begriff der Klasse von Differentialgleichungen *In: Math. Ann.* (73, 1912) 81–93; D. Hilbert: *Gesammelte Abhandlungen 3* (Springer-Verlag, Berlin, 1935)

[31] V. Itskov: Orbit reduction of contact ideals *In: The geometrical study of differential equations* (Washington, DC, 2000, Contemp. Math. 285, Amer. Math. Soc., Providence, 2001) 171–181

[32] T. Ivey and J. M. Landsberg: *Cartan for beginners: differential geometry via moving frames and exterior differential systems, Graduate Studies in Mathematics* (61 American Mathematical Society, Providence, 2003)

[33] E. Kähler: *Einführung in die Theorie der Systeme von Differentialgleichungen* (Teubner, Leipzig, 1934)

[34] N. Kamran and K. Tenenblat: Laplace transformation in higher dimensions *Duke Math. J.* **84(1)** (1996) 237–266

[35] N. Kamran: *Selected topics in the geometrical study of differential equations* (CBMS Conference Series in Mathematics 96, American Mathematical Society, Providence, 2002)

[36] M. Kuranishi: On. E. Cartan's prolongation theorem for exterior differential systems *Amer. J. Math.* **79** (1957) 1–47

[37] J. M. Landsberg: Minimal submanifolds defined by first-order systems of PDE *J. Differential Geom.* **36(2)** (1992) 369–415

[38] S. Lie: Diskussion der differentialgleichung s=F(z) *Arc. Mat.* **6(2)** (1881) 112-124

[39] J. Martinet: Classes caractéristiques des systèmes de Pfaff *In: Géométrie diiférentielle* (Lecture Notes in Mathematics 392, Springer-Verlag, Berlin, 1972) 30–36

[40] B. McKay: Dual curves and pseudoholomorphic curves *Selecta Math. (N.S.)* **9(2)** (2003) 251–311

[41] R. McLean: Deformations of calibrated submanifolds *Comm. Anal. Geom.* **6(4)** (1998) 705–747

[42] M. Rabelo and K. Tenenblat: Submanifolds of constant nonpositive curvature *Mat. Contemp.* **1** (1991) 71–81

[43] D. Spencer: Overdetermined systems of linear partial differential equations *Bull. Amer. Math. Soc.* **75** (1969) 179–239

[44] O. Stormark: *Lie's structural approach to PDE systems* (Encyclopedia of Mathematics and its Applications 80, Cambridge University Press, Cambridge, 2000)

[45] A. Vinogradov: The $\mathcal{C}$-spectral sequence, Lagrangian formalism, and conservation laws. I. The linear theory *J. Math. Anal. Appl.* **100(1)** (1984) 1–40; A. Vinogradov: The $\mathcal{C}$-spectral sequence, Lagrangian formalism, and conservation laws. II. The nonlinear theory *J. Math. Anal. Appl.* **100(1)** (1984) 41–129

[46] S. Wang: Conservation laws for a class of third order evolutionary differential systems *Trans. Amer. Math. Soc.* **356(10)** (2004) 4055–4073

[47] F. W. Warner: *Foundations of differentiable manifolds and Lie groups* (Springer-Verlag, New York, 1983)

[48] D. Yang: Involutive hyperbolic differential systems *Mem. Amer. Math. Soc.* **370** (1987) 1–93

NIKY KAMRAN

Department of Mathematics and Statistics, Burnside Hall, McGill University, 805 Sherbrooke Street West, Montreal, Quebec H3A 2K6, Canada

E-mail: nkamran@math.mcgill.ca

This page intentionally left blank

Demeter Krupka and David Saunders

Weil bundles as generalized jet spaces

Ivan Kolář

Contents

Preface

Having in mind the methods of algebraic geometry, A. Weil introduced an infinitely near point on a smooth manifold M as an algebra homomorphism of the algebra $C^\infty(M, \mathbb{R})$ of smooth functions on M into a local algebra A, [23]. Nowadays A is called Weil algebra and the space $T^A M$ of the corresponding near points on M is said to be a Weil bundle. About 1985, it was deduced that the product preserving bundle functors on the category of all smooth manifolds $\mathcal{M}f$ are just the Weil functors, see [14] or Section 2.6 below. This result clarified that Weil bundles should be a good instrument for differential geometry. In this connection, the so-called covariant approach to Weil bundles was developed, [14]. Under this approach, $T^A M$ is interpreted as a generalization of the bundle $T^r_k M$ of (k, r)-velocities introduced by C. Ehresmann in the framework of his jet theory. So the iteration $T^s_l T^r_k$ of two classical velocities functors is the simplest new example of a Weil functor. In 1999, the fiber product preserving bundle functors on the category $\mathcal{FM}_m$ of fibered manifolds with m-dimensional bases and fibered manifold morphisms with local diffeomorphisms as base maps were also described in terms of Weil algebras, [15]. This clarified that even these functors can be viewed as a reasonable generalization of various types of jet bundles.

In this paper we treat systematically the covariant approach and we present the most interesting geometric results deduced in this way. We aim to differential geometry and its applications in analysis and mathematical physics. So we start with a rather detailed presentation of the algebraic properties of Weil algebras in Section 1. In particular, every Weil algebra of width k and order r is expressed as a factor algebra of the algebra $\mathbb{D}^r_k$ of r-jets of smooth functions on $\mathbb{R}^k$ at 0. Our definition of Weil bundle $T^A M$ in Section 2 is of "practical" character: if A is a Weil (k, r)-algebra, then $T^A M$ can be viewed as a certain factor space of the velocities bundle $T^r_k M$. An important theoretical result reads that the natural transformations $T^{A_1} M \to T^{A_2} M$ of two Weil bundles are in bijection with algebra homomorphisms $\mu : A_1 \to A_2$. Under our approach, each μ can be interpreted as a kind of reparametrization. Even this is suitable for applications. In 2.11 we construct a canonical exchange isomorphism $\varkappa^A_M : T^A TM \to TT^A M$ that can be used for a simple construction of flow prolongations of vector fields. In 2.16 we clarify that the product preserving bundle functors on the category $\mathcal{FM}$ of all fibered manifolds are in bijection with the Weil algebra homomorphisms. The simplest example of these functors are the fiber velocities bundles described in 2.17.

Section 3 is devoted to concrete geometric problems related with T^A-prolongations. First we deduce that each $a \in A$ determines a tensor field of type $(1, 1)$ on $T^A M$. Then we discuss T^A-prolongations of Lie groups, their actions, principal and associated bundles. Next we show that T^A-prolongation preserves the Frölicher-Nijenhuis bracket of tangent valued forms. That is why we use the tangent valued forms as one of the approaches to connections in 3.11. In particular, this yields a general formula for the curvature of the T^A-induced connection and implies some interesting properties of an original concept of a-torsion, $a \in A$. We also discuss a few examples demostrating that the use of the algebra multiplication is very convenient for concrete evaluations.

In Section 4 we study a fiber product preserving bundle functor F on the category $\mathcal{FM}_m$. The basic examples are the r-th jet prolongation, the r-th vertical jet prolongation, the vertical Weil functor and their iterations. The iteration of jet functors leads to nonholonomic jets, whose composition is described in 4.2. The construction of product fibered manifolds is a functor $i : \mathcal{M}f_m \times \mathcal{M}f \to \mathcal{FM}_m$, where $\mathcal{M}f_m$ is the category of m-dimensional manifolds and local diffeomorphisms. If F is of base order r, then $F \circ i$ is characterized by a Weil algebra A and a group homomorphism $H : G^r_m \to \operatorname{Aut} A$ of the r-th jet group in dimension m into the group of all algebra automorphisms of A. Theorem 4.7 reads that the original functor F is identified with a triple (A, H, t), where $t : \mathbb{D}^r_m \to A$ is an equivariant algebra homomorphism. For every fibered manifold $Y \to M$, $m = \dim M$, FY is expressed in the form of a fiber bundle associated to the r-th order frame bundle of M. Then we describe the natural transformations and iterations of such functors. In 4.11 we introduce the general concept of r-th order jet functor as a suitable subfunctor of the nonholonomic one and we characterize it in terms of Weil algebras.

Section 5 is devoted to some applications of Theorem 4.7. In 5.2 we introduce an analogy ψ^F_Y of the flow natural exchange $\varkappa^A_M$ from the manifold case, but the construction of ψ^F_Y is much more sophisticated. Then we show how ψ^F_Y can be applied to F-prolongations of projectable tangent valued forms, the connections being a special case. In 5.7 we generalize the classical theory of jet prolongations of associated fiber bundles to the case of arbitrary F. We find remarkable that F-prolongation of Lie groupoids, studied in 5.8, is essentially based on the expression of FY in the form of an associated bundle from 4.7. Finally, we outline in 5.10 how our velocity-like approach to Weil bundles can be modified to the functional bundle of all smooth maps between the individual fibers of two fibered manifolds over the same base.

Except Section 5.10, we consider the classical smooth manifolds and maps, i.e. all manifolds are finite dimensional and smooth means C^∞-differentiable. All manifolds are assumed to be Hausdorff and separable, [14]. A bundle functor F on the category $\mathcal{M}f$

of all smooth manifolds and all smooth maps transforms every manifold M into a fibered manifold $\pi_M : FM \to M$ and every map $f : M_1 \to M_2$ into a fibered manifold morphism $Ff : FM_1 \to FM_2$ over f. Further, F is assumed to have the localization property: if $i_U : U \hookrightarrow M$ is the injection of an open subset, then $FU = \pi_M^{-1}(U) \subset FM$ and Fi_U is the injection of $\pi_M^{-1}(U)$, [14]. We shall consider the bundle functors also on the following categories:

$\mathcal{M}f_m \subset \mathcal{M}f$ - m-dimensional manifolds and local diffeomorphisms,

$\mathcal{FM}$ - all fibered manifolds and their morphisms,

$\mathcal{FM}_m \subset \mathcal{FM}$ - fibered manifolds with m-dimensional bases and local diffeomorphisms as base maps.

The author acknowledges M. Doupovec, M. Kureš and P. Vašík for reading the manuscript and for several useful comments.

The preparation of this part of the book was supported by the Ministry of Education of Czech Republic under the Research Project MSM 0021622409 and by the Grant Agency of Czech Republic under the Grant 201/05/0523.

1 Weil algebras

1.1 Algebras

An algebra is a vector space V together with a bilinear map $f : V \times V \to V$, which is called algebra multiplication. We write $f(x, y) = xy$. The bilinearity of f implies

$$ox = o\,, \quad xo = o\,, \qquad x \in V,\ o = \text{ the zero vector of } V.$$

Let $(\overline{V}, \overline{f})$ be another algebra. An algebra homomorphism $\mu : (V, f) \to (\overline{V}, \overline{f})$ is a linear map $\mu : V \to \overline{V}$ preserving the multiplications. In what follows all algebras are assumed to be both commutative and associative.

An ideal $I \subset V$ is a linear subspace such that

$$xa \in I \quad \text{for all} \quad x \in I,\ a \in V\,.$$

For every subset $S \subset V$, we denote by $\langle S\rangle$ the ideal generated by S, i.e. the smallest ideal in V containing S. The factor vector space V/I is an algebra with respect to the multiplication

$$(a + I)\,(b + I) = ab + I\,.$$

On the other hand, the kernel of every algebra homomorphism is an ideal.

An element $a \in V$ is said to be nilpotent, if $a^n = o$ for some integer n. It is easy to see that the set N of all nilpotent elements of V is an ideal.

A unit of V is an element $e \neq o$ satisfying $ex = x$ for all $x \in V$. If the unit exists, it is unique. An algebra with unit is said to be unital. The unit defines an injection $\mathbb{R} \hookrightarrow V$, $c \mapsto ce$. In this case, we write $\mathbb{R} \subset V$ and we identify e with $1 \in \mathbb{R}$ and the zero vector o with $0 \in \mathbb{R}$. The homomorphisms of unital algebras $V \to \overline{V}$ are assumed to transform the unit of V into the unit of $\overline{V}$.

1.2 Weil algebras

The following concept was called local algebra in the original paper by A. Weil, [23].

Definition A Weil algebra A is a finite dimensional, commutative, associative and unital algebra of the form

$$A = \mathbb{R} \times N\,,$$

where N is the ideal of all nilpotent elements of A.

In particular, $\mathbb{R}$ is a trivial Weil algebra with $N = 0$. Let $\overline{A} = \mathbb{R} \times \overline{N}$ be another Weil algebra and $\mu : A \to \overline{A}$ be a homomorphism. Then the restriction and corestriction of μ to $\mathbb{R} \subset A$ and $\mathbb{R} \subset \overline{A}$ is the identity and μ transforms N into $\overline{N}$. The zero homomorphism $\mathcal{O} : A \to \overline{A}$ maps N into $0 \in \overline{A}$. We write $\mathrm{Hom}(A, \overline{A})$ for the set of all algebra homomorphisms of A into $\overline{A}$. An ideal $I \subset A$ is assumed to be different from A, so that $I \subset N$.

1.3 The width and the order

In general, for a linear subspace $W \subset V$ we define

$$W^n = \{a_1 \dots a_n + \dots + b_1 \dots b_n;\ a_1, \dots, a_n, \dots, b_1, \dots, b_n \in W\}\,,$$

i.e. the elements of W^n are finite sums of the products of n elements of W.

Since $A = \mathbb{R} \times N$ is finite dimensional, there exists an integer r such that $N^{r+1} = 0$. The smallest r with this property is called the order $\mathrm{ord}\, A$ of A. (A. Weil used the term “depth”, [23].) On the other hand, the dimension wA of the vector space N/N^2 is said to be the width of A. A Weil algebra of width k and order r will be called Weil (k, r)-algebra.

Every algebra homomorphism $\mu : A \to \overline{A}$ induces a linear map

$$\widetilde{\mu} : N/N^2 \to \overline{N}/\overline{N}^2\,, \quad \widetilde{\mu}(a + N^2) = \mu(a) + \overline{N}^2. \tag{1}$$

Clearly, if B is another Weil algebra and $\nu : \overline{A} \to B$ is an algebra homomorphism, then $\widetilde{\nu \circ \mu} = \widetilde{\nu} \circ \widetilde{\mu}$.

1.4 The algebra $\mathbb{D}^r_k$

Let $\mathbb{R}[x_1, \dots, x_k]$ be the algebra of all polynomials in k undetermined. The simplest example of a Weil algebra is

$$\mathbb{D}^r_k = \mathbb{R}[x_1, \dots, x_k]/\langle x_1, \dots, x_k\rangle^{r+1}\,. \tag{2}$$

As a vector space, $\mathbb{D}^r_k$ is the set of all polynomials of degree at most r in k undetermined with the standard addition and multiplication by real scalars. The product of P, $Q \in \mathbb{D}^r_k$ is the “truncated” one: we multiply P and Q as polynomials and we neglect the terms of degree higher than r. Clearly, $\mathrm{ord}(\mathbb{D}^r_k) = r$ and $w(\mathbb{D}^r_k) = k$.

We shall write N^r_k for the nilpotent part of $\mathbb{D}^r_k$. The elements of N^r_k are the polynomials without absolute term.

In particular, $\mathbb{D}^1_1$ is the classical algebra $\mathbb{D}$ of dual (or Study) numbers. Its elements can be written as $a + be$, $a, b \in \mathbb{R}$, with e satisfying $e^2 = 0$.

1.5 A as a factor algebra

The following assertion gives a rather concrete description of Weil algebras.

Proposition *Every Weil algebra A of order $\le r$ and of width k is a factor algebra of $\mathbb{D}^r_k$.*

Proof Choose $a_1, \dots, a_k \in N$ such that

$$a_1 + N^2, \dots, a_k + N^2 \tag{3}$$

is a basis of the vector space N/N^2. Define $\pi : \mathbb{D}^r_k \to A$, $\pi(P) = P(a_1, \dots, a_k)$. First we deduce that π is a homomorphism, i.e. $(PQ)(a_1, \dots, a_k)$
$= P(a_1, \dots, a_k)\, Q(a_1, \dots, a_k)$. Indeed, PQ on the left hand side is the product in $\mathbb{D}^r_k$, while $P(a_1, \dots, a_k)\, Q(a_1, \dots, a_k)$ on the right hand side is the standard product of polynomials in $a_1, \dots, a_k$. But the condition $N^{r+1} = 0$ suppresses the terms of degree $> r$.

It remains to show that π is surjective. Since (3) is a basis, for every $a \in N$ we have

$$a + N^2 = c_1(a_1 + N^2) + \cdots + c_k(a_k + N^2)\,, \quad c_1, \dots, c_k \in \mathbb{R}\,,$$

i.e. $a - c_1 a_1 - \cdots - c_k a_k = n_1 \in N^2$. One verifies directly that the elements $a_i a_j + N^3$ generate linearly N^2/N^3, so that there exist $c_{ij} \in \mathbb{R}$ such that

$$n_1 = \sum_{i,j=1}^{k} c_{ij} a_i a_j + n_2\,, \qquad n_2 \in N^3\,.$$

In the $(l-1)$-st step of such procedure, we obtain

$$n_{l-1} = \sum_{i_1,\dots,i_l=1}^{k} c_{i_1 \dots i_l} a_{i_1} \dots a_{i_l} + n_l\,, \qquad n_l \in N^{l+1}\,.$$

But $N^{r+1} = 0$, so that after r steps we have

$$a = \sum_{i=1}^{k} c_i a_i + \cdots + \sum_{i_1,\dots,i_r=1}^{k} c_{i_1 \dots i_r} a_{i_1} \dots a_{i_r}\,.$$

The c's determine a polynomial $P \in N^r_k$ satisfying $a = P(a_1, \dots, a_k)$. □

This proof yields directly the following assertion. Every Weil (k, r)-algebra A is a factor algebra

$$A = \mathbb{R}[x_1, \dots, x_k]/I\,, \tag{4}$$

where I is an ideal satisfying $\langle x_1, \dots, x_k \rangle^2 \supset I \supset \langle x_1, \dots, x_k \rangle^{r+1}$ with minimal r. We write $\overline{\pi} : \mathbb{R}[x_1, \dots, x_k] \to A$ for the factor projection.

1.6 Reparametrizations

We shall use heavily the following interpretation of $\mathbb{D}^r_k$ in terms of jets. By (2), the elements of $\mathbb{D}^r_k$ are r-jets of functions on $\mathbb{R}^k$ at 0, i.e. $\mathbb{D}^r_k = J^r_0(\mathbb{R}^k, \mathbb{R})$. The addition in $\mathbb{D}^r_k$, the multiplication by reals and the multiplication in $\mathbb{D}^r_k$ are expressed by the formulae

$$j^r_0\gamma + j^r_0\delta = j^r_0(\gamma + \delta)\,, \quad c j^r_0\gamma = j^r_0(c\gamma)\,, \quad (j^r_0\gamma)(j^r_0\delta) = j^r_0(\gamma\delta)\,,$$

$\gamma, \delta : \mathbb{R}^k \to \mathbb{R}$, $c \in \mathbb{R}$.

We denote the composition of jets by the same symbol $\circ$ as the composition of maps. Every r-jet $X \in J_0^r(\mathbb{R}^l, \mathbb{R}^k)_0$ induces an algebra homomorphism $\mathbb{D}_k^r \to \mathbb{D}_l^r$ by

$$Y \to Y \circ X\,, \quad Y \in \mathbb{D}_k^r = J_0^r(\mathbb{R}^k, \mathbb{R})\,. \tag{5}$$

Geometrically, (5) is a reparametrization of the elements of $\mathbb{D}_k^r$.

Proposition *We have* $\mathrm{Hom}(\mathbb{D}_k^r, \mathbb{D}_l^r) = J_0^r(\mathbb{R}^l, \mathbb{R}^k)_0$.

Proof Let $\mu : \mathbb{D}_k^r \to \mathbb{D}_l^r$ be an algebra homomorphism. Then $\mu(x_i) = P_i \in N_l^r$ for every $i = 1, \dots, k$. Hence $P = (P_1, \dots, P_k)$ is a k-tuple of polynomials of degree at most r without the absolute term in l undetermined. This defines an r-jet $P \in J_0^r(\mathbb{R}^l, \mathbb{R}^k)_0$. By the definition of jet composition, $\mu(Y) = Y \circ P$ for all $Y \in \mathbb{D}_k^r$. □

1.7 Supplement to Proposition 1.5

Now we can prove

Proposition *Let A be a Weil (k, r)-algebra and $\pi, \varrho : \mathbb{D}_k^r \to A$ be two surjective algebra homomorphisms. Then there is an algebra isomorphism $\sigma : \mathbb{D}_k^r \to \mathbb{D}_k^r$ satisfying $\pi = \varrho \circ \sigma$.*

Proof Write $a_i = \pi(x_i)$ and choose some $P_i \in \mathbb{D}_k^r$ satisfying $\varrho(P_i) = a_i$. Consider the homomorphism $\sigma : \mathbb{D}_k^r \to \mathbb{D}_k^r$ transforming x_i into P_i. Then $\varrho\big(\sigma(x_i)\big) = \pi(x_i)$, so that $\pi = \varrho \circ \sigma$. Consider the induced maps $\widetilde{\pi}$, $\widetilde{\varrho} : N_k^r/(N_k^r)^2 \to N/N^2$, $\widetilde{\sigma} : N_k^r/(N_k^r)^2 \to N_k^r/(N_k^r)^2$ from 1.3. Both $\widetilde{\pi}$ and $\widetilde{\varrho}$ are linear isomorphisms by the dimension argument, so $\widetilde{\sigma}$ is too. But σ is determined by a reparametrization $Y \mapsto Y \circ X$. One verifies directly that the invertibility of $\widetilde{\sigma}$ is equivalent to the fact that the linear map corresponding to the underlying 1-jet of X is invertible. Hence X is an invertible r-jet, so that σ is an isomorphism. □

1.8 Algebra homomorphisms

Let A and B be two Weil algebras with $k = wA$, $l = wB$, $r = \max(\operatorname{ord} A, \operatorname{ord} B)$ and $\mu : A \to B$ be an algebra homomorphism. Consider two surjective homomorphisms $\pi : \mathbb{D}_k^r \to A$ and $\varrho : \mathbb{D}_l^r \to B$. Analogously to 1.7, one deduces there is an algebra homomorphism $\sigma : \mathbb{D}_k^r \to \mathbb{D}_l^r$ such that the following diagram commutes

$$\begin{array}{ccc} \mathbb{D}_k^r & \xrightarrow{\ \sigma\ } & \mathbb{D}_l^r \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\varrho} \\ A & \xrightarrow{\ \mu\ } & B \end{array} \tag{6}$$

An element $X \in J_0^r(\mathbb{R}^l, \mathbb{R}^k)_0$ such that (6) with $\sigma = X$ commutes will be called μ-admissible.

Let A be expressed by (4) and $P_h(x_1, \dots, x_k)$ be some generators of ideal I, $h = 1, \dots, s$. To determine all algebra homomorphisms $A \to B$, we first consider an arbitrary k-tuple b_i of elements in the nilpotent part of B to be the images of $\overline{\pi}(x_i)$. Then the rule $b_i = \mu\big(\overline{\pi}(x_i)\big)$ generates an algebra homomorphism $\mu : A \to B$, if and only if $P_h(b_1, \dots, b_k) = 0$ for all h. An example can be found in 1.12 below.

1.9 Automorphisms

The group Aut A of all algebra automorphisms of A is a closed subgroup in the group $GL(A)$ of all linear automorphisms of A, so a Lie group. By 1.6, the group Aut $(\mathbb{D}^r_k)$ coincides with the jet group $G^r_k = \text{inv}\, J^r_0(\mathbb{R}^k, \mathbb{R}^k)_0$ of all invertible r-jets of $\mathbb{R}^k$ into $\mathbb{R}^k$ with source and target 0.

A derivation of A is a linear map $f : A \to A$ satisfying $f(a_1a_2) = f(a_1)a_2 + a_1f(a_2)$ for all $a_1, a_2 \in A$. According to the general theory, the Lie algebra $\mathfrak{Aut}A$ of Aut A coincides with the Lie algebra Der A of all derivations of A.

1.10 Sums

For two Weil algebras $A = \mathbb{R} \times N_A$, $B = \mathbb{R} \times N_B$, the vector space

$$A \oplus B = \mathbb{R} \times N_A \times N_B$$

is also a Weil algebra with respect to the multiplication

$$(c_1, a_1, b_1)(c_2, a_2, b_2) = (c_1c_2, c_1a_2 + c_2a_1 + a_1a_2, c_1b_2 + c_2b_1 + b_1b_2)\,,$$

$c_i \in \mathbb{R}$, $a_i \in N_A$, $b_i \in N_B$, $i = 1, 2$. One can say that $A \oplus B$ is the sum of A and B.

If A is expressed by (4) and B analogously by

$$\mathbb{R}[y_1, \dots, y_l]/J\,, \tag{7}$$

then $\langle I, J, x_iy_p\rangle$, $i = 1, \dots, k$, $p = 1, \dots, l$ is an ideal in $\mathbb{R}[x_1, \dots, x_k, y_1, \dots, y_l]$ and we have

$$A \oplus B = \mathbb{R}[x_1, \dots, y_l]/\langle I, J, x_iy_p\rangle\,.$$

In particular, this implies

$$w(A \oplus B) = wA + wB\,, \quad \text{ord}(A \oplus B) = \max(\text{ord}\, A, \text{ord}\, B)\,.$$

1.11 Tensor products

In general, the tensor product $V_1 \otimes V_2$ of two algebras (V_1, f_1) and (V_2, f_2) is also an algebra, whose multiplication f is the tensor product of f_1 and f_2. Thus, for the decomposable tensors $v_1 \otimes v_2$, $\overline{v}_1 \otimes \overline{v}_2$, we have

$$f(v_1 \otimes v_2, \overline{v}_1 \otimes \overline{v}_2) = f_1(v_1, \overline{v}_1) \otimes f_2(v_2, \overline{v}_2)\,.$$

In the case of two Weil algebras $A = \mathbb{R} \times N_A$, $B = \mathbb{R} \times N_B$, $A \otimes B$ is also a Weil algebra with the nilpotent part $N_A \times N_B \times N_A \otimes N_B$. If A is expressed by (4) and B by (7), then $\langle I, J\rangle$ is an ideal in $\mathbb{R}[x_1, \dots, x_k, y_1, \dots, y_l]$ and we have

$$A \otimes B = \mathbb{R}[x_1, \dots, y_l]/\langle I, J\rangle\,.$$

In particular, this implies,

$$w(A \otimes B) = wA + wB\,, \quad \text{ord}(A \otimes B) = \text{ord}\, A + \text{ord}\, B\,.$$

For instance, $\mathbb{D} \otimes \mathbb{D}$ is of the form

$$\mathbb{R}[x, y]/\langle x^2, y^2\rangle\,.$$

1.12 Example

We determine all algebra homomorphisms $\mu : \mathbb{D}\otimes\mathbb{D} \to \mathbb{D}\otimes\mathbb{D}$. Write

$$\mu(x) = c_1x + c_2y + c_3xy\,, \quad \mu(y) = c_4x + c_5y + c_6xy\,.$$

The conditions $\big(\mu(x)\big)^2 = 0$ and $\big(\mu(y)\big)^2 = 0$ imply $c_1c_2 = 0$ and $c_4c_5 = 0$. By 1.8, all algebra homomorphisms $\mathbb{D}\otimes\mathbb{D} \to \mathbb{D}\otimes\mathbb{D}$ form the following 4 four-parameter families

$$c_1 = 0 = c_4\,, \quad c_2 = 0 = c_4\,, \quad c_1 = 0 = c_5\,, \quad c_2 = 0 = c_5$$

with arbitrary c_3 and c_6.

2 Weil bundles

2.1 A-velocities

Having in mind the applications in concrete differential geometric problems, we introduce Weil bundles by using the concept of A-velocity. This generalizes the classical concept of (k,r)-velocity by C. Ehresmann. We recall that the construction of (k,r)-velocities is a bundle functor T^r_k on $\mathcal{M}f$ defined by

$$T^r_kM = J^r_0(\mathbb{R}^k, M)\,,\; T^r_kf(j^r_0\gamma) = j^r_0(f\circ\gamma)\,, \quad \gamma : \mathbb{R}^k \to M$$

for every manifold M and every smooth map $f : M \to N$.

Consider a Weil (k,r)-algebra A together with an algebra homomorphism $\pi : \mathbb{D}^r_k \to A$ from 1.5. By 1.7, π is determined up to an isomorphism $\mathbb{D}^r_k \to \mathbb{D}^r_k$, so that the following definition is independent of π. Let M be a manifold.

Definition Two maps $\gamma, \delta : \mathbb{R}^k \to M$ are said to determine the same A-velocity $j^A\gamma = j^A\delta$, if for every smooth function $\varphi : M \to \mathbb{R}$

$$\pi(j^r_0(\varphi\circ\gamma)) = \pi(j^r_0(\varphi\circ\delta))\,.$$

Proposition *Let $\gamma^i(t_1,\dots,t_k)$ or $\delta^i(t_1,\dots,t_k)$ be the coordinate expressions of γ or δ. Then $j^A\gamma = j^A\delta$ if and only if $j^A\gamma^i = j^A\delta^i$ for all $i = 1,\dots,\dim M$.*

Proof The kernel of π is an ideal I in $\mathbb{D}^r_k$. So $j^A\gamma^i = j^A\delta^i$ means

$$j^r_0\delta^i = j^r_0\gamma^i + j^r_0\varepsilon^i\,, \qquad j^r_0\varepsilon^i \in I\,.$$

Since our assertion depends on $j^r_x\varphi$ only, we may assume φ is a polynomial of degree r. This is a sum of monomials $cx^{i_1}\dots x^{i_l}$. The value of each monomial at $j^r_0\gamma^i + j^r_0\varepsilon^i$ is

$$c(j^r_0\gamma^{i_1} + j^r_0\varepsilon^{i_1})\dots(j^r_0\gamma^{i_l} + j^r_0\varepsilon^{i_l}) = cj^r_0\gamma^{i_1}\dots j^r_0\gamma^{i_l} + \text{ a term in } I.$$

□

2.2 Weil functors

Proposition 2.1 implies

$$T^A\mathbb{R} = \pi(\mathbb{D}_k^r) = A \quad \text{and} \quad T^A\mathbb{R}^m = A^m\,.$$

So $T^A M \to M$ is a fibered manifold that is said to be a Weil bundle. For a smooth map $f : M \to N$, we define

$$T^A f : T^A M \to T^A N \quad \text{by} \quad T^A f(j^A\gamma) = j^A(f \circ \gamma)\,.$$

Using the same argument as in the proof of Proposition 2.1, one verifies that this is a correct definition. Hence we obtain a bundle functor T^A on $\mathcal{M}f$ called Weil functor. Clearly, $T^{\mathbb{D}_k^r} = T_k^r$, so that $T^{\mathbb{D}}$ is the tangent functor T. Our construction yields a surjective map $\pi_M : T_k^r M \to T^A M$ such that the following diagram commutes for every $f : M \to N$

$$\begin{array}{ccc} T_k^r M & \xrightarrow{T_k^r f} & T_k^r N \\ {\scriptstyle \pi_M}\downarrow & & \downarrow{\scriptstyle \pi_N} \\ T^A M & \xrightarrow{T^A f} & T^A N \end{array} \tag{8}$$

A section $M \to T^A M$ is said to be an A-field on M.

2.3 Remarks

Definition 2.1 is a modification of what is called the covariant approach to Weil functors in [14]. Let $\mathcal{E}_k$ be the algebra of germs of smooth functions on $\mathbb{R}^k$ at 0. By (4), A can be viewed as a factor algebra $A = \mathcal{E}_k/\mathcal{I}$, where $\mathcal{I}$ is the ideal determined by I in $\mathcal{E}_k$. Then two maps γ, $\delta : \mathbb{R}^k \to M$ satisfy $j^A\gamma = j^A\delta$, if and only if $\varphi \circ \gamma - \varphi \circ \delta \in \mathcal{I}$ for every germ φ of smooth function on M at $x = \gamma(0) = \delta(0)$.

The original ideas by A. Weil, [23], were inspired by the algebraic geometry. So his approach is of contravariant character. All smooth functions on M form an algebra $C^\infty(M, \mathbb{R})$ with respect to the pointwise multiplication. Weil defined a so-called infinitely near point of type A on M as an algebra homomorphism

$$C^\infty(M, \mathbb{R}) \to A\,.$$

We show that the set of all algebra homomorphisms $\mathrm{Hom}(C^\infty(M, \mathbb{R}), A)$ is canonically isomorphic to $T^A M$.

For every $f \in C^\infty(M, \mathbb{R})$ and every $j^A\gamma \in T^A M$, we define

$$(j^A\gamma)(f) = j^A(f \circ \gamma) \in A\,. \tag{9}$$

This is an algebra homomorphism. Indeed, $(j^A\gamma)(f_1 + f_2) = j^A(f_1 \circ \gamma + f_2 \circ \gamma)$ and $(j^A\gamma)(f_1 f_2) = j^A((f_1 \circ \gamma)(f_2 \circ \gamma))$. Since our operations behave well with respect to localization, it suffices to consider the case $M = \mathbb{R}^m$, so that $T^A\mathbb{R}^m = A^m$. Then Proposition 2.1 implies that (9) establishes a bijection between $T^A\mathbb{R}^m$ and $\mathrm{Hom}(C^\infty(\mathbb{R}^m, \mathbb{R}), A)$.

We also remark that some motivation for the covariant approach came from the so-called synthetic differential geometry, which was developed within the framework of the theory of categories, [8].

2.4 The coordinate expression of $T^A f$

Write $T_x^A f : T_x^A M \to T_{f(x)}^A N$ for the restricted and corestricted map. The map $T_x^A f$ depends on $j_x^r f$ only. Indeed, (8) is a commutative diagram with surjective columns, where the top row depends on $j_x^r f$ only.

Using (8), we deduce the coordinate expression of $T^A f$ in the case of $f : \mathbb{R}^m \to \mathbb{R}^n$ with the components

$$y^p = f^p(x^i)\,, \qquad i = 1, \ldots, m,\ p = 1, \ldots, n\,. \tag{10}$$

Hence $T^A f : A^m \to A^n$. First consider the map $T_k^r f : T_k^r \mathbb{R}^m \to T_k^r \mathbb{R}^n$. This is determined by the Taylor expansions of order r of the components f^p

$$f^p(x^i) + \sum_{|\alpha| \leq r} \frac{1}{\alpha!} D_\alpha f^p(x^i) z^\alpha\,, \quad (z^i) \in \mathbb{R}^m\,,$$

where $\alpha = (\alpha_1, \ldots, \alpha_m)$ denotes a multiindex of range m and $z^\alpha = (z^1)^{\alpha_1} \ldots (z^m)^{\alpha_m}$. Write $x^i + n^i$ for the i-th component in A^m, $n^i \in N_A$, and b^p for the p-th component in A^n. If we express $T_k^r f : (\mathbb{D}_k^r)^m \to (\mathbb{D}_k^r)^n$ in the algebra form and use the homomorphism $\pi : \mathbb{D}_k^r \to A$, we obtain the following expression of $T^A f$

$$b^p = f^p(x^i) + \sum_{|\alpha| \leq r} \frac{1}{\alpha!} D_\alpha f^p(x^i) n^\alpha\,, \tag{11}$$

where $n^\alpha = (n^1)^{\alpha_1} \ldots (n^m)^{\alpha_m}$ with the multiplication in A.

In particular, if $f : \mathbb{R}^m \to \mathbb{R}$ is a real valued function, then (11) with no superscript p expresses the A-valued function $T^A f : T^A \mathbb{R}^m \to A$. We shall need an explicit formula in the simplest case $A = \mathbb{D}$. Write x_1^i for the additional coordinates on $T\mathbb{R}^m$. Then the coordinate form of $Tf : T\mathbb{R}^m \to \mathbb{D}$ is

$$f(x^i) + e\Big(\frac{\partial f}{\partial x^i}\, x_1^i\Big)\,. \tag{12}$$

2.5 Example

Consider the iterated tangent functor TT. The elements of its Weil algebra $\mathbb{D} \otimes \mathbb{D} = \mathbb{R}[t, \tau]/\langle t^2, \tau^2 \rangle$ are of the form $x + ut + v\tau + wt\tau$. The corresponding coordinates on $TT\mathbb{R}^m = (\mathbb{D} \otimes \mathbb{D})^m$ are x^i, u^i, v^i, w^i. From the classical point of view, x^i are the "original" coordinates on $\mathbb{R}^m$, the role of the "first order" coordinates u^i and v^i is more or less symmetric and w^i appear clearly to be the "second order" coordinates.

Further, consider the map (10) and write $y^p + \overline{u}^p t + \overline{v}^p \tau + \overline{w}^p t\tau$ for p-th component of $(\mathbb{D} \otimes \mathbb{D})^n$. Since $t^2 = 0 = \tau^2$, (11) implies

$$y^p + \overline{u}^p t + \overline{v}^p \tau + \overline{w}^p t\tau = f^p + \frac{\partial f^p}{\partial x^i}(u^i t + v^i \tau + w^i t\tau) + \frac{\partial^2 f^p}{\partial x^i \partial x^j}\, u^i v^j t\tau\,.$$

Passing to the individual components, we obtain the standard coordinate expression of TTf.

2.6 Product preserving bundle functors

Consider the product $M \xleftarrow{p_1} M \times N \xrightarrow{p_2} N$ of two manifolds together with the product projections. A bundle functor F on $\mathcal{M}f$ is said to be product preserving, if $F(M \times N) = FM \times FN$. More precisely, this means that

$$FM \xleftarrow{Fp_1} F(M \times N) \xrightarrow{Fp_2} FN$$

is also a product. Clearly, every Weil functor T^A preserves products.

The converse assertion is a fundamental theoretical result. Let F be a product preserving bundle functor on $\mathcal{M}f$. Write pt for one point set and $i_x : pt \to M$ for the map $i_x(pt) = x$, $x \in M$. Since F preserves products, we have $F(pt) = pt$. A natural injection $\nu_M : M \to FM$ is defined by $\nu_M(x) = Fi_x(pt)$. Applying F to the addition $a : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and the multiplication $m : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ of reals, we obtain

$$Fa : F\mathbb{R} \times F\mathbb{R} \to F\mathbb{R}\,, \quad Fm : F\mathbb{R} \times F\mathbb{R} \to F\mathbb{R}\,.$$

One verifies easily that $F\mathbb{R}$ with the addition Fa and the multiplication by real scalars $ca = Fm\big(\nu_{\mathbb{R}}(c), a\big)$, $c \in \mathbb{R}$, $a \in F\mathbb{R}$, is a vector space. The proof of the following assertion can be found in [14].

Theorem *$F\mathbb{R}$ is a Weil algebra with respect to the multiplication Fm and F coincides with the Weil functor $T^{F\mathbb{R}}$.*

A simple example of a bundle functor on $\mathcal{M}f$ that does not preserve products is the second tensor power $\overset{2}{\bigotimes} T$ of the tangent functor T. Indeed, $\dim \overset{2}{\bigotimes} T(M \times N) > \dim \overset{2}{\bigotimes} TM + \dim \overset{2}{\bigotimes} TN$ provided $\dim M > 0$ and $\dim N > 0$.

2.7 Natural transformations

Let $t : T^A \to T^B$ be a natural transformation of functors T^A and T^B. If we apply t to the addition and the multiplication of reals, we obtain two commutative diagrams of the form

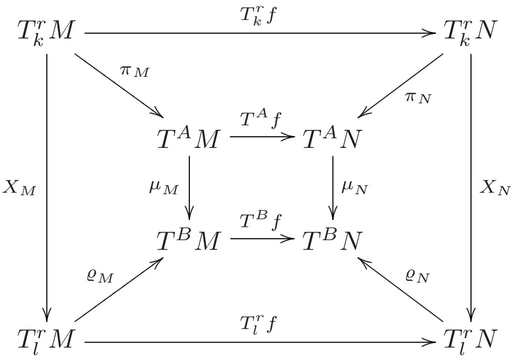
$$\begin{array}{ccc} A \times A & \longrightarrow & A \\ {\scriptstyle t_{\mathbb{R}}}\downarrow \quad \downarrow{\scriptstyle t_{\mathbb{R}}} & & \downarrow{\scriptstyle t_{\mathbb{R}}} \\ B \times B & \longrightarrow & B \end{array}$$

This implies easily that $t_{\mathbb{R}} : A \to B$ is an algebra homomorphism.

The converse assertion is also true. Every algebra homomorphism $\mu : A \to B$ induces a natural transformation (denoted by the same symbol) $\mu : T^A \to T^B$ as follows. Every $X \in J^r_0(\mathbb{R}^l, \mathbb{R}^k)_0 = \mathrm{Hom}(\mathbb{D}^r_k, \mathbb{D}^r_l)$ defines a natural transformation $X_M : T^r_k M \to T^r_l M$ by the reparametrization $Y \to Y \circ X$, $Y \in T^r_k M$. For a general μ, we may consider the situation from (6) with $\sigma = X$. Taking into account that π_M and ϱ_M are surjective, we deduce by (6) that there is a unique map $\mu_M : T^A M \to T^B M$ making the following diagram commutative

$$\begin{array}{ccc} T^r_k M & \xrightarrow{X_M} & T^r_l M \\ {\scriptstyle \pi_M}\downarrow & & \downarrow{\scriptstyle \varrho_M} \\ T^A M & \xrightarrow{\mu_M} & T^B M \end{array} \tag{13}$$

Moreover, for every $f : M \to N$ the following diagram commutes



The inner square yields that μ_M form a natural transformation $T^A \to T^B$. Diagram (13) implies that there exists a map $\overline{\mu} : \mathbb{R}^l \to \mathbb{R}^k$ such that $\mu_M : T^A M \to T^B M$ is of the form

$$\mu_M(j^A \gamma) = j^B(\gamma \circ \overline{\mu}), \quad \gamma : \mathbb{R}^k \to M . \tag{14}$$

Thus we have proved

Proposition *The natural transformations $T^A \to T^B$ are in bijection with the algebra homomorphisms $\mu : A \to B$. If $wA = k$ and $wB = l$, then there exists a map $\overline{\mu} : \mathbb{R}^l \to \mathbb{R}^k$ such that μ_M is of the form (14).*

In other words, even in the case of an arbitrary algebra homomorphism $\mu : A \to B$, the natural transformation μ_M is determined by a reparametrization. The admissibility of μ in the sense of 1.8 depends on $j^A \overline{\mu}$ only.

For example, the natural transformations $TT \to TT$, that correspond to the algebra homomorphisms $\mathbb{D} \otimes \mathbb{D} \to \mathbb{D} \otimes \mathbb{D}$ from 1.12, can be immediately expressed in this way. Each of them can be easily interpreted as a geometric construction on the iterated tangent bundle.

2.8 Multilinear maps

For every vector space V, the vector space structure on $T^A V$ is defined by

$$j^A \gamma + j^A \delta = j^A(\gamma + \delta),\ c j^A \gamma = j^A(c\gamma), \quad c \in \mathbb{R},$$

with pointwise addition and scalar multiplication on the right-hand sides. Consider the map $\otimes : V \times A \to T^A V$,

$$\otimes(v, j^A \varphi(t_1, \dots, t_k)) = j^A(\varphi(t_1, \dots, t_k) v), \quad v \in V,\ \varphi : \mathbb{R}^k \to \mathbb{R} . \tag{15}$$

In coordinates, we have $V = \mathbb{R}^m$, $T^A V = A^m$ and (15) is of the form

$$\big((v_1, \dots, v_n), a\big) \longmapsto (v_1 a, \dots, v_n a), \quad v_i \in \mathbb{R},\ a \in A,\ i = 1, \dots, n .$$

This implies $T^A V = V \otimes A$. In particular, if $A = \mathbb{D}$, then $V \otimes \mathbb{D} = V \times V$ and $TV = V \times V$ is the classical expression of the tangent bundle of a vector space.

Consider another vector space W and a linear map $f : V \to W$. Then

$$T^A f(v \otimes a) = T^A f(j^A(\varphi v)) = j^A(\varphi f(v)) = f(v) \otimes a\,.$$

One verifies directly that $T^A f : T^A V \to T^A W$ is also a linear map. Hence $T^A f = f \otimes \mathrm{id}_A$.

Proposition *Let $f : V_1 \times \cdots \times V_l \to W$ be a multilinear map. Then $T^A f : V_1 \otimes A \times \cdots \times V_l \otimes A \to W \otimes A$ is also multilinear and*

$$T^A f(v_1 \otimes a_1, \ldots, v_l \otimes a_l) = f(v_1, \ldots, v_l) \otimes a_1 \ldots a_l\,,$$

where the product $a_1 \ldots a_l$ is in A.

Proof The multilinearity of f implies

$$T^A f(j^A(\varphi_1 v_1), \ldots, j^A(\varphi_l v_l)) = j^A((\varphi_1 \ldots \varphi_l) f(v_1, \ldots, v_l))$$

with the pointwise product $\varphi_1(t_1, \ldots, t_k) \ldots \varphi_l(t_1, \ldots, t_k)$ in $\mathbb{R}$. □

Corollary *If (V, f) is an algebra, then $T^A V = V \otimes A$ is also an algebra with the multiplication determined by*

$$T^A f(v_1 \otimes a_1, v_2 \otimes a_2) = f(v_1, v_2) \otimes a_1 a_2\,.$$

2.9 Natural transformations on vector spaces

For every vector space V and every algebra homomorphism $\mu : A \to B$,

$$\mu_V : T^A V = V \otimes A \to V \otimes B = T^B V$$

is defined by the reparametrization (14), so that μ_V is a linear map. Applying μ_V to (15), we obtain $\mu_V(v \otimes a) = v \otimes \mu(a)$. Hence we have

$$\mu_V = \mathrm{id}_V \otimes \mu : V \otimes A \to V \otimes B\,.$$

2.10 The iteration

By 2.6, the Weil functors coincide with the product preserving bundle functors on $\mathcal{M}f$. Since the iteration $T^A T^B$ of two Weil functors preserves products as well, this must also be a Weil functor.

Proposition *We have $T^A T^B = T^{B \otimes A}$.*

Proof The Weil algebra of $T^A T^B$ is $(T^A T^B)(\mathbb{R}) = T^A(T^B \mathbb{R}) = T^A B = B \otimes A$. □

We know that every algebra homomorphism induces a natural transformation. In particular, the exchange isomorphism of the tensor product $\mathrm{ex} : B \otimes A \to A \otimes B$ induces an exchange natural equivalence

$$\mathrm{ex}_M : T^A T^B M \to T^B T^A M\,.$$

Geometrically, ex_M can be constructed as follows. Let $t \in \mathbb{R}^k$ and $\tau \in \mathbb{R}^l$. So every $Z \in T^A(T^B M)$ is of the form

$$Z = j^A\big(t \mapsto j^B(\tau \mapsto \delta(t, \tau))\big)\,,$$

where $\delta : \mathbb{R}^k \times \mathbb{R}^l \to M$. Then

$$\mathrm{ex}_M(Z) = j^B\big(\tau \mapsto j^A(t \mapsto \delta(t,\tau))\big)\,.$$

If we write $\pi^A_M : T^A M \to M$ for the bundle projection, then we have

$$T^A\pi^B_M = \pi^B_{T^A M} \circ \mathrm{ex}_M\,, \quad T^B\pi^A_M \circ \mathrm{ex}_M = \pi^A_{T^B M}\,. \tag{16}$$

In the special case $A = \mathbb{D}^r_k$, $B = \mathbb{D}^s_l$, which is important for applications, the natural exchange isomorphism $\mathrm{ex}_M : T^r_k T^s_l M \to T^s_l T^r_k M$ can be expressed in the jet form

$$\mathrm{ex}_M\big(j^r_0(t \mapsto j^s_0(\tau \mapsto \delta(t,\tau)))\big) = j^s_0\big(\tau \mapsto j^r_0(t \mapsto \delta(t,\tau))\big)\,.$$

In the case $k = l = r = s = 1$, we evaluated all algebra homomorphisms $\mathbb{D}\otimes\mathbb{D} \to \mathbb{D}\otimes\mathbb{D}$ in 1.12. Clearly, the exchange isomorphism ex, that determines the well known canonical involution $TTM \to TTM$, corresponds to the values $c_2 = c_4 = 1$, $c_1 = c_3 = c_5 = c_6 = 0$.

2.11 The flow natural exchange

In the case of arbitrary A and $T^B = T$, we obtain a canonical exchange isomorphism

$$\varkappa^A_M : T^A TM \to TT^A M\,.$$

This map is said to be flow natural because of the following important property.

In general, let F be a bundle functor on the category $\mathcal{M}f_m$ of m-dimensional manifolds and local diffeomorphisms. The flow prolongation of a vector field $X : M \to TM$ is a vector field $\mathcal{F}X : FM \to T(FM)$ defined as follows. The flow Fl^X is locally a one-parameter family of diffeomorphisms $Fl^X_t : M \to M$, $t \in \mathbb{R}$. We construct $F(Fl^X_t) : FM \to FM$ for every t and we set

$$\mathcal{F}X = \tfrac{\partial}{\partial t}\big|_0 F(Fl^X_t) : FM \to TFM\,.$$

In the case of a Weil functor T^A, beside $\mathcal{T}^A X : T^A M \to TT^A M$ we can consider the functorial prolongation $T^A X : T^A M \to T^A TM$ of the map X.

Proposition *We have $\mathcal{T}^A X = \varkappa^A_M \circ T^A X$.*

Proof Let $\overline{x} = \varphi(x,t)$ be the flow of X, $x \in M$, $t \in \mathbb{R}$. Hence $X(x) = \frac{\partial}{\partial t}\big|_0 \varphi(x,t)$. For $u = j^A\gamma(\tau) \in T^A M$, consider $\varphi(\gamma(\tau),t) : \mathbb{R}^k \times \mathbb{R} \to M$. Then

$$\tfrac{\partial}{\partial t}\big|_0 j^A\varphi(\gamma(\tau),t) = \tfrac{\partial}{\partial t}\big|_0 (T^A Fl^X_t)(u) = (\mathcal{T}^A X)(u)\,.$$

Of course, $\frac{\partial}{\partial t}\big|_0$ is identified with the first order jet with respect to t. If we exchange the order of j^A and $\frac{\partial}{\partial t}\big|_0$, we obtain

$$j^A\Big(\tfrac{\partial}{\partial t}\big|_0 \varphi(\gamma(\tau),t)\Big) = j^A(X\circ\gamma) = (T^A X)(u)\,.$$

□

So the concrete evaluation of $\mathcal{T}^A X$ for a vector field $X = X^i(x^1,\dots,x^m)\partial/\partial x^i$ is also based on 2.4. For example, if $T^A = T$ is the tangent functor, then (12) implies

$$\mathcal{T}X = X^i\frac{\partial}{\partial x^i} + \Big(\frac{\partial X^i}{\partial x^j}\,x^j_1\Big)\frac{\partial}{\partial x^i_1}\,.$$

2.12 Fiber products

If F and G are two product preserving bundle functors on $\mathcal{M}f$, then the bundle functor $F \oplus G$ on $\mathcal{M}f$ defined by $(F \oplus G)(M) = FM \times_M GM$ and $(F \oplus G)(f) = Ff \times_f Gf$ also preserves products.

Proposition *We have* $T^A \oplus T^B = T^{A \oplus B}$.

Proof $T^A\mathbb{R}$ or $T^B\mathbb{R}$ is the product fibered manifold $\mathbb{R} \times N_A \to \mathbb{R}$ or $\mathbb{R} \times N_B \to \mathbb{R}$, respectively. Hence $(T^A \oplus T^B)(\mathbb{R}) = \mathbb{R} \times N_A \times N_B \to \mathbb{R}$. One verifies easily that the induced multiplication is that one from 1.10. □

2.13 Underlying functors

An interesting feature of the theory of Weil bundles is that every r-th order Weil functor T^A induces the underlying k-th order functors for all $k \leq r = \operatorname{ord} A$.

Clearly, N_A^{k+1} is an ideal in A. Write $\pi_k : A \to A/N_A^{k+1}$ for the factor projection.

Definition The factor algebra $A_k = A/N_A^{k+1}$ is called the underlying Weil algebra of order k. The Weil functor T^{A_k} is said to be the underlying k-th order functor of T^A.

So $(\pi_k)_M : T^A M \to T^{A_k} M$ is a surjective natural transformation. The following lemma is a direct consequence of $\mu(N_A) \subset N_B$.

Lemma *For every algebra homomorphism* $\mu : A \to B$*, we have* $\mu(N_A^k) \subset N_B^k$.

So μ factorizes through an underlying algebra homomorphism $\mu_k : A_k \to B_k$. From the geometric point of view, every natural transformation $t : T^A \to T^B$ is projectable over a natural transformation $t_k : T^{A_k} \to T^{B_k}$ for all $k \leq r$.

For example, the underlying first order functor of the iterated tangent functor TT is $T \oplus T$.

In [9], the following result is deduced.

Proposition $T^A M \to T^{A_{r-1}} M$ *is an affine bundle, whose associated vector bundle is the pullback of* $TM \otimes N^r$ *over* $T^{A_{r-1}} M$.

In the special case of T_k^r, we obtain the classical result that $T_k^r M \to T_k^{r-1} M$ is an affine bundle, whose associated vector bundle is the pullback of $TM \otimes S^r\mathbb{R}^{k*}$ over $T_k^{r-1} M$.

2.14 Regular A-velocities

Consider the vector space $V_A = N_A/N_A^2$. One finds easily that the underlying Weil algebra of the first order is $A_1 = \mathbb{R} \times V_A$ with the zero multiplication in V_A. This implies $T^{A_1} M = TM \otimes V_A$.

Definition An A-velocity $X \in T_x^A M$ is called regular, if the linear map $V_A^* \to T_x M$ determined by $\pi_1(X) \in T_x M \otimes V_A$ is injective.

In the classical case of $X \in (T_k^r M)_x$, $\pi_1(X) \in T_k^1 M$ is identified with a k-tuple of vectors in $T_x M$ and X is regular, if and only if these vectors are linearly independent. In general, one verifies easily that an A-velocity $j^A\gamma$ is regular, if and only if γ is an immersion at $0 \in \mathbb{R}^k$.

2.15 Contact A-elements

We recall that a contact (k, r)-element on a manifold M, as introduced by Ehresmann, is defined as a set $Z \circ G_k^r$, where Z is a regular (k, r)-velocity on M, [14]. The space of all such elements is a fiber bundle $K_k^r M$ over M.

This idea can be directly extended to A-velocities.

Definition A contact A-element on a manifold M is the set $(\mathrm{Aut}\, A)(Z)$, where Z is a regular A-velocity on M.

All contact A-elements on M form a fiber bundle $K^A M \to M$. We remark that the contact A-elements are studied from an algebraic point view in [22].

2.16 Product preserving bundle functors on $\mathcal{FM}$

It is remarkable that the product preserving bundle functors on fibered manifolds can be also characterized in terms of Weil algebras. The following assertion was deduced by W. Mikulski, [20].

Proposition *The product preserving bundle functors on $\mathcal{FM}$ are in bijection with the algebra homomorphisms $\mu : A \to B$.*

On one hand, μ induces two bundle functors T^A, T^B on $\mathcal{M}f$ and a natural transformation $\mu : T^A \to T^B$. For every fibered manifold $p : Y \to M$, we have $T^B p : T^B Y \to T^B M$. Taking into account the map $\mu_M : T^A M \to T^B M$, we construct the induced bundle

$$T^\mu Y = \mu_M^* T^B Y \,,$$

which can be also denoted by $T^\mu Y = T^A M \times_{T^B M} T^B Y$. For every $\mathcal{FM}$-morphism $f : Y \to \overline{Y}$ over $\underline{f} : M \to \overline{M}$, we have $T^B f : T^B Y \to T^B \overline{Y}$, $T^A \underline{f} : T^A M \to T^A \overline{M}$ and we construct the induced map

$$T^\mu f := T^A \underline{f} \times_{T^B \underline{f}} T^B f : T^\mu Y \to T^\mu \overline{Y} \,.$$

This defines a bundle functor T^μ on $\mathcal{FM}$. Clearly, T^μ preserves products.

Conversely, let F be a product preserving bundle functor on $\mathcal{FM}$. Write $pt_M : M \to pt$ for the unique map of M into one point set. There are two canonical injections $i_1, i_2 : \mathcal{M}f \to \mathcal{FM}$ of manifolds into fibered manifolds defined by $i_1 M = (\mathrm{id}_M : M \to M)$, $i_1 f = (f, f)$, $i_2 M = (pt_M : M \to pt)$, $i_2 f = (f, \mathrm{id}_{pt})$ and a natural transformation $t : i_1 \to i_2$, $t_M = (\mathrm{id}_M, pt_M) : i_1 M \to i_2 M$. Applying F, we obtain two bundle functors $F \circ i_1$ and $F \circ i_2$ on $\mathcal{M}f$ and a natural transformation $F \circ t : F \circ i_1 \to F \circ i_2$. By 2.6 and 2.7, there is a Weil algebra homomorphism $\mu : A \to B$ such that $F \circ i_1 = T^A$, $F \circ i_2 = T^B$ and $F \circ t = \mu$. Then $F = T^\mu$. (A rather simple proof can be found in [5].)

Further, if $\nu : C \to D$ is another Weil algebra homomorphism, then the natural transformations $T^\mu \to T^\nu$ are in bijection with the pairs of Weil algebra homomorphisms $\varphi : A \to C$, $\psi : B \to D$ satisfying $\psi \circ \mu = \nu \circ \varphi$. Moreover, the iteration $T^\nu T^\mu$ corresponds to the tensor product $\mu \otimes \nu : A \otimes C \to B \otimes D$.

2.17 Fiber velocities

The simplest examples of product preserving bundle functors on $\mathcal{FM}$ are the fiber velocities bundles, which generalize the classical bundles of (k, r)-velocities. Their definition is based on the idea of (q,s,r)-jet of $\mathcal{FM}$-morphisms, $s \geq q \leq r$. Consider two fibered manifolds $p : Y \to M$ and $\overline{p} : \overline{Y} \to \overline{M}$.

Definition We say that two $\mathcal{FM}$-morphisms $f, g : Y \to \overline{Y}$ with the base maps $\underline{f}, \underline{g} : M \to \overline{M}$ determine the same (q,s,r)-jet $j_y^{q,s,r} f = j_y^{q,s,r} g$ at $y \in Y$, $s \geq q \leq r$, if

$$j_y^q f = j_y^q g\,, \quad j_y^s(f \mid Y_x) = j_y^s(g \mid Y_x)\,, \quad j_x^r \underline{f} = j_x^r \underline{g}\,, \qquad x = p(y)\,.$$

We write $J^{q,s,r}(Y, \overline{Y})$ for the bundle of all (q,s,r)-jets of Y into $\overline{Y}$. The composition of $\mathcal{FM}$-morphisms defines the composition of (q,s,r)-jets. The base maps induce a canonical projection $J^{q,s,r}(Y,\overline{Y}) \to J^r(M,\overline{M})$, where $J^r(M,\overline{M})$ denotes the classical bundle of r-jets of M into $\overline{M}$.

Write $\mathbb{R}^{k,l}$ for the product fibered manifold $\mathbb{R}^k \times \mathbb{R}^l \to \mathbb{R}^k$. We introduce the bundle of fiber velocities of dimension (k,l) and order (q,s,r) on a fibered manifold Y by

$$T_{k,l}^{q,s,r} Y = J_{0,0}^{q,s,r}(\mathbb{R}^{k,l}, Y)\,.$$

For every $\mathcal{FM}$-morphism $f : Y \to \overline{Y}$, $T_{k,l}^{q,s,r} f : T_{k,l}^{q,s,r} Y \to T_{k,l}^{q,s,r}\overline{Y}$ is defined by the jet composition. Clearly, $T_{k,l}^{q,s,r}$ is a product preserving bundle functor on $\mathcal{FM}$ of order (q,s,r).

Let $m = \dim M$ and $m+n = \dim Y$. Then $P^{q,s,r} Y = \operatorname{inv} J_{0,0}^{q,s,r}(\mathbb{R}^{m,n}, Y)$ is a principal bundle over Y with structure group $G_{m,n}^{q,s,r} = \operatorname{inv} J_{0,0}^{q,s,r}(\mathbb{R}^{m,n}, \mathbb{R}^{m,n})_{0,0}$, which is called (q,s,r)-th order frame bundle of Y.

3 On the geometry of T^A-prolongations

3.1 Natural tensor fields of type $(1,1)$

Every $a \in A$ defines a natural tensor field $L(a)_M$ of type $(1,1)$ on $T^A M$ for every manifold M as follows. The multiplication of the tangent vectors of M by reals is a map $\sigma_M : \mathbb{R} \times TM \to TM$. Applying T^A, we obtain $T^A\sigma_M : A \times T^A TM \to T^A TM$. Then we construct

$$\mathcal{T}^A \sigma_M := (\varkappa_M^A)^{-1} \circ T^A \sigma_M \circ (\mathrm{id}_A \times \varkappa_M^A) : A \times TT^A M \to TT^A M$$

and define $L(a)_M = \mathcal{T}^A\sigma_M(a, -)$. Since the multiplication in A is induced by the multiplication of reals, we have

$$L(a_1)_M \circ L(a_2)_M = L(a_1 a_2)_M\,.$$

Clearly, $L(1)_M = \mathrm{id}_{TT^A M}$. The naturality of $L(a)$ means $TT^A f \circ L(a)_M = L(a)_N \circ TT^A f$ for every map $f : M \to N$.

To find the coordinate expression of $L(a)$, we take $M = \mathbb{R}^m$. Then $TT^A\mathbb{R}^m = A^m \times A^m$ and our definition implies directly

$$L(a)_{\mathbb{R}^m}(b_1, \dots, b_m, c_1, \dots, c_m) = (b_1, \dots, b_m, ac_1, \dots, ac_m), \;\; a, b_i, c_i \in A.$$

In the case of the tangent functor, i.e. $A = \mathbb{D} = \{a + be\}$, the map $L(e)_M : TTM \to VTM$ is frequently used in analytical mechanics.

3.2 Natural vector fields

Every element D of the Lie algebra $\mathfrak{Aut}\, A = \operatorname{Der} A$ is of the form

$$D = \tfrac{d}{dt}\big|_0 \gamma\,, \quad \gamma : \mathbb{R} \to \operatorname{Aut} A\,.$$

The natural transformations $\gamma_M(t) : T^A M \to T^A M$ determine a vertical vector field D_M on $T^A M$,

$$D_M(y) = \tfrac{\partial}{\partial t}\big|_0 \gamma(t)_M(y)\,, \quad y \in T^A M\,.$$

For example, in the case $A = \mathbb{D}$ we obtain the classical Liouville vector field on TM and its constant multiples.

We find worth mentioning one of the oldest geometric results deduced by the technique of Weil algebras. The problem was to determine all natural operators transforming every vector field X on a manifold M into a vector field on $T^A M$. In [14] it is proved that every such operator is of the form

$$X \mapsto L(a)_M \circ \mathcal{T}^A X + D_M$$

for all $a \in A$ and $D \in \operatorname{Der} A$, $\mathcal{T}^A X$ being the flow prolongation of X. This general result was new even in the case of (k, r)-velocities.

3.3 T^A-prolongation of functions and vector fields

Every function $f : M \to \mathbb{R}$ induces a vector valued function $T^A f : T^A M \to A$. Every vector field Z on $T^A M$ determines the Lie derivative $Z(T^A f) : T^A M \to A$ of this vector valued function. Given $a \in A$, we define $aT^A f : T^A M \to A$ by multiplying in A. We shall shorten $L(a)_M \circ Z$ to $L(a)Z$.

We start with some simple formulae, [2]. For every vector field X on M, we have

$$\mathcal{T}^A X(aT^A f) = aT^A(Xf)\,, \quad \big(L(a)\mathcal{T}^A X\big)T^A f = aT^A(Xf)\,.$$

For the bracket of vector fields, we have

$$\big[L(a_1)\mathcal{T}^A X_1, L(a_2)\mathcal{T}^A X_2\big] = L(a_1 a_2)\mathcal{T}^A\big([X_1, X_2]\big)\,.$$

3.4 T^A-prolongation of Lie groups

The following assertions are easy to verify.

Let G be a Lie group with composition $\varphi : G \times G \to G$ and $\mathfrak{g} = \operatorname{Lie} G$ be its Lie algebra. Then $T^A\varphi : T^A G \times T^A G \to T^A G$ is also a Lie group and $T^A\mathfrak{g} = \mathfrak{g} \otimes A$ is its Lie algebra. The bracket in $T^A\mathfrak{g}$ is the T^A-prolongation of the bracket in $\mathfrak{g}$. Using 2.8, we obtain

$$[v_1 \otimes a_1, v_2 \otimes a_2]_{\mathfrak{g}\otimes A} = [v_1, v_2]_{\mathfrak{g}} \otimes a_1 a_2\,,$$

$v_1, v_2 \in \mathfrak{g}$, $a_1, a_2 \in A$, the product $a_1 a_2$ being in A. This formula is powerful even in the case of (k, r)-velocities, in which $\operatorname{Lie}(T^r_k G) = \mathfrak{g} \otimes \mathbb{D}^r_k$. For example, in the case of the tangent functor T we obtain immediately the well known expression for the bracket of $\mathrm{Lie}(TG) = \mathfrak{g} \times \mathfrak{g}$

$$\big[(v_1, \overline{v}_1), (v_2, \overline{v}_2)\big] = \big([v_1, v_2], [v_1, \overline{v}_2] + [\overline{v}_1, v_2]\big)\,.$$

If $\exp_G : \mathfrak{g} \to G$ is the exponential map of G, then the exponential map of $T^A G$ is

$$\exp_{T^A G} = T^A(\exp_G) : T^A\mathfrak{g} \to T^A G\,.$$

The Maurer-Cartan form $\omega_G : TG \to \mathfrak{g}$ of G induces the Maurer-Cartan form $\omega_{T^A G} : TT^A G \to T^A\mathfrak{g}$ by

$$\omega_{T^A G} = T^A \omega_G \circ (\varkappa^A_G)^{-1}\,.$$

For a group homomorphism $h : G_1 \to G_2$, $T^A h : T^A G_1 \to T^A G_2$ is also a group homomorphism. If $\chi : \mathfrak{g}_1 \to \mathfrak{g}_2$ is the induced Lie algebra homomorphism, then $T^A \chi = \chi \otimes \mathrm{id}_A : T^A\mathfrak{g}_1 \to T^A\mathfrak{g}_2$ is the Lie algebra homomorphism determined by $T^A h$.

For every Weil algebra homomorphism $\mu : A \to B$, the natural transformation $\mu_G : T^A G \to T^B G$ is a group homomorphism. The induced Lie algebra homomorphism is $\mathrm{id}_{\mathfrak{g}} \otimes \mu : \mathfrak{g} \otimes A \to \mathfrak{g} \otimes B$.

3.5 T^A-prolongation of actions

Let $l : G \times M \to M$ be a left action of G on a manifold M. One verifies directly that $T^A l : T^A G \times T^A M \to T^A M$ is a left action of $T^A G$ on $T^A M$. For every algebra homomorphism $\mu : A \to B$, the natural transformations $\mu_G : T^A G \to T^B G$ and $\mu_M : T^A M \to T^B M$ form a morphism of actions.

The infinitesimal action $\lambda : \mathfrak{g} \times M \to TM$ of l is defined by

$$\lambda = Tl \circ (i_G \times 0_M)\,,$$

where $i_G : \mathfrak{g} \to TG$ is the canonical injection and $0_M : M \to TM$ is the zero section. We write $\lambda(v) = \lambda(v, -) : M \to TM$ for the fundamental vector field on M determined by $v \in \mathfrak{g}$. One finds easily that the infinitesimal action of $T^A l$, which will be denoted by $\mathcal{T}^A \lambda : T^A\mathfrak{g} \times T^A M \to TT^A M$, is of the form

$$\mathcal{T}^A \lambda = \varkappa^A_M \circ T^A \lambda\,.$$

Every $v \otimes a \in \mathfrak{g} \otimes A$ defines the fundamental vector field $(\mathcal{T}^A \lambda)(v \otimes a)$ on $T^A M$. On the other hand, $\lambda(v)$ is a vector field on M and we can construct its flow prolongation $\mathcal{T}^A\big(\lambda(v)\big)$. The proof of the following assertion, which relates the previous concept in an interesting way, can be found in [10].

Proposition *We have*

$$(\mathcal{T}^A \lambda)(v \otimes a) = L(a)_M \circ \mathcal{T}^A\big(\lambda(v)\big)\,.$$

3.6 The linear case

Consider the case $M = V$ is a vector space. Then $TV = V \times V$ and the first component of $\lambda : \mathfrak{g} \times V \to V \times V$ is the product projection $\mathfrak{g} \times V \to V$. The second component will be denoted by

$$\overline{\lambda} : \mathfrak{g} \times V \to V\,.$$

Since $T^A V = V \otimes A$ is also a vector space, we have $T^A TV = V \otimes A \times V \otimes A$ and $TT^A V = V \otimes A \times V \otimes A$. Under these identifications, $\varkappa_V^A$ is the identity of $V \otimes A \times V \otimes A$.

For the infinitesimal action $\mathcal{T}^A \lambda : T^A \mathfrak{g} \times T^A V \to TT^A V$, we have

$$\overline{\mathcal{T}^A \lambda} : T^A \mathfrak{g} \times T^A V \to T^A V \,.$$

Then our previous results yield

Proposition *We have*

$$\overline{\mathcal{T}^A \lambda} = T^A \overline{\lambda} : T^A \mathfrak{g} \times T^A V \to T^A V \,.$$

In particular, let l be a linear action of G on V, so that $\overline{\lambda}$ is the classical representation of Lie algebra $\mathfrak{g}$ on V. Hence $\overline{\lambda}$ is a bilinear map. By 2.8, $\overline{\mathcal{T}^A \lambda}$ is of the form

$$\overline{\mathcal{T}^A \lambda}(v \otimes a_1, z \otimes a_2) = \overline{\lambda}(v, z) \otimes a_1 a_2 \,,$$

$v \in \mathfrak{g}$, $z \in V$, the product $a_1 a_2$ being in A.

3.7 Vector bundles

For a vector bundle $p : E \to M$, $T^A p : T^A E \to T^A M$ is also a vector bundle. If $X_1, X_2 \in T^A E$ satisfy $T^A p(X_1) = T^A p(X_2)$, we may write $X_1 = j^A \varphi_1$, $X_2 = j^A \varphi_2$ with $p \circ \varphi_1 = p \circ \varphi_2$, so that $\varphi_1(u)$ and $\varphi_2(u)$ are in the same fiber of $E \to M$ for all $u \in \mathbb{R}^k$. Then we define $X_1 + X_2$ by $j^A\big(\varphi_1(u) + \varphi_2(u)\big)$. Similarly, $c\big(j^A \varphi(u)\big) = j^A\big(c\varphi(u)\big)$, $c \in \mathbb{R}$. Further, if $\overline{p} : \overline{E} \to \overline{M}$ is another vector bundle and $f : E \to \overline{E}$ is a linear morphism over $\underline{f} : M \to \overline{M}$, then $T^A f : T^A E \to T^A \overline{E}$ is a linear morphism over $T^A \underline{f} : T^A M \to T^A \overline{M}$.

3.8 Principal and associated bundles

Let $P(M, G)$ be a principal bundle with structure group G and projection $p : P \to M$. Write $\varrho_P : P \times G \to P$ for the right action of G on P. Then $T^A p : T^A P \to T^A M$ is a principal bundle with structure group $T^A G$ and $\varrho_{T^A P} = T^A \varrho_P : T^A P \times T^A G \to T^A P$.

Consider a fiber bundle $E = P[S, l]$ associated to P with respect to an action $l : G \times S \to S$ of G on the standard fiber S and write $q : E \to M$ for the bundle projection. Then $T^A q : T^A E \to T^A M$ is an associated bundle $T^A E = T^A P[T^A S, T^A l]$.

For every fibered manifold $Y \to M$, the vertical Weil bundle $V^A Y$ is the union

$$V^A Y = \bigcup_{x \in M} T^A(Y_x) \,, \qquad V^A Y \subset T^A Y$$

of the Weil bundles of the individual fibers of Y. Clearly, $V^A P$ is a principal bundle $V^A P(M, T^A G)$ and $V^A E$ is an associated bundle $V^A E = V^A P[T^A S, T^A l]$.

3.9 Tensor fields of type $(1, k)$

A tensor field C of type $(1, k)$ on a manifold M can be considered as a map

$$C : TM \underbrace{\times_M \cdots \times_M}_{k\text{-times}} TM \to TM \,.$$

Applying functor T^A, we obtain

$$T^A C : T^A TM \times_{T^A M} \cdots \times_{T^A M} T^A TM \to T^A TM\,.$$

Using the canonical exchange $\varkappa_M^A$, we construct

$$\mathcal{T}^A C = \varkappa_M^A \circ T^A C \circ \big((\varkappa_M^A)^{-1} \times \cdots \times (\varkappa_M^A)^{-1}\big)\,.$$

This is a tensor field of type $(1,k)$ on $T^A M$, which is called the complete lift of C to $T^A M$. In the special case $k=0$, C is a vector field on M and $\mathcal{T}^A C$ is its flow prolongation, see 2.11.

For every k vector fields $X_1,\dots,X_k$ on M, $C(X_1,\dots,X_k)$ is also a vector field. The following assertion is deduced in [7].

Proposition *For every $a_1,\dots,a_k \in A$, we have*

$$\mathcal{T}^A C\big(L(a_1)\mathcal{T}^A X_1,\dots,L(a_k)\mathcal{T}^A X_k\big) = L(a_1\dots a_k)\mathcal{T}^A\big(C(X_1,\dots,X_k)\big)\,.$$

3.10 The Frölicher-Nijenhuis bracket

An antisymmetric tensor field P of type $(1,k)$ on M is said to be a tangent valued k-form on M. The Frölicher-Nijenhuis bracket is an important geometric operation on the tangent valued forms, see e.g. [14]. If Q is another tangent valued l-form on M, then the Frölicher-Nijenhuis bracket $[P,Q]$ is a tangent valued $(k+l)$-form on M. A tangent valued 0-form is a vector field. If both P and Q are tangent valued 0-forms, then $[P,Q]$ coincides with the classical bracket of vector fields.

The identity of TM can be interpreted as a tangent valued 1-form on M. For every P, we have

$$\big[\,\mathrm{id}_{TM}, P\big] = 0\,. \tag{17}$$

Given a tangent valued form S on $T^A M$ and $a \in A$, $L(a)_M \circ S =: L(a)S$ is a tangent valued form on $T^A M$, too.

Theorem *For every tangent valued k-form P, every tangent valued l-form Q on M and every $a_1, a_2 \in A$, we have*

$$\big[L(a_1)\mathcal{T}^A P, L(a_2)\mathcal{T}^A Q\big] = L(a_1 a_2)\mathcal{T}^A\big([P,Q]\big)\,.$$

The proof is based on a formula by M. Modugno and P. W. Michor that expresses $[P,Q]$ in terms of the bracket of vector fields, [2].

For $a_1 = a_2 = 1$ we obtain

Corollary *The T^A-prolongation of tangent valued forms commutes with the Frölicher-Nijenhuis bracket.*

3.11 Connections

The first jet prolongation $J^1 Y$ of a fibered manifold $p: Y \to M$ is the bundle of 1-jets of local sections of Y. The elements of $J^1 Y$ are identified with the m-dimensional linear subspaces in TY complementary to the vertical tangent space, $m = \dim M$. Hence a connection on an arbitrary fibered manifold Y can be interpreted as a section $\Gamma : Y \to J^1Y$.

The projection $T_yY \to V_yY$ in the direction of $\Gamma(y)$, $y \in Y$, defines the connection form $\omega_\Gamma : TY \to VY \subset TY$. Clearly, ω_Γ is a tangent valued 1-form on TY satisfying $\omega_\Gamma(X) = X$ for all $X \in VY$. If we denote by $i_{VY} : VY \to TY$ the injection, this can be expressed by

$$\omega_\Gamma \circ i_{VY} = i_{VY} \,. \tag{18}$$

Conversely, if ω is a tangent valued 1-form on Y satisfying $\omega(TY) \subset VY$ and (18), then the kernels of ω determine a unique connection $\Gamma = \mathrm{Ker}\ \omega$ such that $\omega_\Gamma = \omega$.

The curvature C_Γ of Γ can be identified with the Frölicher-Nijenhuis bracket, see [14],

$$C_\Gamma = \frac{1}{2}\big[\omega_\Gamma, \omega_\Gamma\big] \,.$$

The T^A-prolongation $\mathcal{T}^A\omega_\Gamma$ of ω_Γ is a tangent valued 1-form on T^AY. Taking into account $\varkappa_Y^A(T^AVY) = V(T^AY \to T^AM)$, one verifies easily

$$\mathcal{T}^A\omega_\Gamma \circ i_{V(T^AY\to T^AM)} = i_{V(T^AY\to T^AM)} \,.$$

Hence there is a unique connection $\mathcal{T}^A\Gamma$ on $T^AY \to T^AM$ such that $\mathcal{T}^A(\omega_\Gamma) = \omega_{\mathcal{T}^A\Gamma}$. It will be called the T^A-prolongation of Γ.

Corollary 3.10 yields a formula for the curvature $C_{\mathcal{T}^A\Gamma} = \frac{1}{2}\big[\omega_{\mathcal{T}^A\Gamma}, \omega_{\mathcal{T}^A\Gamma}\big]$.

Proposition *We have* $C_{\mathcal{T}^A\Gamma} = \mathcal{T}^A(C_\Gamma)$.

3.12 Vector valued forms

Let V be a vector space. A V-valued k-form ω on M can be interpreted as a map

$$\omega : TM \times_M \cdots \times_M TM \to V \,.$$

Applying T^A and using $\varkappa_M^A$, we obtain

$$\mathcal{T}^A\omega : TT^AM \times_{T^AM} \cdots \times_{T^AM} TT^AM \to T^AV \,,$$

which is a $V \otimes A$-valued k-form on T^AM. Taking into account 3.3, one finds easily that this operation commutes with the exterior differentiation, i.e.

$$\mathcal{T}^A(d\omega) = d(\mathcal{T}^A\omega) \,.$$

We describe the coordinate form of $\mathcal{T}^A\omega$. Let $M = \mathbb{R}^m$, $V = \mathbb{R}^n$ and

$$y^p = f^p_{i_1\dots i_k}(x^1, \dots, x^m)\, dx^{i_1} \wedge \cdots \wedge dx^{i_k}$$

be the coordinate expression of ω. We have $TT^A\mathbb{R}^m = A^m \times A^m$ and we write a^i, da^i for the corresponding algebra coordinates. The coordinate formula for $T^Af^p_{i_1\dots i_k}$ is described in 2.4. Then 2.8 implies that $\mathcal{T}^A\omega$ is of the form

$$\big(T^Af^p_{i_1\dots i_k}\big)da^{i_1} \wedge \cdots \wedge da^{i_k}$$

with all products in A.

For example, consider the tangent functor T and $n = 1$, $k = 2$, so that

$$\omega = f_{i_1 i_2}(x^1, \dots, x^m)\, dx^{i_1} \wedge dx^{i_2}\,.$$

Using (12), we obtain

$$\begin{aligned}\mathcal{T}\omega &= \Big(f_{i_1 i_2} + e\tfrac{\partial f_{i_1 i_2}}{\partial x^i}\, x^i_1\Big)(dx^{i_1} + edx^{i_1}_1) \wedge (dx^{i_2} + edx^{i_2}_1)\\ &= f_{i_1 i_2} dx^{i_1} \wedge dx^{i_2}\\ &\quad + e\Big(\Big(\tfrac{\partial f_{i_1 i_2}}{\partial x^i}\, x^i_1\Big)dx^{i_1} \wedge dx^{i_2} + f_{i_1 i_2}(dx^{i_1} \wedge dx^{i_2}_1 + dx^{i_1}_1 \wedge dx^{i_2})\Big)\end{aligned}$$

In coordinates, a tangent valued k-form P looks like a vector valued k-form. So the procedure of finding the coordinate expression of $\mathcal{T}^A P$ is the same.

3.13 Connections in the lifting form

Taking into account the projection $T_y Y \to \Gamma(y)$ in the direction of $V_y Y$, we can interpret Γ as the lifting map (denoted by the same symbol)

$$\Gamma : Y \times_M TM \to TY\,.$$

Conversely, let $\Phi : Y \times_M TM \to TY$ be a map linear in TM and such that $\pi_Y \circ \Phi = pr_1$, $Tp \circ \Phi = pr_2$, where $\pi_Y : TY \to Y$ is the bundle projection. Then there is a unique connection Γ on Y such that Φ is its lifting map.

The lifting map of $\mathcal{T}^A\Gamma$ is

$$\mathcal{T}^A\Gamma := \varkappa^A_Y \circ \big(T^A\Gamma \circ (\mathrm{id}_{T^A Y} \times_{T^A M} (\varkappa^A_M)^{-1})\big)\,,$$

where $T^A\Gamma : T^A Y \times_{T^A M} T^A TM \to T^A TY$.

The algebra multiplication enables us to evaluate the equations of $\mathcal{T}^A\Gamma$ in the following simple way, [3]. If we have the product bundle $\mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m$ with the canonical coordinates x^i, y^p, then the equations of Γ are

$$dy^p = F^p_i(x, y)\, dx^i\,. \tag{19}$$

The Weil bundle $T^A(\mathbb{R}^m \times \mathbb{R}^n) \to T^A\mathbb{R}^m$ is $A^m \times A^n \to A^m$ with the algebra coordinates a^i, b^p. Then $\mathcal{T}^A(dx^i) = da^i$ and $\mathcal{T}^A(dy^p) = db^p$ are the additional coordinates on $TT^A(\mathbb{R}^m \times \mathbb{R}^m)$. Applying T^A to (19), we obtain the equations of $\mathcal{T}^A\Gamma$ in the form

$$db^p = (T^A F^p_i)\, da^i\,,$$

where the A-valued functions $T^A F^p_i$ can be evaluated by (11) and the products on the right hand side are in A.

For example, consider the tangent functor T. By (12),

$$TF^p_i = F^p_i + e\Big(\tfrac{\partial F^p_i}{\partial x^j}\, x^j_1 + \tfrac{\partial F^p_i}{\partial y^q}\, y^q_1\Big)$$

and we have $da^i = dx^i + edx^i_1$, $db^p = dy^p + edy^p_1$. Then

$$(dy^p + edy^p_1) = \Big(F^p_i + e\Big(\tfrac{\partial F^p_i}{\partial x^j}\, x^j_1 + \tfrac{\partial F^p_i}{\partial y^q}\, y^q_1\Big)\Big)\,(dx^i + edx^i_1)$$

implies that the equations of $\mathcal{T}\Gamma$ are (19) and

$$dy_1^p = \left(\frac{\partial F_i^p}{\partial x^j}\, x_1^j + \frac{\partial F_i^p}{\partial y^q}\, y_1^q\right) dx^i + F_i^p\, dx_1^i\,.$$

Of course, the aim of this simplest example is to illustrate the procedure. The power of the Weil algebra technique appears properly if we consider a more complicated algebra A.

3.14 Curvature in the lifting form

Consider Γ in the lifting form. The Γ-lift of a vector field X on M is a projectable vector field ΓX on Y defined by $\Gamma X = \Gamma(-, X)$. Then the curvature C_Γ is identified with a morphism

$$C_\Gamma : Y \times_M TM \times_M TM \to VY \tag{20}$$

defined as follows. Two vectors $\xi_1, \xi_2 \in T_{p(y)}M$ are extended into some vector fields X_1, X_2 on M and

$$C_\Gamma(y, \xi_1, \xi_2) = [\Gamma X_1, \Gamma X_2](y) - \Gamma\big([X_1, X_2]\big)(y)\,.$$

Using the lifting form of Γ, we construct

$$T^A C_\Gamma : T^A Y \times_{T^A M} T^A TM \times_{T^A M} T^A TM \to T^A VY\,.$$

Then we have the following characterization of the curvature of $\mathcal{T}^A\Gamma$.

Proposition *If C_Γ is in the form (20), then*

$$C_{\mathcal{T}^A\Gamma} = \varkappa_Y^A \circ T^A C_\Gamma \circ \big(\mathrm{id}_{T^A Y} \times_{T^A M} (\varkappa_M^A)^{-1} \times_{T^A M} (\varkappa_M^A)^{-1}\big)\,.$$

3.15 Principal and linear connections

Connections in the sense of 3.11 represent a generalization of the classical concept of connection. In the classical theories, connections are studied on principal bundles and are assumed to be right-invariant. Nowadays, a right-invariant connection on P is said to be principal. Similarly, if $E \to M$ is a vector bundle, so that $J^1E \to M$ is also a vector bundle, a connection $\Gamma : E \to J^1E$ is said to be linear, if Γ is a linear morphism.

Consider a connection Γ on a principal bundle $P(M, G)$. Each fiber of $VP \to P$ is identified with $\mathfrak{g}$ in the standard way. Hence the connection form ω_Γ of Γ can be interpreted as a 1-form on P with values in the vector space $\mathfrak{g}$. If Γ is principal, then $\omega_\Gamma : TP \to \mathfrak{g}$ is the classical connection form of Γ. One deduces easily

Proposition *If Γ is a principal connection on P with connection form $\omega_\Gamma : TP \to \mathfrak{g}$, then $\mathcal{T}^A\Gamma$ is a principal connection on $T^AP \to T^AM$ with connection form $\omega_{\mathcal{T}^A\Gamma} = \mathcal{T}^A(\omega_\Gamma)$.*

Analogously, let Γ be a linear connection on a vector bundle $E \to M$. Then the induced connection $\mathcal{T}^A\Gamma$ on $T^AE \to T^AM$ is also linear.

3.16 a-torsions

The Frölicher-Nijenhuis bracket can be applied in the theory of torsions of connections on a Weil bundle T^AM. This was pointed out in [17]. We present the basic ideas in a slightly more general setting.

For every $a \in A$, $L(a)_M$ is a natural tangent valued 1-form on $T^A M$. Consider a natural fibered manifold $T^A M \to Q$. (For example, Q is M or one of the underlying Weil bundles from 2.13. In the case of a fibered manifold $Y \to M$, we can consider $T^A Y \to T^A M$.) Then every connection Γ on $T^A M \to Q$ can be viewed as a tangent valued 1-form ω_Γ on $T^A M$.

Definition The Frölicher-Nijenhuis bracket $\left[L(a)_M, \omega_\Gamma\right]$ is called a-torsion of Γ.

Since $L(1)_M = \mathrm{id}_{TT^A M}$ and $[\mathrm{id}, \omega_\Gamma] = 0$ by (17), only the elements from the nilpotent part of A are interesting.

We have to explain how our definition generalizes the classical notion of torsion. The Weil algebra of the tangent functor T is $\mathbb{D}$ with one-dimensional nilpotent part generated by e, see 1.4. Let Γ be a classical linear connection on $TM \to M$. Using simple evaluation, one deduced that $\left[L(e)_M, \omega_\Gamma\right]$ is the classical torsion of Γ.

Consider the T^A-prolongation $\mathcal{T}^A\Gamma$ of a connection Γ on $Y \to M$. Then $\mathcal{T}^A\Gamma$ is torsion free in the following sense.

Proposition *We have* $\left[L(a)_Y, \omega_{\mathcal{T}^A\Gamma}\right] = 0$ *for all* $a \in A$.

Proof We have $L(a)_Y = L(a) \circ \mathrm{id}_{TT^A Y}$. Hence Theorem 3.10 and (17) imply $\left[L(a)_Y, \omega_{\mathcal{T}^A\Gamma}\right] = L(a)\mathcal{T}^A\big([\mathrm{id}_{TY}, \omega_\Gamma]\big) = 0$. □

3.17 Some further structures

The first paper dealing systematically with T^A-prolongation of various geometric structures is by A. Morimoto, [21]. Beside some subjects we already mentioned, he studied T^A-prolongation of almost complex structures and the prolongation of a classical linear connection on M into a classical linear connection on $T^A M$.

Even the paper [7] by J. Gancarzewicz, W. Mikulski and Z. Pogoda is devoted, beside the general theory, to T^A-prolongation of further geometric structures. Main attention is paid to Riemannian and pseudo-Riemannian metrics, symplectic and almost symplectic structures, almost tangent structures and Kählerian structures.

3.18 Remark

All results of this paper are valid for an arbitrary Weil algebra. However, we have to point out that there are further geometric problems, in which certain special kinds of Weil algebras are specified. We mention solely a paper by M. Kureš and W. Mikulski, [19], on the natural operators transforming vector fields from a manifold M into vector fields on the bundle $K^A M$ of contact A-elements introduced in 2.15. In [19], the best results are deduced for homogeneous Weil algebras. We recall that an ideal $I \subset \mathbb{R}[x_1, \ldots, x_k]$ is said to be homogeneous, if $P \in I$ implies that all homogeneous components of polynomial P are also in I. A Weil algebra is called homogeneous, if it can be expressed in the form $\mathbb{R}[x_1, \ldots, x_k]/I$ with a homogeneous ideal I. Clearly, all algebras $\mathbb{D}^r_k$ and their tensor products are homogeneous. Examples of nonhomogeneous Weil algebras are constructed in [19].

4 Fiber product preserving bundle functors

4.1 Jet functors

There are 3 bundle functors, whose construction is based on the concept of r-jet only.

For every two manifolds M and N, $J^r(M,N)$ is the bundle of all r-jets of M into N. Let $f : M \to \overline{M}$ be a local diffeomorphism and $g : N \to \overline{N}$ be a map. Then the induced map $J^r(f,g) : J^r(M,N) \to J^r(\overline{M},\overline{N})$ is defined by

$$J^r(f,g)(X) = (j^r_y g) \circ X \circ (j^r_x f)^{-1}\,,$$

where x or y is the source or the target of $X \in J^r(M,N)$. Hence J^r is a bundle functor defined on the product category $\mathcal{M}f_m \times \mathcal{M}f$, $m = \dim M$.

For every fibered manifold $p : Y \to M$, J^rY is the bundle of r-jets of local sections of Y. If $\overline{p} : \overline{Y} \to \overline{M}$ is another fibered manifold and $f : Y \to \overline{Y}$ is an $\mathcal{FM}$-morphism such that the base map $\underline{f} : M \to \overline{M}$ is a local diffeomorphism, then the map $J^r(\underline{f}, f) : J^r(M,Y) \to J^r(\overline{M},\overline{Y})$ transforms J^rY into $J^r\overline{Y}$. The restricted and corestricted map $J^rf : J^rY \to J^r\overline{Y}$ is called the r-th jet prolongation of f. Hence J^r is a functor on the category $\mathcal{FM}_m$. In concrete problems, it is always clear which functor J^r is under consideration and we have no intention to change this convention. But in this theoretical section we have to distinguish. So, in this section we shall write J^r_h in the case of $\mathcal{FM}_m$.

The r-th vertical jet prolongation J^r_vY, which is used e.g. in the theory of higher order absolute differentiation, is defined by

$$J^r_vY = \bigcup_{x\in M} J^r_x(M, Y_x)\,.$$

The restriction and corestriction of $J^r(\underline{f}, f)$ defines $J^r_vf : J^r_vY \to J^r_v\overline{Y}$. Hence J^r_v is a bundle functor on $\mathcal{FM}_m$. In this context, we also say that J^r_hY is the r-th horizontal jet prolongation of Y.

The construction of product fibered manifolds is a functor $i : \mathcal{M}f_m \times \mathcal{M}f \to \mathcal{FM}_m$, $i(M \times N) = (M \times N \to M)$ and $i(f \times g)$ is $f \times g$ with the base map f. We have

$$J^r_h(M \times N) = J^r(M,N) = J^r_v(M \times N)$$

and both functors $J^r_h \circ i$ and $J^r_v \circ i$ coincide with J^r.

4.2 Nonholonomic jets

The r-th nonholonomic prolongation $\widetilde{J}^r_hY$ of a fibered manifold Y is defined by the iteration

$$\widetilde{J}^r_hY = J^1_h(\widetilde{J}^{r-1}_hY \to M)\,,$$

$\widetilde{J}^1_hY = J^1_hY$. For every $f : Y \to \overline{Y}$ in $\mathcal{FM}_m$, the iteration determines

$$\widetilde{J}^r_hf = J^1_h(\widetilde{J}^{r-1}_hf) : \widetilde{J}^r_hY \to \widetilde{J}^r_h\overline{Y}\,.$$

Hence $\widetilde{J}^r_h$ is a bundle functor on $\mathcal{FM}_m$. The canonical inclusion $J^r_hY \hookrightarrow \widetilde{J}^r_hY$ is defined by the iteration $j^r_xs \mapsto j^1_x(u \mapsto j^{r-1}_us)$ for every local section s of Y, $u \in M$.

The restriction $\widetilde{J}^r_h \circ i$ yields a functor $\widetilde{J}^r$ on $\mathcal{M}f_m \times \mathcal{M}f$. The space

$$\widetilde{J}^r(M,N) = \widetilde{J}^r_h(M \times N \to M)$$

is the bundle of nonholonomic r-jets of manifold M into manifold N defined by Ehresmann. The elements of $J^r(M,N) \subset \widetilde{J}^r(M,N)$ are also said to be holonomic r-jets. Let Q be a third manifold. Ehresmann introduced the composition of nonholonomic r-jets by the following induction. For $r = 1$, we have the composition of 1-jets. Write $\beta : \widetilde{J}^{r-1}(M,N) \to N$ for the canonical projection. Let $X = j^1_x s(u) \in \widetilde{J}^r_x(M,N)_y$, $u \in M$, and $Z = j^1_y \sigma \in \widetilde{J}^r_y(N,Q)_z$, $y = \beta\big(s(x)\big)$. Then

$$Z \circ X := j^1_x\big(\sigma\big(\beta(s(u))\big) \circ s(u)\big) \in \widetilde{J}^r_x(M,Q)_z \qquad (21)$$

with the composition of nonholonomic $(r-1)$-jets on the right hand side. If X and Z are holonomic r-jets, then (21) coincides with the classical composition. The composition of nonholonomic r-jets is associative.

The r-th vertical nonholonomic prolongation of Y is defined by

$$\widetilde{J}^r_v Y = \bigcup_{x \in M} \widetilde{J}^r_x(M, Y_x)\,.$$

Even this is a bundle functor on $\mathcal{FM}_m$. Similarly to the holonomic case, we have $\widetilde{J}^r_h \circ i = \widetilde{J}^r = \widetilde{J}^r_v \circ i$.

4.3 Bundle functors in the product case

We start with some properties of bundle functors on $\mathcal{M}f \times \mathcal{M}f$ that are needed for the main subject of this section. First we introduce some notation.

Let F be a bundle functor on $\mathcal{M}f \times \mathcal{M}f$. We write $F_x(M,N)$ or $F_x(M,N)_y$ for the submanifold of all elements of $F(M,N)$ over $x \in M$ or $(x,y) \in M \times N$, respectively. For $g : M \to \overline{M}$ and $f : N \to \overline{N}$, we write $F_x(g,f) : F_x(M,N) \to F_{g(x)}(\overline{M},\overline{N})$ and $F_x(g,f)_y : F_x(M,N)_y \to F_{g(x)}(\overline{M},\overline{N})_{f(y)}$ for the restricted and corestricted maps.

Definition We say that F preserves products in the second factor, if $F(M, N_1 \times N_2) = F(M,N_1) \times_M F(M,N_2)$. We say that F has order r in the first factor, if for every $g, \overline{g} : M \to \overline{M}$ and $f : N \to \overline{N}$, $j^r_x g = j^r_x \overline{g}$ implies

$$F_x(g,f) = F_x(\overline{g},f) : F_x(M,N) \to F_{g(x)}(\overline{M},\overline{N})\,.$$

Clearly, if F has order r in the standard sense, i.e. $j^r_{x,y}(g,f) = j^r_{x,y}(\overline{g},\overline{f})$ implies $F_x(g,f)_y = F_x(\overline{g},\overline{f})_y$, then F has order r in the first factor.

4.4 The case of $\mathcal{M}f_m \times \mathcal{M}f$

Let F be a bundle functor on $\mathcal{M}f_m \times \mathcal{M}f$ that preserves products in the second factor. We define an associated bundle functor G^F on $\mathcal{M}f$ by $G^F(N) = F_0(\mathbb{R}^m, N)$ and $G^F(f) = F_0(\mathrm{id}_{\mathbb{R}^m}, f) : G^F N \to G^F \overline{N}$. Clearly, G^F preserves products, so that $G^F = T^A$ for a Weil algebra A.

Assume further that F has order r in the first factor. For $X \in G^r_m$, $X = j^r_0\gamma$ and every manifold N, we set

$$H^F(X)_N = F_0(\gamma, \mathrm{id}_N) : G^F N \to G^F N\,.$$

Since the following diagram commutes

$$\begin{array}{ccc} (\mathbb{R}^m, N) & \xrightarrow{(\mathrm{id}_{\mathbb{R}^m}, f)} & (\mathbb{R}^m, \overline{N}) \\ \Big\downarrow{\scriptstyle (\gamma, \mathrm{id}_N)} & & \Big\downarrow{\scriptstyle (\gamma, \mathrm{id}_{\overline{N}})} \\ (\mathbb{R}^m, N) & \xrightarrow{(\mathrm{id}_{\mathbb{R}^m}, f)} & (\mathbb{R}^m, \overline{N}) \end{array}$$

each $H^F(X)$ is a natural equivalence $T^A \to T^A$. By 2.7, $H^F(X)$ corresponds to an element of $\mathrm{Aut}\, A$. Clearly, $H^F : G^r_m \to \mathrm{Aut}\, A$ is a group homomorphism.

Conversely, consider a Weil algebra A and a group homomorphism $H : G^r_m \to \mathrm{Aut}\, A$. For every manifold N, we have the induced left action H_N of G^r_m on $T^A N$, so that we can construct the associated bundle $P^r M[T^A N, H_N] =: (A, H)(M, N)$ for every m-dimensional manifold M. We underline that the elements of $P^r M[T^A N, H_N]$ are the equivalence classes

$$\{u, Z\}\,, \qquad u \in P^r M,\ Z \in T^A N\,.$$

For every local diffeomorphism $g : M \to \overline{M}$, we have the induced morphism $P^r g : P^r M \to P^r \overline{M}$ of principal bundles and every map $f : N \to \overline{N}$ induces a G^r_m-equivariant map $T^A f : T^A N \to T^A \overline{N}$. So we can construct the morphism of associated bundles $(A, H)(g, f) := P^r g[T^A f] : P^r M[T^A N, H_N] \to P^r \overline{M}[T^A \overline{N}, H_{\overline{N}}]$. Clearly, (A, H) is a functor. Thus, we have proved

Proposition *The above construction establishes a bijection between the bundle functors on $\mathcal{M}f_m \times \mathcal{M}f$ that preserve products in the second factor and have order r in the first factor and the pairs (A, H) of a Weil algebra A and a group homomorphism $H : G^r_m \to$ $\mathrm{Aut}\, A$.*

In particular, if $M = \mathbb{R}^m$, then $P^r\mathbb{R}^m = \mathbb{R}^m \times G^r_m$ and the associated bundle $(A, H)(\mathbb{R}^m, N)$ is identified with $\mathbb{R}^m \times T^A N$. Given $f : N \to \overline{N}$, we have

$$(A, H)(\mathrm{id}_{\mathbb{R}^m}, f) = \mathrm{id}_{\mathbb{R}^m} \times T^A f : \mathbb{R}^m \times T^A N \to \mathbb{R}^m \times T^A \overline{N}\,.$$

If $\overline{F} = (\overline{A}, \overline{H})$ is another bundle functor on $\mathcal{M}f_m \times \mathcal{M}f$ of order r in the first factor, then the natural transformations $\tau : F \to \overline{F}$ are in bijection with the equivariant algebra homomorphisms $\mu : A \to \overline{A}$, i.e. $\mu\big(H(X)(a)\big) = \overline{H}(X)\big(\mu(a)\big)$ for all $a \in A$ and $X \in G^r_m$. We have

$$\tau_{M,N} = (\mathrm{id}_{P^r M}, \mu_N) : P^r M[T^A N, H_N] \to P^r M[T^{\overline{A}} N, \overline{H}_N]\,.$$

In the case $M = \mathbb{R}^m$, $\tau_{\mathbb{R}^m, N} = \mathrm{id}_{\mathbb{R}^m} \times \mu_N : \mathbb{R}^m \times T^A N \to \mathbb{R}^m \times T^{\overline{A}} N$.

4.5 The algebra $\widetilde{\mathbb{D}}^r_m$

We describe the functor $\widetilde{J}^r$ in this way. The space $\widetilde{T}^r_m N := \widetilde{J}^r_0(\mathbb{R}^m, N)$ is called bundle of nonholonomic (m, r)-velocities on N. Using the translations on $\mathbb{R}^m$, one identifies $\widetilde{T}^r_m N$ with the iteration $T^1_m \underbrace{\big(\dots}_{r\text{-times}}(T^1_m N)\dots\big)$. Since the Weil algebra of T^1_m is $\mathbb{D}^1_m$, 2.10 implies that the Weil algebra $\widetilde{\mathbb{D}}^r_m$ of $\widetilde{J}^r$ is $\mathbb{D}^1_m \underbrace{\otimes \dots \otimes}_{r\text{-times}} \mathbb{D}^1_m$. The action of $G^r_m = \operatorname{inv} J^r_0(\mathbb{R}^m, \mathbb{R}^m)_0$ on $\widetilde{\mathbb{D}}^r_m$ is given by the composition of nonholonomic jets.

4.6 The base order of F

Let F be a bundle functor on $\mathcal{FM}$. The definition of the order of F is based on the concept of (q, s, r)-jet, see 2.17. We say that F is of order (q, s, r), $s \geq q \leq r$, if for every $\mathcal{FM}$-morphism $f : Y \to \overline{Y}$, the restriction $Ff \mid F_y Y$ depends on $j^{q,s,r}_y f$ only, $y \in Y$.

The integer r is called the base order of F.

4.7 The main result

Let F be a fiber product preserving bundle (in short: f.p.p.b.) functor on $\mathcal{FM}_m$, i.e. $F(Y_1 \times_M Y_2) = FY_1 \times_M FY_2$ for every two fibered manifolds Y_1 and Y_2 over the same base M, $\dim M = m$. Its restriction $F \circ i$ to $\mathcal{M}f_m \times \mathcal{M}f$ preserves products in the second factor. In [15], it is deduced that F has finite order. The base order r of F coincides with the order of $F \circ i$ in the second factor. By 4.4, $F \circ i = (A, H)$, so that $F(M \times N) = P^r M[T^A N, H_N]$.

Further, F determines a natural transformation $\widetilde{t}_Y : J^r_h Y \to FY$. Every element $X \in J^r_h Y$ is of the form $j^r_x s$. We interpret the local section s of Y as a local $\mathcal{FM}_m$-morphism $\widetilde{s}$ of the trivial fibered morphism $\mathrm{id}_M : M \to M$ (denoted by $i_1 M$ in 2.16) into Y and we set

$$\widetilde{t}_Y(X) = (F\widetilde{s})(x) \in FY\,. \tag{22}$$

In the product case $Y = \mathbb{R}^m \times N$, we have $J^r_h(\mathbb{R}^m \times N) = \mathbb{R}^m \times T^r_m N$ and $F(\mathbb{R}^m \times N) = \mathbb{R}^m \times T^A N$. Write

$$t_N : T^r_m N \to T^A N$$

for the restricted and corestricted map over $0 \in \mathbb{R}^m$. This is a natural transformation, so that it corresponds to an algebra homomorphism $t : \mathbb{D}^r_m \to A$. By naturality of $\widetilde{t}$, t is a G^r_m-equivariant algebra homomorphism, i.e.

$$t(X \circ g) = H(g)\big(t(X)\big)\,, \qquad X \in \mathbb{D}^r_m,\ g \in G^r_m\,.$$

The inclusion $Y \hookrightarrow M \times Y$, $y \mapsto \big(p(x), y\big)$ is a base preserving morphism. Applying F, we obtain an inclusion

$$FY \hookrightarrow F(M \times Y) = P^r M[T^A Y, H_Y]\,.$$

Since $P^rM \subset T^r_mM$, t defines a natural map (denoted by the same symbol) $t_M : P^rM \to T^AM$. According to [15], FY is the space of all equivalence classes $\{u, Z\}$ satisfying $t_M(u) = T^Ap(Z)$, $u \in P^rM$, $Z \in T^AY$.

Conversely, such a triple (A, H, t) defines a bundle functor $F = (A, H, t)$ on $\mathcal{FM}_m$ by

$$FY = \big\{\{u, Z\}; u \in P^rM, Z \in T^AY, t_M(u) = T^Ap(Z)\big\} . \tag{23}$$

For an $\mathcal{FM}_m$-morphism $f : Y \to \overline{Y}$ over $\underline{f} : M \to \overline{M}$, $(A, H)(\underline{f}, f) = P^r\underline{f}[T^Af]$ maps FY into $F\overline{Y}$ and Ff is its restriction and corestriction. This implies

Theorem *The f.p.p.b. functors of base order r on $\mathcal{FM}_m$ are in bijection with the triples (A, H, t), where A is a Weil algebra, $H : G^r_m \to \operatorname{Aut} A$ is a group homomorphism and $t : \mathbb{D}^r_m \to A$ is an equivariant algebra homomorphism. The natural transformations $(A, H, t) \to (\overline{A}, \overline{H}, \overline{t})$ are in bijection with the equivariant algebra homomorphisms $\mu : A \to \overline{A}$ satisfying $\overline{t} = \mu \circ t$.*

The second assertion follows from the fact that $\overline{t} = \mu \circ t$ implies that the natural transformation $(A, H) \to (\overline{A}, \overline{H})$ maps $(A, H, t)(Y)$ into $(\overline{A}, \overline{H}, \overline{t})(Y)$.

If we use the inclusions $J^r_hY \subset P^rM[T^r_mY]$ and $FY \subset P^rM[T^AY]$, then the map $\widetilde{t}_Y : J^r_hY \to FY$ is of the form

$$\widetilde{t}_Y\big(\{u, X\}\big) = \big\{u, t_Y(X)\big\} . \tag{24}$$

4.8 The basic examples

The simplest examples of f.p.p.b. functors on $\mathcal{FM}_m$ are J^r_h, J^r_v and V^A. So every iteration of these functors is also a f.p.p.b. functor on $\mathcal{FM}_m$.

By 1.9, $\operatorname{Aut} \mathbb{D}^r_m = G^r_m$. Write C for the corresponding action of G^r_m on $\mathbb{D}^r_m$. One finds easily $J^r_h = (\mathbb{D}^r_m, C, \mathrm{id}_{\mathbb{D}^r_m})$ and $J^r_v = (\mathbb{D}^r_m, C, \mathcal{O})$, where $\mathcal{O} : \mathbb{D}^r_m \to \mathbb{D}^r_m$ is the zero homomorphism. Functor V^A has base order 0. In this case, $G^0_m = \{e\}$ is the one element group, $\mathbb{D}^0_m = \mathbb{R}$ and H maps e into $\mathrm{id}_{\mathbb{R}}$.

4.9 The iteration

Let $E = (B, K, u)$ be another f.p.p.b. functor on $\mathcal{FM}_m$ of base order s, so that $K : G^s_m \to \operatorname{Aut} B$ and $u : \mathbb{D}^s_m \to B$. The composition $F \circ E$ preserves fiber products as well. In [6], the following expression $F \circ E = (C, L, v)$ is deduced. Clearly, the Weil algebra is $C = B \otimes A$.

Then we construct the group homomorphism $L : G^{r+s}_m \to \operatorname{Aut}(B \otimes A)$. Write $\iota^{r,s}_m : G^{r+s}_m \to T^r_mG^s_m$ for the map

$$j^{r+s}_0\gamma \mapsto j^s_0\big(y \mapsto j^r_0(\gamma \circ \tau_y)\big) , \quad y \in \mathbb{R}^m ,$$

where τ_y is the translation on $\mathbb{R}^m$ transforming 0 into y. For every $g \in G^{r+s}_m$, we denote by $\beta_r(g) \in G^r_m$ the underlying r-jet. We interpret K or L as a map $K : G^s_m \times B \to B$ or $L : G^{r+s}_m \times T^AB \to T^AB$, respectively. Then $T^AK : T^AG^s_m \times T^AB \to T^AB$. By 3.4, t induces a group homomorphism $t_{G^s_m} : T^r_mG^s_m \to T^AG^s_m$. According to [6], we have

$$L(g, Z) = H\big(\beta_r(g)\big)_B\big(T^AK(t_{G^s_m}(\iota^{r+s}_m(g)), Z)\big) , \quad Z \in T^AB .$$

To find the algebra homomorphism $v : \mathbb{D}_m^{r+s} \to B \otimes A$, we consider the injection $i_m^{r,s} : \mathbb{D}_m^{r+s} \to T_m^r \mathbb{D}_m^s$,

$$j_0^{r+s}\gamma \mapsto j_0^r\big(y \mapsto j_0^s(\gamma \circ \tau_{\gamma(y)})\big)\,, \quad y \in \mathbb{R}^m\,.$$

Then we construct $T_m^r u : T_m^r \mathbb{D}_m^s \to T_m^r B$ and $t_B : T_m^r B \to T^A B$. By [6], we have

$$v = t_B \circ T_m^r u \circ i_m^{r,s}\,.$$

4.10 Applications

We present two simple geometric applications of 4.9. Consider the vertical Weil functor V^B in the role of E. The base order of both iterations $V^B F$ and FV^B is r. In the first or the second case, the action of G_m^r on $A \otimes B$ or $B \otimes A$ is $H \otimes \mathrm{id}_B$ or $\mathrm{id}_B \otimes H$ and the algebra homomorphism $\mathbb{D}_m^r \to A \otimes B$ or $\mathbb{D}_m^r \to B \otimes A$ is $t \otimes \mathrm{id}_B$ or $\mathrm{id}_B \otimes t$, respectively. Hence the exchange algebra homomorphism $\mathrm{ex} : A \otimes B \to B \otimes A$ is equivariant and satisfies $\mathrm{ex} \circ (t \otimes \mathrm{id}_B) = \mathrm{id}_B \otimes t$. Thus ex determines a canonical natural equivalence

$$\varkappa_Y^{B,F} : V^B(FY) \to F(V^B Y)\,. \tag{25}$$

The special case $F = J_h^r$ and $V^B = V$ is heavily used e.g. in the variational calculus on fibered manifolds.

Another result, which is proved in [6], clarifies certain "rigidity" properties of jet functors.

Proposition *The only natural transformation $J_h^r J_h^s \to J_h^r J_h^s$ is the identity.*

In particular, for $r = s = 1$ we obtain that there is no natural exchange map $\widetilde{J}_h^2 Y \to \widetilde{J}_h^2 Y$.

4.11 The general concept of jet functor

Several special kinds of nonholonomic r-jets are known. Ehresmann himself defined the r-th semiholonomic jet prolongation $\overline{J}_h^r Y$ of a fibered manifold $Y \to M$ by induction starting with $\overline{J}_h^1 Y = J_h^1 Y$. Then $\overline{J}_h^r Y$ is the space of 1-jets $j_x^1 s$, where s is a local section of $\overline{J}_h^{r-1} Y \to M$ satisfying $s(x) = j_x^1(\beta_{r-1} \circ s)$, $\beta_{r-1} : \overline{J}_h^{r-1} Y \to \overline{J}_h^{r-2} Y$ being the canonical projection. Clearly, we have $J_h^r Y \subset \overline{J}_h^r Y$. In the product case, $\overline{J}_h^r(M \times N) =: \overline{J}^r(M, N)$ is the bundle of semiholonomic r-jets of M into N. The composition of two semiholonomic r-jets in the sense of 4.2 is a semiholonomic r-jet as well. We can also construct the r-th vertical semiholonomic jet prolongation of Y by

$$\overline{J}_v^r Y = \bigcup_{x \in M} \overline{J}_x^r(M, Y_x)\,.$$

Using the viewpoint of Weil bundles, we introduced the general concept of r-th order jet functor, [11]. This is a bundle functor G on $\mathcal{M}f_m \times \mathcal{M}f$ satisfying

$$J^r \subset G \subset \widetilde{J}^r$$

and preserving products in the second factor. By 4.4, for the corresponding Weil algebra A we have $\mathbb{D}_m^r \subset A \subset \widetilde{\mathbb{D}}_m^r$, so that there is a canonical action C_A of G_m^r on A. Conversely, every Weil algebra A with this property defines an r-th order jet functor G. Using 4.7, we can construct its horizontal version $G_h = (A, C_A, i)$, where $i : \mathbb{D}_m \to A$ is the above injection, and the vertical version $G_v = (A, C_A, \mathcal{O})$, where $\mathcal{O} : \mathbb{D}^r_m \to A$ is the zero homomorphism. We have $J^r_h \subset G_h \subset \widetilde{J}^r_h$ and $J^r_v \subset G_v \subset \widetilde{J}^r_h$.

4.12 Remark

The f.p.p.b. functors were studied on the category $\mathcal{FM}_m$ because of the relations to jet bundles. Of course, one can be also interested in the f.p.p.b. functors on the whole category $\mathcal{FM}$. Even these functors can be characterized in terms of Weil algebras. The first results on this subject are presented in a recent paper [16].

5 Some applications

5.1 F-prolongation of vector bundles

Consider a fibered manifold $p : C \to M$ and a Lie group K. We say that C is a group bundle of type K, if each fiber is a Lie group and for every $x \in M$ there exists a neighbourhood U such that $p^{-1}(U) = U \times K$. The group compositions form a base preserving morphism $\nu : C \times_M C \to C$. For every f.p.p.b. functor $F = (A, H, t)$, $F\nu : FC \times_M FC \to FC$ endows $FC \to M$ with the structure of group bundle of type $T^A K$. If $\overline{C} \to \overline{M}$ is another group bundle, an $\mathcal{FM}$-morphism $f : C \to \overline{C}$ is called group bundle morphism, if its restriction to each fiber is a group homomorphism. If $f : C \to \overline{C}$ is a group bundle morphism with local diffeomorphism as base map, then $Ff : FC \to F\overline{C}$ is also a group bundle morphism.

In particular, if $E \to M$ is a vector bundle, then FE is a bundle of Abelian groups. The multiplication of the elements of E by reals can be interpreted as a base preserving morphism $\sigma : (M \times \mathbb{R}) \times_M E \to E$. By 4.7, $F(M \times \mathbb{R}) = P^r M[A, H]$. Hence

$$F\sigma : P^r M[A, H] \times_M FE \to FE\,.$$

But $\mathbb{R} \subset A$ is an H-invariant subspace, so that $M \times \mathbb{R}$ is a subbundle of $P^r M[A, H]$. This defines the multiplication by reals on FE and one verifies easily that $FE \to M$ is a vector bundle, too. Clearly, the bundle projection $FE \to E$ is a linear morphism. Further, if $\overline{E} \to \overline{M}$ is another vector bundle and $f : E \to \overline{E}$ is a linear morphism with local diffeomorphism as base map, then $Ff : FE \to F\overline{E}$ is also a linear morphism.

5.2 The flow natural map

In the case of a f.p.p.b. functor $F = (A, H, t)$ on $\mathcal{FM}_m$, we have the following analogy of the flow natural exchange map from 2.11. First consider a vector field ξ on M. Its flow prolongation $\mathcal{P}^r\xi$ is a right-invariant vector field on r-th order frame bundle $P^r M$, whose value at every $u \in P^r_x M$ depends on $j^r_x\xi$ only. This defines a map $i : P^r M \times_M J^r TM \to TP^r M$.

For a fibered manifold $p : Y \to M$, we shall consider TY as a fibered manifold $TY \to M$. Then $Tp : TY \to TM$ is a base preserving morphism that induces $FTp : FTY \to FTM$. Taking into account the natural transformation $\widetilde{t}_{TM} : J^r TM \to FTM$

from 4.7, we construct the fiber product

$$J^rTM \times_{FTM} FTY\,. \tag{26}$$

By 4.7, we have $FTY \subset P^rM[T^ATY]$. Consider $(X, \{u, Z\})$ from (26), $X \in J^r_xTM$, $u \in P^r_xM$, $Z \in T^ATY$. Write $i(u, X) = (\partial/\partial t)_0\gamma(t)$, $\gamma : \mathbb{R} \to P^rM$. By (26), $\varkappa^A_Y(Z) \in TT^AY$ can be expressed as $(\partial/\partial t)_0\zeta(t)$, where $\zeta : \mathbb{R} \to T^AY$ satisfies $t_M\big(\gamma(t)\big) = T^Ap\big(\zeta(t)\big)$ for all t. So $\{\gamma(t), \zeta(t)\}$ is a curve on FY and we define

$$\psi^F_Y\big(X, \{u, Z\}\big) = \tfrac{\partial}{\partial t}\big|_0\big\{\gamma(t), \zeta(t)\big\}\,.$$

By right-invariancy, this is independent of the choice of u. Hence we obtain a map

$$\psi^F_Y : J^rTM \times_{FTM} FTY \to TFY\,.$$

By 5.1, FTY is a vector bundle over $FY \times_M FTM$ and one finds easily that ψ^F_Y is linear in both J^rTM and FTY.

A projectable vector field η on Y over ξ on M can be interpreted as a base preserving morphism $\eta : Y \to TY$. Then we construct the functorial prolongation $F\eta : FY \to FTY$ as well as the r-th jet prolongation $j^r\xi : M \to J^rTM$. The values of $j^r\xi \times_{\mathrm{id}_M} F\eta$ are in (26). The proof of the following assertion can be found in [10].

Proposition *The flow prolongation $\mathcal{F}\eta$ of η satisfies*

$$\mathcal{F}\eta = \psi^F_Y \circ (j^r\xi \times_{\mathrm{id}_M} F\eta)\,.$$

It is useful to introduce also a modified map

$$\widetilde{\psi}^F_Y : J^rTM \times_{FTM} FTY \to J^rTM \times_{TM} TFY\,,$$
$$\widetilde{\psi}^F_Y\big(X, \{u, Z\}\big) = \big(X, \psi^F_Y(X, \{u, Z\})\big)\,.$$

One finds easily that $\widetilde{\psi}^F_Y$ is a diffeomorphism.

In the case $F = J^r$, we have $FTM = J^rTM$, so that

$$\psi^{J^r}_Y : J^rTY \to TJ^rY\,.$$

This map was constructed in another way by L. Mangiarotti and M. Modugno.

We present a local expression of ψ^F_Y. We take $Y = \mathbb{R}^m \times V$, where we may assume V is a vector space. The group homomorphism $H : G^r_m \to \mathrm{Aut}\,A$ induces a Lie algebra homomorphism $h : \mathfrak{g}^r_m \to \mathrm{Der}\ A$. We have $J^rT\mathbb{R}^m = T\mathbb{R}^m \times \mathfrak{g}^r_m$. Since $\widetilde{t}_{T\mathbb{R}^m}$ maps $\mathfrak{g}^r_m$ into $(N_A)^m$, it is

$$J^rT\mathbb{R}^m \times_{FT\mathbb{R}^m} FT(\mathbb{R}^m \times V) = T\mathbb{R}^m \times \mathfrak{g}^r_m \times V \otimes A \times V \otimes A\,.$$

On the other hand, $TF(\mathbb{R}^m \times V) = T\mathbb{R}^m \times V \otimes A \times V \otimes A$. For $z \in T\mathbb{R}^m$, $u \in \mathfrak{g}^r_m$, $v \otimes a \in V \otimes A$ and $w \in V \otimes A$, one finds

$$\psi^F_{\mathbb{R}^m \times V}(z, u, v \otimes a, w) = \big(z, v \otimes a, w + v \otimes h(u)(a)\big)\,. \tag{27}$$

Even in the classical case $F = J^r$, this formula is convenient for evaluating $\mathcal{J}^r\eta$.

5.3 F-prolongation of connections

The flow natural map ψ_Y^F can be used for constructing the F-prolongation of projectable tangent valued forms on Y. However, in this situation we need an auxiliary linear r-th order connection on the base M, i.e. a linear base preserving morphism

$$\Lambda : TM \to J^r TM$$

satisfying $\beta \circ \Lambda = \mathrm{id}_{TM}$. This fact is well known from the theory of connections, [14]. Let F be an arbitrary bundle functor on the category of local isomorphisms of fibered manifolds of base order r and Γ be a connection on Y. The flow prolongation of the lifted vector field ΓX depends on the r-jets of vector field $X : M \to TM$. This defines a map

$$\mathcal{F}\Gamma : FY \times_M J^r TM \to TFY\,.$$

If we add Λ to the second factor, we obtain the lifting map $FY \times_M TM \to TFY$ of a connection $\mathcal{F}(\Gamma, \Lambda)$ on FY, which is called the F-prolongation of Γ with respect to Λ.

5.4 Tangent valued forms

For the sake of simplicity, we consider a projectable tangent valued 1-form $Q : TY \to TY$ over $\underline{Q} : TM \to TM$, that means $Tp \circ Q = \underline{Q} \circ Tp$. We have to interpret both Q and $\underline{Q}$ as base preserving morphisms over M. Then we construct the induced map

$$J^r \underline{Q} \times_{F\underline{Q}} FQ : J^r TM \times_{FTM} FTY \to J^r TM \times_{FTM} FTY\,.$$

Consider the following diagram

$$\begin{array}{ccc} J^r TM \times_{FTM} FTY & \xrightarrow{\;J^r\underline{Q} \times_{F\underline{Q}} FQ\;} & J^r TM \times_{FTM} FTY \\ \Big\downarrow {\scriptstyle \widetilde{\psi}_Y^F} & & \Big\downarrow {\scriptstyle \psi_Y^F} \\ J^r TM \times_{TM} TFY & \xrightarrow{\qquad\qquad} & TFY \end{array}$$

Since $\widetilde{\psi}_Y^F$ is invertible, the bottom arrow defines

$$\mathcal{F}Q = \psi_Y^F \circ (J^r \underline{Q} \times_{F\underline{Q}} FQ) \circ (\widetilde{\psi}_Y^F)^{-1} : J^r TM \times_{TM} TFY \to TFY\,.$$

Then 5.2 implies the following property of $\mathcal{F}Q$.

Proposition *Let η be a projectable vector field on Y over ξ on M. Then*

$$\mathcal{F}\big(Q(\eta)\big) = \mathcal{F}Q \circ (j^r\xi \times_{\mathrm{id}_{TM}} \mathcal{F}\eta)\,.$$

Definition For every linear r-th order connection $\Lambda : TM \to J^r TM$, the tangent valued 1-form on FY

$$\mathcal{F}(Q, \Lambda) := \mathcal{F}Q \circ (\Lambda \times_{\mathrm{id}_{TM}} \mathrm{id}_{TFY}) : TFY \to TFY$$

is called the F-prolongation of Q with respect to Λ.

For a projectable tangent valued k-form Q on Y, we construct a projectable tangent valued k-form $\mathcal{F}(Q, \Lambda)$ on FY in the same way.

We remark that ψ^F can be applied in an interesting way for constructing the F-prolongations of Lie algebroids and their actions, [12].

5.5 Remarks

If ω_Γ is the connection form of a connection Γ on Y, see 3.11, then $\mathcal{F}(\omega_\Gamma, \Lambda)$ is the connection form of the connection $\mathcal{F}(\Gamma, \Lambda)$.

However, finding the curvature of $\mathcal{F}(\Gamma, \Lambda)$ is a much more complicated problem than in 3.11. This can be illustrated on some special cases discussed in detail in [4].

In general, the F-prolongation of tangent valued forms with respect to Λ does not preserve the Frölicher-Nijenhuis bracket. Some very special cases, in which this bracket is preserved, are characterized in [4].

5.6 Weak principal bundles

First we recall the basic properties of r-th jet prolongations of associated bundles, which were clarified already by Ehresmann. Consider a fiber bundle $E = P[S, l]$ associated to a principal bundle $P(M, G)$. By 3.4, G^r_m acts on $T^r_m G$ by group isomorphisms. Hence we can construct the semidirect group product

$$W^r_m G = G^r_m \rtimes T^r_m G\,.$$

Then $W^r P := (P^r M \times_M J^r P) \to M$ is a principal bundle with structure group $W^r_m G$. (More details will be given in 5.7 in a more general setting.) $W^r P$ is called the r-th principal prolongation of principal bundle P, [14].

An action $l : G \times S \to S$ induces an action $W^r_m l : W^r_m G \times T^r_m S \to T^r_m S$,

$$W^r_m l\big((g, X), Z\big) = \big(T^r_m l(X, Z)\big) \circ g^{-1}\,,$$

$g \in G^r_m$, $X \in T^r_m G$, $Z \in T^r_m S$. Then $J^r E$ has a canonical structure of associated bundle

$$J^r E = W^r P[T^r_m S, W^r_m l]\,.$$

Some geometric properties of $J^r P$ can be described by using the following concept. Consider a group bundle $C \to M$.

Definition A fibered manifold $Q \to M$ is called a weak principal bundle with structure group bundle $C \to M$, if we are given a base preserving morphism $\varrho_Q : Q \times_M C \to Q$ such that each group C_x acts simply transitively on the right on Q_x.

The principal bundle is a weak principal bundle, the group bundle of which is a product $M \times K$.

By 4.7, we have $J^r(M \times G) = P^r M[T^r_m G]$. This is a group bundle of type $T^r_m G$. Applying J^r to $\varrho_P : P \times_M (M \times G) \to P$, we obtain

$$J^r \varrho_P : J^r P \times_M P^r M[T^r_m G] \to J^r P\,.$$

This defines a weak principal bundle structure on $J^r P$.

5.7 Principal F-prolongations

Consider the general case $F = (A, H, t)$. Applying F to ϱ_Q, we obtain

$$\varrho_{FQ} := F\varrho_Q : FQ \times_M FC \to FQ\,.$$

This endows $FQ \to M$ with the structure of a weak principal bundle with structure group bundle $FC \to M$.

The construction of $W^r P$ from 5.6 can be extended to the general case. By 3.4, the action H_G of G^r_m on $T^A G$ is by group homomorphisms. Hence we can construct the semidirect group product $W^A_H G = G^r_m \rtimes T^A G$ with the composition

$$(g_1, X_1)(g_2, X_2) = \big(g_1 \circ g_2, T^A\varphi\big(H_G(g_2^{-1})(X_1), X_2\big)\big)\,,$$

where φ is the group composition of G. For an action $l : G \times S \to S$, we define

$$W^A_H l : W^A_H G \times T^A S \to T^A S\,, \quad W^A_H l\big((g, X), Z\big) = H_S(g)\big(T^A l(X, Z)\big)\,,$$

$g \in G^r_m$, $X \in T^A G$, $Z \in T^A S$. This is an action, too. Clearly,

$$F(M \times G) = P^r M[T^A G, H_G]$$

is a group bundle of type $T^A G$. We have

$$F\varrho_P : FP \times_M P^r M[T^A G, H_G] \to FP\,.$$

We introduce $W^F P = P^r M \times_M FP$ and we define an action of $W^A_H G$ on $W^F P$ by

$$(u, Z)(g, X) = \big(u \circ g, F\varrho_P\big(Z, \{u \circ g, X\}\big)\big) \tag{28}$$

with $u \in P^r_x M$, $Z \in F_x P$, $g \in G^r_m$, $X \in T^A G$, so that $\{u \circ g, X\} \in P^r M[T^A G, H_G]$. One verifies directly

Proposition *$W^F P(M, W^A_H G)$ is a principal bundle. For an associated bundle $E = P[S, l]$, FE is an associated bundle $W^F P[T^A S, W^A_H l]$.*

We say that $W^F P$ is the principal F-prolongation of principal bundle P. In the case $F = J^r$, we have $W^{J^r} P = W^r P$.

Remark In [13], a more general construction of a principal bundle from a weak principal bundle and a suitable group bundle is discussed. Then a construction of the Lie algebroid of the principal bundle in question in terms of an action of a Lie algebroid on a Lie algebra bundle by derivations is described. In particular, the Lie algebroid of $W^F P$ can be determined in this way.

5.8 Lie groupoids

It is interesting that the description of FY as a subset of $P^r M[T^A Y]$, deduced in 4.7, is unavoidable for constructing the F-prolongations of Lie groupoids.

In the algebraic sense, a groupoid is a category in which all elements are invertible. We write a or b for the right or left unit map (also called source or target), respectively. A smooth groupoid $\Phi \overset{a}{\underset{b}{\rightrightarrows}} M$ is a groupoid such that Φ and M are manifolds, both $a, b : \Phi \to M$ are surjective submersions and the partial composition law

$$\varphi : \Phi^a \times_M \Phi^b \to \Phi\,, \quad \Phi^a := (\Phi \xrightarrow{a} M)\,, \quad \Phi^b := (\Phi \xrightarrow{b} M)$$

as well as the unit injection $e : M \to \Phi$ are smooth maps. The product groupoid is of the form

$$M \times G \times M\,,$$

where M is a manifold, G is a Lie group, $a = pr_3$, $b = pr_1$, $e(x) = (x, e_G, x)$, where e_G is the unit of G, and

$$\varphi\big((x_3, g_2, x_2)(x_2, g_1, x_1)\big) = (x_3, g_2 g_1, x_1)\,,$$

the product $g_2 g_1$ being in G. A smooth groupoid is called a Lie groupoid, if it is locally isomorphic to the product one.

For every Lie groupoid Φ and every $x \in M$,

$$\Phi_x := \{\theta \in \Phi, a\theta = x\}$$

is a principal bundle, whose structure group G_x is the isotropy group of Φ over x. Conversely, if $P(M, G)$ is a principal bundle, then the space of all equivalence classes

$$PP^{-1} = P \times P/\sim\,, \quad (v, u) \sim (vg, ug)\,, \quad u, v \in P,\; g \in G\,,$$

is a Lie groupoid over M with respect to the composition $\{w, v\}\{v, u\} = \{w, u\}$. We say that PP^{-1} is the gauge groupoid of P.

One verifies easily that for every Lie groupoid $\Phi \underset{b}{\overset{a}{\rightrightarrows}} M$, $T^A\Phi \underset{T^A b}{\overset{T^A a}{\rightrightarrows}} T^A M$ is also a Lie groupoid with partial composition law $T^A\varphi$ and unit injection $T^A e : T^A M \to T^A\Phi$. If $\Phi = PP^{-1}$, then $T^A\Phi = T^A P(T^A P)^{-1}$, where $T^A P \to T^A M$ is the principal bundle from 3.8.

However, in the case $F = (A, H, t)$ we cannot apply F to φ, for φ is not an $\mathcal{FM}_m$-morphism. But we can consider the F-prolongation of $\Phi \overset{a}{\to} M$. We write $\pi : F(\Phi^a) \to \Phi^a$ for the bundle projection and $\overline{a} = a \circ \pi$, $\overline{b} = b \circ \pi : F\Phi^a \to M$. Further consider the groupoid $\Pi^r M \underset{\beta}{\overset{\alpha}{\rightrightarrows}} M$ of all invertible r-jets of M into M. Hence $(\overline{a}, \overline{b}) : F\Phi^a \to M \times M$ and $(\alpha, \beta) : \Pi^r M \to M \times M$, so that we can construct the fiber product $\Pi^r M \times_{M\times M} F\Phi^a$.

Definition The F-prolongation of a Lie groupoid Φ is the subset

$$\mathcal{F}\Phi \subset \Pi^r M \times_{M\times M} F\Phi^a \tag{29}$$

of all pairs $(v \circ u^{-1}, \{u, Z\})$ satisfying $t_M v = T^A b(Z)$.

Hence the elements of $\mathcal{F}\Phi$ are the equivalence classes, with respect to the action of G^r_m,

$$\{v, Z, u\} \quad \text{satisfying} \quad T^A a(Z) = t_M u\,,\; T^A b(Z) = t_M v\,,$$

$u, v \in P^r M$, $Z \in T^A\Phi$. Given another $\{w, \overline{Z}, v\} \in \mathcal{F}\Phi$, we define the composition $*$ by

$$\{w, \overline{Z}, v\} * \{v, Z, u\} = \{w, T^A\varphi(\overline{Z}, Z), u\}\,. \tag{30}$$

Write $\widetilde{a}, \widetilde{b} : \mathcal{F}\Phi \to M$ for the projections determined by (29). Define $\widetilde{e} : M \to \mathcal{F}\Phi$ by $\widetilde{e}(x) = \{u, j^A\widehat{e(x)}, u\}$, where $\widehat{\ }$ denotes the constant map of $\mathbb{R}^k$ into $e(x)$. Then one verifies directly

Proposition $\mathcal{F}\Phi \overset{\widetilde{a}}{\underset{\widetilde{b}}{\rightrightarrows}} M$ *with the partial composition law* (30) *and the unit injection* $\widetilde{e}$ *is a Lie groupoid over* M.

For a principal bundle P, we have $\mathcal{F}(PP^{-1}) = (W^F P)(W^F P)^{-1}$, where $W^F P$ is the principal bundle from 5.7.

5.9 F-prolongation of actions

An action of a Lie groupoid Φ on a fibered manifold $p : Y \to M$ is a map $\psi : \Phi^a \times_M Y \to Y$ such that

$$p\big(\psi(\theta, y)\big) = b(\theta)\,,\ \psi\big(\varphi(\theta_2, \theta_1), y\big) = \psi\big(\theta_2, \psi(\theta_1, y)\big)\,,\ \psi\big(e(x), y\big) = y\,.$$

Every left group action $l : G \times S \to S$ induces an action of $M \times G \times M$ on the product fibered manifold $M \times S \to M$ by

$$\psi\big((x_2, g, x_1), (x_1, s)\big) = \big(x_2, l(g, s)\big)\,.$$

In the principal bundle form, $M \times S$ is a fiber bundle associated to $M \times G$.

If $\psi : \Phi^a \times_M Y \to Y$ is an action of Φ on Y, then $T^A\psi : (T^A\Phi)^{T^A a} \times_{T^A M} T^A Y \to T^A Y$ is an action of $T^A\Phi$ on $T^A Y \to T^A M$. In the case $F = (A, H, t)$, we have $FY \subset P^r M[T^A Y]$ and we define $\mathcal{F}\psi : (F\Phi)^{\widetilde{a}} \times_M FY \to FY$ by

$$\mathcal{F}\psi\big(\{v, Z, u\}, \{u, Q\}\big) = \big\{v, T^A\psi(Z, Q)\big\}\,,$$

$u, v \in P^r M$, $Z \in T^A\Phi$, $Q \in T^A Y$.

Analogously to 5.8, one verifies directly

Proposition $\mathcal{F}\psi$ *is an action of* $\mathcal{F}\Phi$ *on* $FY \to M$.

In the case $\Phi = PP^{-1}$, we obtain the associated bundle structure on FY described in 5.7.

5.10 Remark

The theory of Weil bundles can be extended to some infinite dimensional spaces. The basic information can be found in [18]. We describe the case of a rather simple functional bundle and we present one more advanced result.

Two fibered manifolds $p_1 : Y_1 \to M$, $p_2 : Y_2 \to M$ over the same base define an infinite dimensional bundle

$$\bigcup_{x \in M} C^\infty(Y_{1x}, Y_{2x}) =: \mathcal{F}(Y_1, Y_2) \xrightarrow{\ \pi\ } M$$

of all smooth maps between the individual fibers over the same base point. This is a smooth space in the sense of Frölicher. We describe the basic ideas of this theory only. Given a manifold Q, a map $f : Q \to \mathcal{F}(Y_1, Y_2)$ is said to be smooth in the sense of Frölicher, if $\underline{f} = \pi \circ f : Q \to M$ is smooth, so that we can construct the pullback

$$\underline{f}^* Y_1 = \big\{(q, y) \in Q \times Y_1, \pi\big(f(q)\big) = p_1(y)\big\}\,,$$

and the associated map

$$\widetilde{f} : \underline{f}^*Y_1 \to Y_2\,, \quad \widetilde{f}(q,y) = f(q)(y)$$

is also smooth.

Write $T^A_X Y_i = (T^A p_i)^{-1}(X)$, $X \in T^A M$, $i = 1, 2$. Consider two maps $f, g : \mathbb{R}^k \to \mathcal{F}(Y_1, Y_2)$ that are smooth in the sense of Frölicher and satisfy $j^A(\pi \circ f) = j^A(\pi \circ g) = X$. Then we can construct a map $\mathbf{j}^A f : T^A_X Y_1 \to T^A_X Y_2$,

$$(\mathbf{j}^A f)\big(j^A\big(\gamma(u)\big)\big) = j^A\big(f(u)\big(\gamma(u)\big)\big)\,, \quad u \in \mathbb{R}^k\,,$$

and the same for g. If the maps $\mathbf{j}^A f$ and $\mathbf{j}^A g$ coincide, we say that f and g determine the same A-velocity on $\mathcal{F}(Y_1, Y_2)$. This defines an infinite dimensional bundle $T^A\mathcal{F}(Y_1, Y_2) \to T^A M$. Since each algebra homomorphism $\mu : A \to B$ can be interpreted as a reparametrization, we have an induced map

$$\mu_{\mathcal{F}(Y_1,Y_2)} : T^A\mathcal{F}(Y_1, Y_2) \to T^B\mathcal{F}(Y_1, Y_2)\,.$$

There are no flows in this situation. However, if $X : \mathcal{F}(Y_1, Y_2) \to T\mathcal{F}(Y_1, Y_2)$ is a vector field that is differentiable in the sense of Frölicher, then the formula

$$\mathcal{T}^A X := \varkappa^A_{\mathcal{F}(Y_1,Y_2)} \circ T^A X$$

defines a vector field on $T^A\mathcal{F}(Y_1, Y_2)$. The proof of the fact that this operation preserves the bracket of vector fields is heavily based on the Weil algebra technique, [1].

Further, if F is a f.p.p.b. functor on $\mathcal{FM}_m$, then the decomposition $F = (A, H, t)$ enables us to construct an infinite dimensional bundle $F\mathcal{F}(Y_1, Y_2) \to M$ analogously to 4.7, [1]. The r-jet prolongation $J^r\mathcal{F}(Y_1, Y_2)$ can be also interpreted in a direct geometric way. A section $s : M \to \mathcal{F}(Y_1, Y_2)$ smooth in the sense of Frölicher is identified with a base-preserving morphism $Y_1 \to Y_2$. Then $J^r\mathcal{F}(Y_1, Y_2)$ is the space of all fiber r-jets $\mathbf{j}^r_x s$, $x \in M$, introduced in [14].

Bibliography

[1] A. Cabras, J. Janyška and I. Kolář: Functorial prolongations of some functional bundles *Annales Acad. Paed. Cracoviensis, Folia 23, Studia Mahematica IV* (2003) 17–30

[2] A. Cabras and I. Kolář: Prolongation of tangent valued forms to Weil bundles *Arch. Math. (Brno)* **31** (1995) 139–145

[3] A. Cabras and I. Kolář: On the covariant exterior differentiation on Weil bundles *In: Differential Geometry and Its Applications* (Proceedings, Silesian University at Opava, 2002) 63–72

[4] A. Cabras and I. Kolář: Flow prolongation of some tangent valued forms *Czechoslovak Math. J* (to appear)

[5] M. Doupovec and I. Kolář: On the jets of fibered manifold morphisms *Cahiers Topo. Géom. Differ. Catégoriques* **XL** (1999) 21–30

[6] M. Doupovec and I. Kolář: Iteration of fiber product preserving bundle functors *Monatsh. Math.* **134** (2001) 39–50

[7] J. Gancarzewicz, M. Mikulski and Z. Pogoda: Lifts of some tensor fields and connections to product preserving bundles *Nagoya Math. J.* **135** (1994) 1–41

[8] A. Kock: *Synthetic Differential Geometry* (London Math. Soc. Lecture Notes Series 51, Cambridge University Press, 1981)

[9] I. Kolář: Affine structure on Weil bundles *Nagoya Math. J.* **158** (2000) 99–106

[10] I. Kolář: On the geometry of fiber product preserving bundle functors *In: Differential Geometry and Its Applications* (Proceedings, Silesian University at Opava, 2002) 85–92

[11] I. Kolář: A general point of view to nonholonomic jet bundles *Cahiers Topo. Géom. Differ. Catégoriques* **XLIV** (2003) 149–160

[12] I. Kolář: Functorial prolongations of Lie algebroids *In: Proceedings Conf. Prague 2004* (Charles University Prague, 2005) 305–314

[13] I. Kolář and A. Cabras: On the functorial prolongations of principal bundles *Comment. Math. Univ. Carolin* **47** (2006) 719–731

[14] I. Kolář, P. W. Michor and J. Slovák: *Natural Operations in Differential Geometry* (Springer-Verlag, 1993)

[15] I. Kolář and W. M. Mikulski: On the fiber product preserving bundle functors *Differential Geometry and Its Applications* **11** (1999) 105–115

[16] I. Kolář and W. M. Mikulski: Fiber product preserving bundle functors on all morphisms of fibered manifolds *Arch. Math. (Brno)* **42** (2006) 285–293

[17] I. Kolář and M. Modugno: Torsions of connections on some natural bundles *Differential Geometry and Its Applications* **2** (1992) 1–16

[18] A. Kriegl and P. W. Michor: *The Convenient Setting of Global Analysis* (AMS, Mathematical Surveys and Monographs, Vol. 53, 1997)

[19] M. Kureš and W. M. Mikulski: Natural operators lifting vector fields to bundles of Weil contact elements *Czechoslovak Math. J.* **54** (2004) 855–867

[20] W. M. Mikulski: Product preserving bundle functors on fibered manifolds *Arch. Math. (Brno)* **32** (1996) 307–316

[21] A. Morimoto: Prolongations of connections to bundles of infinitely near points *J. Diff. Geometry* **11** (1976) 479–498

[22] J. Muñoz, J. Rodriguez and F. J. Muriel: Weil bundles and jet spaces *Czechoslovak Math. J.* **50** (2000) 721–748

[23] A. Weil: Théorie des points proches sur les variétes différentielles *In: Colloque de topol. et géom. diff.* (Strasbourg, 1953) 111–117

IVAN KOLÁŘ

Department of Algebra and Geometry, Masaryk University,
Janáčkovo nám. 2a, CZ 602 00 Brno, Czech Republic

E-mail: kolar@math.muni.cz

Handbook of Global Analysis
Demeter Krupka and David Saunders

Distributions, vector distributions, and immersions of manifolds in Euclidean spaces[1]

Július Korbaš

Contents

1 Introduction

The term *distribution* has two different usages in global analysis. In one of them (see Sections 2 and 3), distributions are sometimes also called *Schwartzian distributions*; a popular example is Dirac's delta function.

In the other usage (see Section 4), a k-dimensional smooth distribution on a smooth manifold M is a smooth assignment of a k-dimensional subspace of the tangent vector space $T(M)_p$ to each point $p \in M$; "smooth" means of class C^∞ here and elsewhere in this text. A distribution in the latter sense defines a smooth k-dimensional subbundle of the tangent bundle $T(M)$, and vice versa. We shall call it – to make a clear distinction – a k-dimensional *vector distribution*. We shall concentrate on general (mainly existence) results on vector distributions and on the following question, known as the *vector field problem*: When does the tangent bundle of a smooth finite-dimensional manifold admit a trivial subbundle of a given dimension? There are also interesting and important results on *completely integrable distributions*, i.e., on *foliations*; but we restrict ourselves to just a passing mention of them: they are covered in a recent handbook-survey by R. Barre and A. El Kacimi Alaoui [15].

Finally, a generalized version of the above question reads: When does a given vector bundle over a smooth manifold admit a trivial subbundle of a given dimension? Hirsch

[1] The author was supported in part by two grants of VEGA (Slovakia).

and Smale's theory provides a bridge between this question and *immersions of manifolds* in Euclidean spaces (see Section 5).

However different the three topics announced in the title seem, they have much in common. Indeed, on the one hand, *Schwartzian distributions* play an important rôle in Atiyah and Singer's proof, published in [13], of their index formula (for the formula, see D. Bleecker's contribution in this Handbook). On the other hand, as we shall see in Sections 4 and 5, the index formula itself implies results on *vector distributions* and *immersions of manifolds* in Euclidean spaces. So the Atiyah-Singer index formula is a kind of node at which the three topics mentioned in the title are joined. In Section 5, we shall also show some interplay between vector distributions, immersions, and the Lyusternik-Shnirel'man category.

We adopt the following **convention**: Unless specified otherwise, we shall use the word "manifold" to mean a smooth, Hausdorff, paracompact, connected, finite-dimensional manifold without boundary, equipped with a smooth Riemannian metric whenever needed. In particular, if such a manifold is compact, we call it *closed*. Maps between manifolds will be smooth, all maps between topological spaces will be continuous. We do not distinguish between diffeomorphic manifolds. Finally, $\mathbb{C}$ denotes the complex numbers, $\mathbb{R}$ the reals, $\mathbb{Q}$ the rational numbers, $\mathbb{Z}$ the integers, and $\mathbb{Z}_p$ the integers modulo the prime number p.

We remark that some of the concepts and questions considered here in the smooth context also have their C^r-analogues for $r \neq \infty$ (but their behaviour may or may not be analogous!).

By no means is this text comprehensive. We apologize in advance for omission of many names, considerable ideas and important contributions.

We thank Martin Čadek, Michael Kunzinger, Tibor Macko, Jim Stasheff, and (last but not least) Peter Zvengrowski for their comments.

2 Distributions on Euclidean spaces

2.1 Preliminaries

Unless stated differently, U denotes an open subset of $\mathbb{R}^n$. If $1 \leq p < \infty$, then $\mathcal{L}^p(U)$ (briefly $\mathcal{L}^p$) denotes the standard Lebesgue space of (equivalence classes of) measurable functions f (on U) such that $|f|^p$ is integrable; $\mathcal{L}^\infty(U)$ is the vector space of (equivalence classes of) functions measurable and essentially bounded on U. Let $\mathcal{E}(U)$ be the set of all smooth functions $U \to \mathbb{C}$. With the usual point-wise addition and point-wise multiplication by scalars, $\mathcal{E}(U)$ is a vector space. With a suitable topology (usually defined by a family of semi-norms), $\mathcal{E}(U)$ becomes a Fréchet space (i.e., a complete, metrizable, locally convex topological vector space).

For non-negative integers $\alpha_1, \dots, \alpha_n$, let $\alpha = (\alpha_1, \dots, \alpha_n)$ denote a multi-index of order $|\alpha| = \alpha_1 + \cdots + \alpha_n$. Let $D_x^\alpha = D_1^{\alpha_1} \dots D_n^{\alpha_n}$ be differentiation operators, where $D_j^{\alpha_j} = (\frac{1}{i}\frac{\partial}{\partial x_j})^{\alpha_j}$. The factor $\frac{1}{i}$ (where $i \in \mathbb{C}$ is the imaginary unit) is included just for convenience, to simplify some formulae (e.g., for Fourier transformations). We note that $D_x^{(0,\dots,0)} f(x) = f(x)$. We define a subset $\mathcal{D}(U) \subset \mathcal{E}(U)$ by

$$\mathcal{D}(U) = \{f \in \mathcal{E}(U); \operatorname{supp}(f) \text{ is compact}\},$$

where for any function g on U, $\operatorname{supp}(g)$, the *support* of g, is the closure in U of $\{x \in U; g(x) \neq 0\}$. With a suitable topology (which is not the subspace topology), $\mathcal{D}(U)$

becomes a locally convex nonmetrizable space; but for any compact $K \subset U$, the set $\mathcal{D}(K)$, consisting of elements in $\mathcal{D}(U)$ with support contained in K, is a Fréchet space. In $\mathcal{D}(U)$, a sequence $\{h_i\}$ converges to the zero-function 0 (we write $\lim_{i\to\infty} h_i = 0$) if there exists some compact $K_0 \subset U$ such that $\mathrm{supp}(h_i) \subset K_0$ for every i, and for each α, the sequence $\{D^\alpha h_i\}$ converges uniformly to 0.

In general, if V is a (complex) locally convex topological vector space and V' its topological dual, we denote by $\langle\ ,\ \rangle : V' \times V \to \mathbb{C}$ the bilinear dual pairing. The vector space V' can be equipped with several topologies, but for our purposes it suffices to consider two of them: the (weak) w^*-topology or the (strong) s^*-topology (see, e.g., Chap. 4 in K. Yosida, Functional analysis, Springer 1965). In the sequel, topological duals of locally convex topological vector spaces will be mostly (tacitly) taken with the w^*-topology; in many cases, the w^*-topology and the s^*-topology are equivalent, but in some cases, it is important to take the s^*-topology.

2.2 The definition of distributions, examples, and historical remarks

Distributions, in the sense of Schwartz, or *Schwartzian distributions*, on U are defined to be elements of the topological dual vector space $(\mathcal{D}(U))'$, briefly denoted by $\mathcal{D}'(U)$. The elements of $\mathcal{D}(U)$ are then called *test functions* for the distributions from $\mathcal{D}'(U)$. Sometimes, the elements of $\mathcal{D}'(U)$ are referred to as *distributions over the test space* $\mathcal{D}(U)$. If we do not wish to emphasize U, we write just $\mathcal{D}$ instead of $\mathcal{D}(U)$, and $\mathcal{D}'$ instead of $\mathcal{D}'(U)$. More precisely, we defined here *complex* distributions. Real distributions will not be considered here.

For a distribution $s \in \mathcal{D}'$, its value at $f \in \mathcal{D}$ will be written, using the dual pairing, as $\langle s, f\rangle$; but sometimes we may write it simply $s(f)$. A linear functional $s : \mathcal{D}(U) \to \mathbb{C}$ is a distribution on U if it is continuous, that is, if $\lim_{i\to\infty} h_i = 0$ implies $\lim_{i\to\infty} s(h_i) = 0$.

To give examples, denote by $\mathcal{L}^p_{\mathrm{loc}}(U)$ (briefly $\mathcal{L}^p_{\mathrm{loc}}$) the space of locally $\mathcal{L}_p$-functions on U, i.e., of functions f such that $\varphi f \in \mathcal{L}^p(U)$ for each $\varphi \in \mathcal{D}(U)$. Let $f \in \mathcal{L}^1_{\mathrm{loc}}(U)$. Then we associate with f a distribution T_f, defined by

$$\langle T_f, u\rangle = \int_U f(x)u(x)dx,$$

for all $u \in \mathcal{D}(U)$. It can be proved that $T_f = T_g$ if and only if f and g coincide almost everywhere. Therefore the distribution T_f is frequently identified with the function f (one often writes just f instead of T_f), and in this sense, distributions are generalizations of (locally integrable) functions.

The need for generalizations of functions (and their derivatives) had arisen since the end of the 1920's. For instance, P. Dirac and other physicists, in their works on quantum mechanics, were successfully using (see [112]) a "function"(later known as the *Dirac delta-function*) which was mathematically impossible. Indeed, they required, e.g., the following properties: $\delta(x) = 0$ for $x \neq 0$, $\delta(0) = \infty$, and $\int_{-\infty}^{\infty} \delta(x)dx = 1$. In addition to this, in the framework of Heaviside's symbolic calculus, invented in the theory of electrical circuits, the *Heaviside function* $y : \mathbb{R} \to \mathbb{R}$, defined by $y(x) = 0$ for $x \leq 0$, $y(x) = 1$ for $x > 0$, should have had the Dirac delta-function as its first derivative: $y'(x) = \delta(x)$. But there were also purely mathematical works (mainly on partial differential equations) for which the standard notions of functions and derivatives appeared to be too narrow. So it was, for instance, about 1932, with the finite parts of divergent integrals, applied by J.

Hadamard in his works on the fundamental solutions of the wave equation. From others who encountered similar problems with functions (and also tried to solve them in some way) in the 1930's - 1940's, we mention S. Bochner, T. Carleman, A. Beurling, J. Leray, K. Friedrichs, C. Morrey, and – last but not least – S. L. Sobolev. The latter, in 1936 (in Mat. Sb. 1), introduced and successfully applied *generalized* derivatives and *generalized* solutions of differential equations in his study of the Cauchy problem for hyperbolic equations.

A short time after World War II, L. Schwartz ([112], [113]) started to present a new, systematic basis for the whole variety of the generalized functions that had appeared in the meantime. The definition of distributions, test functions, and test spaces above is due to him. It is interesting that his idea was in a sense anticipated by A. Weil in his book L'intégration dans les groupes topologiques et ses applications, Hermann 1940. Indeed, in Weil's approach to integration on locally compact groups, Radon measures on a group G are considered as continuous linear functionals on the space of those continuous functions on G vanishing on the complement of some compact subset.

Thanks to Schwartz's theory of distributions, many old problems or discrepancies disappeared. So, e.g., the problematic "Dirac's function" was converted into the rigorous *Dirac's distribution* δ on $\mathbb{R}$ defined by

$$\langle \delta, f \rangle = f(0)$$

for every $f \in \mathcal{D}(\mathbb{R})$. The same is also an example of a "pure" distribution, because it is not (associated with) a function. At the same time (see, e.g., [59, 4.1.5]), for any distribution $s \in \mathcal{D}'(U)$ there is a sequence of smooth functions on U converging to s (with the w^*-topology on $\mathcal{D}'(U)$).

Schwartz's distributions turned out to be suitable for describing *distributions* of various material quantities. This explains why Schwartz, perhaps also influenced by the fact that, in the beginning, physicists were more accepting of his ideas than mathematicians, preferred the term *distributions* (although some other mathematicians have continued calling them "generalized functions"). In the subsequent years, the theory of distributions was further developed in many directions. It has found many applications in differential equations ([59]), pseudodifferential operators ([13]), representations of Lie groups ([35]), in mathematical physics (see, e.g., books by Y. Choquet-Bruhat, C. DeWitt-Morette, and M. Dillard-Bleick or by V. S. Vladimirov), but also in engineering. In the 1950's, G. de Rham ([106]) inspired by Schwartz's approach, developed his theory of *currents* on manifolds, and used it to prove that the de Rham cohomology groups and the singular cohomology groups (for manifolds) are isomorphic. The notion of currents includes, as special cases, both the notion of differentiable forms and the notion of singular chains. Schwartz's distributions can also be defined on manifolds; they are in fact a special class of de Rham's currents, or still from another point of view, a special class of *generalized sections of vector bundles* (see 3.2).

There are also other important types of generalizations of the notion of function, e.g., *Sato's hyperfunctions* (see M. Morimoto, An introduction to Sato's hyperfunctions, American Mathematical Society (1993)). The following subsections, until the end of Section 2, will be devoted to some of the basic properties of Schwartzian distributions on open subsets of Euclidean space, as a kind of preparation for passing to distributions on manifolds. A detailed and far-reaching exposition of the theory of distributions on Euclidean spaces can be found, e.g., in L. Hörmander's book [59] (having also a chapter on hyperfunctions); a recent brief source is [99].

2.3 Local properties of distributions

Let U' be an open subset of an open subset $U \subset \mathbb{R}^n$. For every $f \in \mathcal{D}(U')$, we have its obvious smooth extension $\tilde{f} \in \mathcal{D}(U)$: $\tilde{f}(x) = f(x)$ if $x \in U'$, and $\tilde{f}(x) = 0$ if $x \in U \setminus U'$. Then for every distribution $s \in \mathcal{D}'(U)$ one defines its restriction $s_{U'}$ to U' by putting $\langle s_{U'}, f\rangle = \langle s, \tilde{f}\rangle$, for every $f \in \mathcal{D}(U')$. Two distributions $s,\, t \in \mathcal{D}'(U)$ are defined to be equal on U' if $s_{U'} = t_{U'}$. Then $s,\, t \in \mathcal{D}'(U)$ are equal if each point in U has a neighbourhood, where s and t are equal.

The following is sometimes called the *gluing principle* for distributions. Given an open covering $\{U_i; i \in I\}$ (where I is some set) of U and a set of distributions $\{s_i \in \mathcal{D}'(U_i); i \in I\}$ such that s_i and s_j coincide in $U_i \cap U_j$ for any $i,\, j \in I$, then there is a unique distribution $s \in \mathcal{D}'(U)$ such that s coincides with s_i on U_i for each i. This implies that for each distribution $s \in \mathcal{D}'(U)$ there exists an open (in U) subset $U_s \subset U$ which is largest, with respect to set inclusion, with the property that s vanishes on it (in the sense that $\langle s, f\rangle = 0$ for each function $f \in \mathcal{D}(U_s)$). The complement, $U \setminus U_s$, is defined to be the *support of the distribution* s; we denote it by $\mathrm{supp}(s)$. For instance, we have $\mathrm{supp}(\delta) = \{0\}$. For any distribution T_f associated with a continuous function f, $\mathrm{supp}(T_f)$ is the same as $\mathrm{supp}(f)$.

Similarly, one can define the singular support of a distribution $s \in \mathcal{D}'(U)$. Let V be an open subset of U. Let V_0 be the largest open subset in U such that s_{V_0} is a smooth function on V_0. Then the complement $U \setminus V_0$ is defined to be the *singular support* of s; we denote it by $\mathrm{singsupp}(s)$. For instance, we have $\mathrm{singsupp}(\delta) = \{0\}$.

2.4 Spaces of distributions pertinent to various test spaces

For a given problem (e.g., on partial differential equations), it may be necessary to work with a proper subset of the distributions from $\mathcal{D}'(U)$. Indeed, it might be better to choose a test space different from $\mathcal{D}(U)$, hence to consider the topological dual to some other topological vector space.

So for instance, if we extend the test space from $\mathcal{D}(U)$ to $\mathcal{E}(U)$, then the topological dual, $(\mathcal{E}(U))'$, which we denote by $\mathcal{E}'(U)$, may be considered as a set of distributions. Indeed ([59, 2.3.1]), $\mathcal{E}'(U)$ can be identified with the set of all distributions in $\mathcal{D}'(U)$ with compact support.

To mention another important example, let $\mathcal{S}(\mathbb{R}^n)$ be the Schwartz space of (rapidly decreasing) smooth functions $f : \mathbb{R}^n \to \mathbb{C}$ such that $p_{\alpha,\beta}(f) = \sup_{x\in\mathbb{R}^n} |x^\beta D^\alpha f(x)| < \infty$ for all multi-indices α and β, where $x^\beta = x_1^{\beta_1} \cdots x_n^{\beta_n}$. The semi-norms $p_{\alpha,\beta}$ define a topology on $\mathcal{S}(\mathbb{R}^n)$, converting it into a Fréchet space. The space $\mathcal{S}(\mathbb{R}^n)$ can be continuously injected into $\mathcal{L}^p(\mathbb{R}^n)$ for all p; we shall consider it as a subset of $\mathcal{L}^p(\mathbb{R}^n)$ (note that it is dense if $1 \le p < \infty$).

The topological dual, $(\mathcal{S}(\mathbb{R}^n))'$, will be denoted by $\mathcal{S}'(\mathbb{R}^n)$. The space $\mathcal{S}'(\mathbb{R}^n)$ can be interpreted as a subspace of $\mathcal{D}'(\mathbb{R}^n)$; its elements are known as *slowly increasing* or *tempered distributions*. One has (see [59, p. 164]) $\mathcal{L}^p(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n)$ for all p. We can write

$$\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset \mathcal{E}(\mathbb{R}^n), \ \ \mathcal{E}'(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n).$$

2.5 Derivatives of distributions

If f is a smooth function on U, then (using integration by parts) we obtain $\langle T_{D^\alpha f}, u\rangle = (-1)^{|\alpha|}\langle T_f, D^\alpha u\rangle$ for any $u \in \mathcal{D}(U)$. Motivated by this, for any $s \in \mathcal{D}'(U)$, we define $D^\alpha s : \mathcal{D}(U) \to \mathbb{C}$ by

$$(D^\alpha s)(u) = (-1)^{|\alpha|}\langle s, D^\alpha u\rangle,$$

$u \in \mathcal{D}(U)$. It is readily seen that $D^\alpha s \in \mathcal{D}'(U)$; $D^\alpha s$ is called the *derivative* (of order $|\alpha|$) *of the distribution* s. The map $D^\alpha : \mathcal{D}'(U) \to \mathcal{D}'(U)$ is continuous.

On the one hand, using this definition, one can calculate any derivatives of those locally integrable functions (interpreted as distributions) which were not differentiable in the standard sense. On the other hand, the distributional derivatives of smooth functions can be seen to coincide with the usual partial derivatives. Thanks to the definition given above, *each* distribution can be considered infinitely differentiable (in the generalized sense). If f is a locally integrable function (on some U) and if, for some α, the distributional derivative $D^\alpha f$ coincides with T_g (see 2.2) for some locally integrable function g on U, then $D^\alpha f$ is a *generalized derivative of function type*. For example, for the Heaviside function y (see 2.1) we immediately obtain $D^{(1)}y = \delta$, hence a rigorous form of what Dirac and other physicists were using in Heaviside's symbolic calculus in the 1920's. At the same time, since the Dirac distribution $\delta \in \mathcal{D}'(\mathbb{R})$ is not associated with a function (see 2.2), $D^{(1)}y$ is not a generalized derivative of function type. The space $\mathcal{S}'(\mathbb{R}^n)$ is closed with respect to differentiation ([59, p. 164]).

2.6 Products and Colombeau's algebras

There is no problem with defining the product $fs = sf$ for $s \in \mathcal{D}'(U)$ and f smooth on U: one puts $(fs)(u) = \langle s, fu\rangle$, $u \in \mathcal{D}(U)$. Then $fs \in \mathcal{D}'(U)$. Additionally ([59, p. 164]), $\mathcal{S}'(\mathbb{R}^n)$ is closed with respect to products with polynomials or functions from $\mathcal{S}(\mathbb{R}^n)$. However, as shown by Schwartz [114], there are difficulties with extending this product, if it should have reasonable properties (e.g., associativity), to *arbitrary* distributions. For products of distributions under certain restrictions see, e.g., [59, 8.2], [97].

A radical step was undertaken by J.-F. Colombeau in the early 1980's. To give a meaning to *any* finite product of distributions, he constructed a commutative and associative algebra of "new generalized functions", for any open subset $U \subset \mathbb{R}^n$. There are various versions of the original construction, so that one may speak about various Colombeau's algebras (see [25]). The *simplified* or *special Colombeau algebra* $\mathcal{G}(\mathbb{R}^n)$ presented in [25] can be defined in simple terms. Indeed, let $\mathcal{M}(\mathbb{R}^n)$ ($\mathcal{M}$ comes from "moderate") be the algebra of all nets $(f_\varepsilon)_{\varepsilon>0}$ of smooth functions $\mathbb{R}^n \to \mathbb{C}$ such that for every compact $K \subset \mathbb{R}^n$ and for any α there exist a non-negative integer N, a real $\kappa > 0$, and a real $c > 0$ such that

$$\sup_{x\in K} |D^\alpha_x f_\varepsilon(x)| \leq c \cdot \varepsilon^{-N}$$

whenever $0 < \varepsilon < \kappa$. Let $\mathcal{N}(\mathbb{R}^n)$ ($\mathcal{N}$ comes from "negligible") be the ideal in $\mathcal{M}(\mathbb{R}^n)$ consisting of $(f_\varepsilon)_{\varepsilon>0}$ such that for every compact $K \subset \mathbb{R}^n$, for any α, and for any non-negative integer N there exist $\kappa > 0$ and $c > 0$ such that

$$\sup_{x\in K} |D^\alpha_x f_\varepsilon(x)| \leq c \cdot \varepsilon^N$$

whenever $0 < \varepsilon < \kappa$. Then the Colombeau algebra $\mathcal{G}(\mathbb{R}^n)$ is the quotient $\mathcal{G}(\mathbb{R}^n) = \mathcal{M}(\mathbb{R}^n)/\mathcal{N}(\mathbb{R}^n)$.

For $\mathcal{G}(\mathbb{R}^n)$ it is true that every element of $\mathcal{G}(\mathbb{R}^n)$ admits partial differentiation to any order, the set of Schwartzian distributions on $\mathbb{R}^n$, $\mathcal{D}'(\mathbb{R}^n)$, is contained (linearly embedded) in $\mathcal{G}(\mathbb{R}^n)$, the algebra of smooth functions, $\mathcal{E}(\mathbb{R}^n)$, is a subalgebra of $\mathcal{G}(\mathbb{R}^n)$, and every function in $\mathcal{G}(\mathbb{R}^n)$ admits values at points (but these are *generalized numbers*). In addition to this, $\mathcal{G}(\mathbb{R}^n)$ is an algebra which is closed under partial differentiation. In an analogous way, one can describe the Colombeau algebra $\mathcal{G}(U)$ for any open subset $U \subset \mathbb{R}^n$. As applications, Colombeau used his construction for explaining some heuristic formulae from physics (dealing with the Hamiltonian and Lagrangian densities). He also has shown that in $\mathcal{G}(U)$, partial differential equations may have new solutions.

Various modifications and applications (in particular, to nonlinear problems) of Colombeau's theory have also been studied, or other nonlinear theories have been suggested, by several authors (J. Aragona, V. Boie, Yu. V. Egorov, M. Kunzinger, M. Oberguggenberger, S. Pilipović, E. Rosinger, V. M. Shelkovich, R. Steinbauer, to name at least some). Now we come back to the framework of the linear Schwartzian distribution theory (later, in the final part of 3.2, we again shall include a remark on Colombeau's theory).

2.7 Behaviour of distributions under diffeomorphisms

Let U_x and V_y be open subsets of $\mathbb{R}^n$; $x = (x_1, \ldots, x_n)$ is a generic variable point in U_x, and $y = (y_1, \ldots, y_n)$ is a generic variable point in V_y. Let $\varphi : U_x \to V_y$ be a diffeomorphism. So we have a (smooth) coordinate change,

$$x_1 = (\varphi^{-1})_1(y_1, \ldots, y_n), \ldots, x_n = (\varphi^{-1})_n(y_1, \ldots, y_n).$$

We denote by $J(\varphi^{-1})$ the determinant of the Jacobian matrix $d(\varphi^{-1})$. For $g \in \mathcal{L}^1_{\text{loc}}(V_y)$, we define $\varphi^*(g)$ on U_x by

$$\varphi^*(g)(x) = g \circ \varphi(x).$$

At the same time, for $f \in \mathcal{L}^1_{\text{loc}}(U_x)$ we define $\varphi_\#(f)$ on V_y by

$$\varphi_\#(f)(y) = |J(\varphi^{-1})(y)|(\varphi^{-1})^*(f)(y).$$

One readily sees that $\varphi^* : \mathcal{L}^1_{\text{loc}}(V_y) \to \mathcal{L}^1_{\text{loc}}(U_x)$ is a linear isomorphism, and it restricts to linear isomorphisms $\varphi^* : \mathcal{E}(V_y) \to \mathcal{E}(U_x)$, $\varphi^* : \mathcal{D}(V_y) \to \mathcal{D}(U_x)$. Similarly, $\varphi_\# : \mathcal{L}^1_{\text{loc}}(U_x) \to \mathcal{L}^1_{\text{loc}}(V_y)$ is a linear isomorphism, and it restricts to linear isomorphisms $\varphi_\# : \mathcal{E}(U_x) \to \mathcal{E}(V_y)$, $\varphi_\# : \mathcal{D}(U_x) \to \mathcal{D}(V_y)$.

The operator $\varphi_\#$ acts "covariantly" on standardly interpreted functions, while the operator φ^* acts "contravariantly" on functions usually interpreted as distributions (hence on objects dual to standard functions). Indeed, for any $g \in \mathcal{L}^1_{\text{loc}}(V_y)$ we have the associated distribution $T_g \in \mathcal{D}'(V_y)$. So we obtain (recall, from 2.2, that we may identify functions and the associated distributions) that

$$\langle \varphi^*(T_g), f \rangle = \langle T_g, \varphi_\#(f) \rangle$$

for all $f \in \mathcal{D}(U_x)$. It is then natural to use this as a pattern for defining

$$\varphi^* : \mathcal{D}'(V_y) \to \mathcal{D}'(U_x), \ \ \langle \varphi^*(s), f \rangle = \langle s, \varphi_\#(f) \rangle,$$

for all $f \in \mathcal{D}(U_x)$; this map is continuous (in the w^*-topology). The distribution $\varphi^*(s)$ is sometimes called the *pullback of the distribution* s by φ. Of course, if we have diffeomorphisms $\varphi : U_x \to V_y$ and $\psi : V_y \to W_z$, then $(\psi \circ \varphi)^* = \varphi^* \circ \psi^*$.

2.8 The Schwartz kernel theorem

Let $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ be open sets; now we equip $\mathcal{D}'(V)$ with the s^*-topology. Let $L(\mathcal{D}(U), \mathcal{D}'(V))$ denote the space of continuous linear operators $\mathcal{D}(U) \to \mathcal{D}'(V)$, with the topology of uniform convergence on bounded sets. The following is known as the *Schwartz kernel theorem* (see [59, 5.2]): Every distribution $k \in \mathcal{D}'(V \times U)$ defines a continuous linear operator $\Lambda_k \in L(\mathcal{D}(U), \mathcal{D}'(V))$ by

$$\Lambda_k(u)(v) = \langle k, v \otimes u \rangle,$$

$u \in \mathcal{D}(U)$, $v \in \mathcal{D}(V)$, where $v \otimes u : V \times U \to \mathbb{C}$, $(v \otimes u)(y, x) = v(y)u(x)$. Conversely, every $\Lambda \in L(\mathcal{D}(U), \mathcal{D}'(V))$ can be expressed as Λ_k for some distribution $k \in \mathcal{D}'(V \times U)$. In addition to this, the map

$$\mathcal{D}'(V \times U) \to L(\mathcal{D}(U), \mathcal{D}'(V)), \ \ k \mapsto \Lambda_k,$$

is bijective.

The dual pairing $\langle \ , \ \rangle : \mathcal{D}'(V \times U) \times \mathcal{D}(V \times U) \to \mathbb{C}$ is sometimes written using the integral sign (see 3.1). Using this convention, one may write

$$\Lambda_k(u)(v) = \langle k, v \otimes u \rangle$$

as

$$\Lambda_k(u)(v) = \int k \cdot v \otimes u \ dxdy.$$

This converts – of course, in general just optically – Λ_k into an *integral operator*, and that is why the distribution $k \in \mathcal{D}'(V \times U)$ is called a *generalized* or *Schwartz kernel*, and why the above theorem is called the kernel theorem.

2.9 Fourier transforms of distributions

In 2.4, we introduced tempered distributions. They behave nicely with respect to the Fourier transform which is useful, e.g., for dealing with pseudodifferential operators (see Bleecker's contribution in this Handbook).

For $f \in \mathcal{L}^1(\mathbb{R}^n_x)$, the *Fourier transform* is defined by

$$\mathcal{F}f(\xi) = \int e^{-ix\cdot\xi} f(x)dx,$$

where $x = (x_1, \ldots, x_n) \in \mathbb{R}^n_x$, $\xi = (\xi_1, \ldots, \xi_n) \in R^n_\xi$, and $x \cdot \xi$ is the inner product of (the vector) x and (the covector) ξ: $x \cdot \xi = x_1\xi_1 + \cdots + x_n\xi_n$. We shall also write $\widehat{f}(\xi)$, or simply $\widehat{f}$, instead of $\mathcal{F}f(\xi)$. The *inverse Fourier transform* $\mathcal{F}^{-1}g$ of a function $g \in \mathcal{L}^1(\mathbb{R}^n_\xi)$ is given by

$$\mathcal{F}^{-1}g(x) = (2\pi)^{-n} \int e^{ix\cdot\xi} g(\xi)d\xi.$$

Fourier transforms of functions from the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ (which is dense in $\mathcal{L}^1(\mathbb{R}^n)$) have remarkable properties. For instance, using integration by parts, one readily verifies that if $u \in \mathcal{S}(\mathbb{R}^n)$, then one has the following nice formula, enabling one to convert problems about differential operators into algebraic problems:

$$\mathcal{F}(D^\alpha u) = \xi^\alpha \mathcal{F}u.$$

Note that this formula would be more complicated if the factor $\frac{1}{i}$ in the definition of $D_k^{\alpha_k}$ had not been included (see 2.1). Another useful formula is

$$\mathcal{F}((-1)^{|\alpha|}x^\alpha u) = D^\alpha \mathcal{F}u.$$

One of the basic facts about the Fourier transform is the following theorem ([59, 7.1.5]): The Fourier transformation $\mathcal{F} : \mathcal{S}(\mathbb{R}^n_x) \to \mathcal{S}(\mathbb{R}^n_\xi)$ is a linear topological isomorphism, its inverse being the inverse Fourier transformation $\mathcal{F}^{-1}$.

By the Parseval-Plancherel theorem, the operator $\mathcal{F} : \mathcal{S}(\mathbb{R}^n_x) \to \mathcal{S}(\mathbb{R}^n_\xi)$ can be continuously extended to an isomorphism

$$\mathcal{F} : \mathcal{L}^2(\mathbb{R}^n_x) \to \mathcal{L}^2(\mathbb{R}^n_\xi)$$

such that for all $f,\, g \in \mathcal{L}^2(\mathbb{R}^n)$ we have that

$$(f, g) = (2\pi)^{-n}(\widehat{f}, \widehat{g});$$

here the inner product (f, g) is defined to be

$$\int f(x)\overline{g(x)}dx.$$

From the Fourier transforms on $\mathcal{S}(\mathbb{R}^n)$, we pass to the Fourier transforms on the space of tempered distributions (see 2.4). For $u \in \mathcal{L}^1(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n_x)$ (identifying u with T_u), one readily obtains that

$$\langle \widehat{u}, v\rangle = \langle u, \widehat{v}\rangle,$$

$v \in \mathcal{S}(\mathbb{R}^n)$. This motivates one to define

$$\langle \mathcal{F}(s), u\rangle = \langle s, \mathcal{F}(u)\rangle,$$

$s \in \mathcal{S}'(\mathbb{R}^n_x)$, $u \in \mathcal{S}(\mathbb{R}^n_\xi)$. Then (see [59, 7.1.9]) $\mathcal{F}(s) = \widehat{s}$ is in $\mathcal{S}'(\mathbb{R}^n_\xi)$, and the Fourier transformation defines an isomorphism, $\mathcal{F} : \mathcal{S}'(\mathbb{R}^n_x) \to \mathcal{S}'(\mathbb{R}^n_\xi)$. As one readily calculates, we have $\mathcal{F}(D_k s) = \xi_k \mathcal{F}s$, $s \in \mathcal{S}'(\mathbb{R}^n)$, $k = 1, \ldots, n$.

3 Distributions and related concepts on manifolds

3.1 Preliminaries

Here and elsewhere in this text, manifolds will mostly be denoted by M, N or similarly. Frequently, when we wish to emphasize or recall that, e.g., N is of dimension d, we use the "more detailed" notation N^d.

In this section, M will always denote a manifold of dimension n. A smooth structure on M is given by a fixed maximal atlas (see [57]) of local coordinate systems (or charts) $\{(U_\alpha, \varphi_\alpha)\}_{\alpha \in A}$, where U_α (covering M) are open subsets and each φ_α is a homeomorphism from U_α to an open subset $\varphi_\alpha(U_\alpha) \subset \mathbb{R}^n$. Then for any α, $\beta \in A$ the mapping $\varphi_\alpha \circ \varphi_\beta^{-1} : \varphi_\beta(U_\alpha \cap U_\beta) \to \varphi_\alpha(U_\alpha \cap U_\beta)$ is a diffeomorphism. If U is an open subset of M, we shall denote (similarly to the Euclidean case) by $\mathcal{E}(U)$ the set of all smooth functions $U \to \mathbb{C}$. The subset of those $f \in \mathcal{E}(U)$ having compact support will be denoted by $\mathcal{D}(U)$. For any map g on U such that the meaning of $g(x) \neq 0$ is well defined (e.g., g may be a function or a section of a vector bundle), the support of g is defined to be the closure of the set $\{x \in U; g(x) \neq 0\}$.

If $\alpha = (E, \pi, B)$ is a vector bundle with the total space E, projection π and base space B, we refer to it as α or as E, and the fibre over a point $x \in B$ is denoted by α_x or by E_x. If U is a subset of B, the restriction of E over U will be denoted by $\pi^{-1}(U)$, by $E_{|U}$, or by $\alpha_{|U}$. The expressions "vector bundle of dimension k", "vector bundle of rank k", and "k-plane bundle" for $k \geq 2$ or "line bundle" for $k = 1$ are used as synonymous. If α is a vector bundle of dimension k, we sometimes denote it by α^k, to indicate its dimension. All vector bundles over M will be (supposed) finite dimensional.

We shall work with the *category of smooth complex (alternatively: real) vector bundles*; it will be denoted by $\mathbf{VB}(\mathbb{F})$ (where $\mathbb{F} = \mathbb{C}$ or $\mathbb{F} = \mathbb{R}$; see the paragraph after the next for the definition of a vector bundle morphism). We note that each vector bundle of class C^r over a C^∞-manifold has a C^∞-structure compatible with its C^r-structure, and this C^∞-structure is unique up to a C^∞-isomorphism. For this and other C^p-to-C^q questions in the category $\mathbf{VB}(\mathbb{F})$ see, e.g., [57]. Useful sources of information on vector bundles are also, e.g., [122], [60], [98], [94].

For any smooth k-plane bundle (E, π, B) we can find an open covering of B by coordinate neighbourhoods V_i (where (V_i, ψ_i) belongs to an atlas of the smooth structure on B) such that $\varphi_i : \pi^{-1}(V_i) \to V_i \times \mathbb{F}^k$ is a trivialization of E over V_i, and the composition $(\psi_i \times \mathrm{id}) \circ \varphi_i$ is a fibre-preserving diffeomorphism between $\pi^{-1}(V_i)$ and the product bundle $\psi_i(V_i) \times \mathbb{F}^k$.

In the category $\mathbf{VB}(\mathbb{F})$, a morphism from (F, π, Y) to (E, π', X) is a pair of smooth maps $(f, \bar{f}) : (F, Y) \to (E, X)$ such that $\pi' \circ f = \bar{f} \circ \pi$, and f restricts for each $y \in Y$ to a linear map $f_{|F_y} : F_y \to E_{\bar{f}(y)}$. Such a pair $(f, \bar{f})$ is called a *vector bundle morphism* from F to E. We may also speak of a vector bundle morphism f from F to E, and write it simply as $f : F \to E$, because the map $\bar{f}$ is completely determined by f (via $\bar{f}(\pi(u)) = \pi'(f(u))$). In this category, the composition of morphisms is given simply by the usual composition of maps. For each smooth manifold X, we have a subcategory **X-VB**$(\mathbb{F})$: its objects are the smooth $\mathbb{F}$-vector bundles over X, and its morphisms from (F, π, X) to (E, π', X) are the vector bundle morphisms $(f, \bar{f}) : (F, X) \to (E, X)$ such that $\bar{f} = \mathrm{id}_X$; we shall refer to them briefly as X-morphisms. In particular, if there is an X-isomorphism between two $\mathbb{F}$-vector bundles over X, then we say that they are isomorphic.

We often do not distinguish between isomorphic vector bundles. Any $\mathbb{F}$-vector bundle over X isomorphic to the product bundle $X \times \mathbb{F}^k$ will be denoted by ε^k, and will be called a *trivial k-dimensional vector bundle* over X.

For any (complex or real) vector bundle $\alpha = (E, \pi, M)$, we denote by $\Gamma(M, E)$ (briefly: $\Gamma(E)$ or $\Gamma(\alpha)$) the topological vector space of all *smooth sections* of E (that is, of smooth maps $s : M \to E$ such that $\pi \circ s = \mathrm{id}_M$). Similarly, $\Gamma_c(M, E)$ (briefly: $\Gamma_c(E)$ or $\Gamma_c(\alpha)$) will be the space of the smooth sections of E with compact support. Recall that if $f : N \to M$ is a smooth map and $s \in \Gamma(M, E)$, then we have the pullback bundle $f^*(E)$ over N and the induced section $t \in \Gamma(N, f^*(E))$, $t(y) = (y, s \circ f(y))$. In particular, if N is a subspace of M and $f : N \to M$ is the inclusion, then $f^*(E)$ is the *restriction of the vector bundle* E to N, denoted by E_N (or similarly), and the section t of E_N is then the *restriction of the section* s to N, denoted by $s_{|N}$. We note that if E is a complex line bundle over $\mathbb{R}^n$, then E is of course trivial, and the spaces $\Gamma(E)$ and $\Gamma_c(E)$ coincide with $\mathcal{E}(\mathbb{R}^n)$ and $\mathcal{D}(\mathbb{R}^n)$, respectively.

3.2 Generalized sections, distributions, and distributional densities

Let $T(M)_p$ be the tangent space to M at $p \in M$. The union of all the spaces $T(M)_p$ can be converted, in a natural way, into a smooth manifold, $T(M)$, called the *tangent manifold* of M. The manifold $T(M)$ becomes the total space of a vector bundle called the *tangent bundle* of M; the fibre over $p \in M$ is $T(M)_p$. The tangent bundle $T(M)$ is trivial over any U_α, if $(U_\alpha, \varphi_\alpha)$ is a local coordinate system. The corresponding transition functions are $g_{\alpha,\beta} : U_\alpha \cap U_\beta \to GL(n, \mathbb{R})$, $g_{\alpha,\beta}(p) = d(\varphi_\alpha \circ \varphi_\beta^{-1})(\varphi_\beta(p))$; the latter is the Jacobian matrix, at the point $\varphi_\beta(p)$, of the diffeomorphism $\varphi_\alpha \circ \varphi_\beta^{-1}$ (see 2.7). At the same time, transition functions determining the cotangent bundle $T^*(M)$ are $l_{\alpha,\beta} : U_\alpha \cap U_\beta \to GL(n, \mathbb{R})$, $l_{\alpha,\beta}(p) = d^*(\varphi_\alpha \circ \varphi_\beta^{-1})(\varphi_\beta(p))$, where the latter is the transpose of the inverse matrix to $d(\varphi_\alpha \circ \varphi_\beta^{-1})(\varphi_\beta(p))$. Of course, the cotangent bundle of M is the dual vector bundle to the tangent bundle of M.

Now we denote by $\Omega(M)$ (briefly Ω) the complexification of the real line bundle $\Omega_\mathbb{R}$ over M determined by transition functions $h_{\alpha,\beta} : U_\alpha \cap U_\beta \to GL(1, \mathbb{R})$, $h_{\alpha,\beta}(p) = |J(\varphi_\alpha \circ \varphi_\beta^{-1})(\varphi_\beta(p))|^{-1}$; here (as in 2.7) $J(\varphi_\alpha \circ \varphi_\beta^{-1})$ is the determinant of the Jacobian matrix $d(\varphi_\alpha \circ \varphi_\beta^{-1})$. The vector bundle $\Omega = \Omega_\mathbb{R} \otimes \mathbb{C}$ is known as the complex *line bundle of densities* on M (see [134, Chap. VII] or L. Loomis, S. Sternberg, Advanced calculus, Jones and Bartlett Publ. 1990). [Note that in the real theory, $\Omega_\mathbb{R}$ is the real line bundle of densities.] The fibre of Ω over a point p consists of functions $f : \Lambda^n(T(M)_p) \setminus \{0\} \to \mathbb{C}$ such that $f(kv) = |k| f(v)$ for all $k \in \mathbb{R} \setminus \{0\}$ and all $v \in \Lambda^n(T(M)_p) \setminus \{0\}$; here $\Lambda^n(T(M)_p)$ is the nth exterior power of $T(M)_p$. Smooth sections of Ω are called *smooth densities* on M; $\Gamma(\Omega)$ is the space of all smooth densities, and $\Gamma_c(\Omega)$ is the space of smooth densities with compact support on M.

For a given local coordinate system $(U_\alpha, \varphi_\alpha)$, let $X_i : U_\alpha \to \mathbb{R}$ be the ith coordinate function, $X_i = r_i \circ \varphi$, where $r_i : \mathbb{R}^n \to \mathbb{R}$, $r_i(x_1, \dots, x_n) = x_i$. Then $\frac{\partial}{\partial X_1} \wedge \dots \wedge \frac{\partial}{\partial X_n}$ is a non-vanishing section of $\Lambda^n(T(M))$ restricted over U_α. A smooth section of Ω restricted over U_α is always of the form $g(x)|dX|$, where g is a smooth complex-valued function on U_α, and $|dX| = |dX_1 \dots dX_n|$ has the value $|k|$ on $k\frac{\partial}{\partial X_1} \wedge \dots \wedge \frac{\partial}{\partial X_n}$. Using standard means (among them, a suitable partition of unity), one readily shows that elements of $\Gamma_c(\Omega)$ can be integrated on M, in an invariant manner ([134, Chap. VII]).

Let E be a complex vector bundle over M, and let E^* be the dual vector bundle; the fibre E^*_x over $x \in M$ is the dual complex vector space to the fibre E_x. Let

$$(\quad , \quad) : E_x \times (E^*_x \otimes \Omega_x) \to \Omega_x$$

denote the obvious pairing. Given $s \in \Gamma(E)$ and $t \in \Gamma(E^* \otimes \Omega)$, we define a smooth section ω of the line bundle Ω by

$$\omega(p) = (s(p), t(p)).$$

So we obtain a bilinear map $\Gamma_c(E) \times \Gamma(E^* \otimes \Omega) \to \Gamma(\Omega)$ which, when composed with the integration operator, yields a (separately continuous) bilinear map $\varphi : \Gamma_c(E) \times \Gamma(E^* \otimes \Omega) \to \mathbb{C}$. Then we have an obvious linear map $\tilde{\varphi} : \Gamma(E^* \otimes \Omega) \to \Gamma_c(E)'$ (defined by $\tilde{\varphi}(u)(v) = \varphi(v, u)$). Taking $E^* \otimes \Omega$ in the rôle of E and identifying $(E^* \otimes \Omega)^* \otimes \Omega$ with E (note that we can canonically identify E^{**} with E; in addition to this, for any line bundle L, the tensor product $L^* \otimes L$ is canonically isomorphic to the trivial line bundle), we obtain a linear map

$$\Gamma(E) \to (\Gamma_c(E^* \otimes \Omega))'.$$

This map is injective and its image is dense in $(\Gamma_c(E^* \otimes \Omega))'$. So it is natural to call elements of $(\Gamma_c(E^* \otimes \Omega))'$ *generalized sections* of E; note that they are also called *distributional sections* of E (see [13]), or *E-valued distributions*. The space of generalized sections of E over M is denoted by $\mathcal{D}'(M, E)$, or just $\mathcal{D}'(E)$ (if M is clear from the context). Similarly, elements of $(\Gamma(E^* \otimes \Omega))'$ are defined to be *generalized sections* (or *distributional sections*) of E *with compact support*. The space of such generalized sections is denoted by $\mathcal{E}'(M, E)$, or just $\mathcal{E}'(E)$.

We remark that, for a suitable complex vector bundle E, complex de Rham currents on M (see 2.1) can also be defined as elements of $\mathcal{D}'(M, E)$. For a real vector bundle V over M, one can introduce a set of generalized sections, $\mathcal{D}'(M, V)$, in an analogous way to the complex case. Again, for a suitable V, the elements of $\mathcal{D}'(M, V)$ are the de Rham real currents basically treated in [106]; for their recent applications see [51], [52].

Now we extend the notion of distributions to manifolds. We define a *distribution* u on M to be a family $u = \{u_\alpha\}_{\alpha \in A}$ of distributions $u_\alpha \in \mathcal{D}'(\varphi_\alpha(U_\alpha))$ such that, for all α, $\beta \in A$,

$$u_\beta = (\varphi_\alpha \circ \varphi_\beta^{-1})^*(u_\alpha);$$

(for the right-hand side, see 2.7). Using the gluing principle (see 2.3), one verifies (cf. [59, 6.3]) that for defining a distribution on M, in a unique way, it suffices to give a family $\{u_\alpha\}$ of distributions $u_\alpha \in \mathcal{D}'(\varphi_\alpha(U_\alpha))$, where the family $\{(U_\alpha, \varphi_\alpha)\}$ is just an *atlas* of local coordinate systems on M, if the condition

$$u_\beta = (\varphi_\alpha \circ \varphi_\beta^{-1})^*(u_\alpha)$$

is fulfilled for any $(U_\alpha, \varphi_\alpha)$, (U_β, φ_β) from that atlas. The set of all distributions on M can be identified with the space of all generalized sections of the trivial complex line bundle $M \times \mathbb{C}$ over M, hence with the space $(\Gamma_c(\Omega))' = \mathcal{D}'(M, M \times \mathbb{C})$ which we denote by

$\mathcal{D}'(M)$. [In the terminology of Guillemin and Sternberg [48], the elements of $\mathcal{D}'(M)$ are called *generalized functions* on M.] We observe that there is an algebraic isomorphism

$$\mathcal{D}'(M, E) \cong \mathcal{D}'(M) \otimes \Gamma(M, E)$$

of modules over the ring of smooth functions on M. This justifies the terminology in which the elements of $\mathcal{D}'(M, E)$ are called *sections with distributional coefficients*. In particular, a (complex) current on M can be represented, with respect to every local coordinate system, as a ($\mathbb{C}$-valued) differential form on M with distributions as coefficients. We also note that distributions on M can be identified with de Rham currents of a certain type.

We define *distributional densities* on M to be continuous linear functionals (or forms) on the space $\mathcal{D}(M)$ (with a suitable topology) of all smooth $\mathbb{C}$-valued functions with compact support on M. The space of distributional densities on M will be denoted by $\mathcal{D}'_{dd}(M)$. Any smooth density ω on M (hence $\omega \in \Gamma(\Omega(M))$) can be considered as a distributional density on M. Indeed, it defines a continuous linear functional ω on $\mathcal{D}(M)$ by

$$\langle \omega, u \rangle = \int_M u \cdot \omega$$

(note that then $u \cdot \omega$ is a smooth density with compact support, hence can be invariantly integrated on M). [For this reason, and also because we can identify $\mathcal{D}'_{dd}(M)$ with the space of *generalized* sections of Ω, $(\Gamma_c(\Omega^* \otimes \Omega))' = \mathcal{D}'(\Omega)$, distributional densities are sometimes also called *generalized densities*; cf. [48].]

If $f : M \to N$ is a smooth map, then we have a continuous linear map $f^* : \mathcal{E}(N) \to \mathcal{E}(M),\ f^*(u) = u \circ f$. Now if $f : M \to N$ is a proper smooth map (that is, the pre-image of each compact subset is compact), and $s \in \mathcal{D}'_{dd}(M)$, then one defines a distributional density $f_*(s) \in \mathcal{D}'_{dd}(N)$ by

$$\langle f_*(s), u \rangle = \langle s, f^*(u) \rangle.$$

The support $\mathrm{supp}(s)$ of $s \in \mathcal{D}'_{dd}(M)$ is defined by: $x \notin \mathrm{supp}(s)$ if there exists a neighbourhood U of x such that $\langle s, v \rangle = 0$ if $\mathrm{supp}(v) \subset U$. If $s \in \mathcal{D}'_{dd}(M)$ is a distributional density with compact support, then it defines a linear functional on the space $\mathcal{E}(M)$, by

$$\langle s, v \rangle = \langle s, \varphi \cdot v \rangle,$$

for $v \in \mathcal{E}(M)$, where $\varphi : M \to \mathbb{R}$ is a smooth function with compact support such that $\varphi(p) = 1$ for every p from some open neighbourhood of $\mathrm{supp}(s)$ (the definition does not depend on the choice of φ). With this pairing between distributional densities with compact support and smooth functions on M, if s is a distributional density with compact support and if $f : M \to N$ is any smooth map, then again

$$\langle f_*(s), u \rangle = \langle s, f^*(u) \rangle$$

defines $f_*(s)$ as a distributional density (even if f is not a proper map). The subspace in $\mathcal{D}'_{dd}(M)$ consisting of distributions with compact support is denoted by $\mathcal{E}'_{dd}(M)$.

We write the pairing between $s \in \mathcal{D}'(\Omega) = \mathcal{D}'_{dd}(M)$ and $f \in \mathcal{D}(M)$ as $\langle s, f \rangle$. Sometimes, mainly in applications to physics, the same is written (perhaps more suggestively) as

$$\int_M f \cdot s.$$

The latter is really an integral only if s is a smooth density, hence if we have $s \in \Gamma(\Omega)$.

If E is a complex vector bundle over M, we can identify

$$\mathcal{D}'(E^* \otimes \Omega) = (\Gamma_c(E^{**} \otimes \Omega^* \otimes \Omega))' = (\Gamma_c(E))'.$$

We write the corresponding pairing between $s \in \mathcal{D}'(E^* \otimes \Omega)$ and $f \in \Gamma_c(E)$ again, as in the case of generalized densities, as $\langle s, f \rangle$. Sometimes (again, mainly in physics oriented literature) the notation

$$\int_M f \cdot s$$

is also used; of course, the integral appears here, in general, just formally.

We remark that the Colombeau algebras of "new generalized functions" (see 2.6) can also be defined on smooth manifolds. About 2001, the problem of constructing a theory in which the embedding of the distributions (hence the embedding of $(\Gamma_c(\Omega))' = \mathcal{D}'(M)$) commutes with smooth local coordinate changes has been solved by M. Grosser, E. Farkas, M. Kunzinger, R. Steinbauer, and J. Vickers. The solution, together with other interesting material (also on concepts presented in this subsection), can be found in the book [47]. In [83], the authors employ generalized tensor analysis in the sense of Colombeau's construction in order to introduce a nonlinear distributional pseudo-Riemannian geometry. They also give applications to general relativity and compare some of their concepts with those presented by J. Marsden in [88]. On the Colombeau algebras on manifolds, the reader can also consult, e.g., publications by H. Balasin, J.-F. Colombeau, R. Hermann, J. Jelínek.

Now we come back to generalized sections of vector bundles as introduced earlier in this subsection.

3.3 Some properties of generalized sections of vector bundles

If $M = \mathbb{R}^n$, then both distributional densities and distributions on M can be identified with the Schwartz distributions on $\mathbb{R}^n$ (in the sense of 2.2). At the same time, generalized sections of vector bundles are generalizations of both distributional densities and distributions.

Before commenting on some properties of generalized sections, we recall a category of smooth vector bundles described, e.g., by Guillemin and Sternberg [48, App. II]. Its objects are again (as in the category $\mathbf{VB}(\mathbb{F})$) smooth complex (alternatively: real) vector bundles. A morphism from (F, π_F, Y) to (E, π_E, X) is any pair (f, r), where $f : Y \to X$ is a smooth map and r is a smooth section of the vector bundle $\mathrm{Hom}(f^*(E), F)$. In the latter, $f^*(E)$ is the induced vector bundle, hence we can identify $f^*(E)_y$ with $E_{f(y)}$ for any $y \in Y$, and the fibre of $\mathrm{Hom}(f^*(E), F)$ over $y \in Y$ is then $\mathrm{Hom}(E_{f(y)}, F_y)$, hence the vector space of linear maps $E_{f(y)} \to F_y$. (Note that any smooth section of $\mathrm{Hom}(f^*(E), F)$ defines a smooth Y-morphism from $f^*(E)$ to F.) We shall call (f, r) a smooth *sectional vector bundle morphism* (briefly: sectional morphism) from F to E; we shall write $(f, r) : (F, \pi_F, Y) \to (E, \pi_E, X)$ (or briefly $(f, r) : F \to E$). To describe the composition law $\circ$ for morphisms, let us have, in addition to (f, r), a smooth sectional morphism (g, s) from (G, π_G, Z) to (F, π_F, Y). Then we put $(g, s) \circ (f, r) = (f \circ g, r \circ s)$, where g and f compose in the standard way, while the smooth section

$$r \circ s : Z \to \mathrm{Hom}((f \circ g)^*(E), G)$$

is defined by

$$(r \circ s)(z) = s(z) \circ r(g(z)),$$

for $z \in Z$. We denote this category by **sectVB**($\mathbb{F}$), where $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$; we shall call it a *sectional category of smooth* $\mathbb{F}$*-vector bundles*. We observe that this is in fact the opposite category to the category described in 7.2.6 of N. Bourbaki, Éléments de mathématique, Fasc. XXXIII, Hermann 1967.

For example, it is clear that $(\mathrm{id}_X, r) : (F, \pi_F, X) \to (E, \pi_E, X)$ is a smooth sectional vector bundle morphism precisely when r defines a smooth X-morphism from E to F. To give another example, let $f : X \to Y$ be a smooth map, and let $df : T(X) \to T(Y)$ be its *differential* (or *tangent map*). Then df is a smooth vector bundle morphism from $T(X)$ to $T(Y)$. At the same time, let $df^*(x) : T^*(Y)_{f(x)} \to T^*(X)_x$ be the dual of the linear map $df_x : T(X)_x \to T(Y)_{f(x)}$. Then $df^* : X \to \mathrm{Hom}(f^*(T^*(Y)), T^*(X))$ is a smooth section, and the pair (f, df^*) is a smooth sectional vector bundle morphism from $T^*(X)$ to $T^*(Y)$, hence an example of a morphism in the category **sectVB**($\mathbb{R}$)).

Now if X and Y are closed manifolds and $(f, r) : (E, \pi_E, X) \to (F, \pi_F, Y)$ is a smooth sectional morphism from E to F, then we have a linear map, $(f, r)^* : \Gamma(Y, F) \to \Gamma(X, E)$, defined by $(f, r)^*(u)(x) = r(x)(u(f(x)))$, $x \in X$. Using the pairing mentioned in 3.2, we obtain a linear map,

$$(f, r)_* : \mathcal{D}'(E^* \otimes \Omega(X)) \to \mathcal{D}'(F^* \otimes \Omega(Y)),$$

defined (for $s \in \Gamma(F)$) by

$$\langle (f, r)_*(\omega), s\rangle = \langle \omega, (f, r)^*(s)\rangle.$$

So if $(f, r) : E \to F$ is a smooth sectional morphism, then smooth sections of F pull back, and generalized sections of $E^* \otimes \Omega(X)$ push forward.

Many other results on generalized sections (e.g., constructions of pullbacks of generalized sections under submersions, that is, under smooth maps $f : X \to Y$ such that $df_x : T(X)_x \to T(Y)_{f(x)}$ is surjective for each $x \in X$), together with interesting applications, can be found in [48].

3.4 The wave front sets

An important rôle in the theory of generalized sections and in applications is played by wave fronts introduced by L. Hörmander in 1971; for instance, they have been used in *micro-local analysis*. The latter is sometimes called a geometric theory of distributions; it is an analysis on the cotangent bundle, applied, e.g., to the study of systems of differential and integral equations.

Let $U \subset \mathbb{R}^n$ be open, and let $s \in \mathcal{D}'(U)$. We define the *wave front set* $WF(s) \subset U \times (\mathbb{R}^n \setminus \{0\})$ of s in the following way. A point $(x_0, \xi_0) \in U \times (\mathbb{R}^n \setminus \{0\})$ does not lie in $WF(s)$ (and s is considered to be smooth in a neighbourhood of the point (x_0, ξ_0)) if there is a function $f \in \mathcal{D}(U)$, equal to 1 in a neighbourhood of x_0, and an open conic neighbourhood Γ^0 of ξ_0 (hence $\Gamma^0 \subset \mathbb{R}^n$ is an open neighbourhood of ξ_0, and if $\xi \in \Gamma^0$, then also $\rho\xi \in \Gamma^0$ for all $\rho > 0$) such that for every $N \geq 0$ there exists a constant $C_N \geq 0$ such that

$$|(\widehat{fs})(\xi)| \leq C_N(1 + |\xi|)^{-N},$$

$\xi \in \Gamma^0$. Here (see 2.9) $\widehat{(fs)}(\xi) = \langle s(x), f(x)e^{-ix\cdot\xi}\rangle$ is the Fourier transform of fs. The set $WF(s)$ is closed and conic in $U \times (\mathbb{R}^n \setminus \{0\})$ (that is, $(x,\xi) \in WF(s)$ implies that $(x, t\xi) \in WF(s)$ for all $t > 0$).

If M is a manifold and s is a generalized section of a smooth vector bundle F over M, then $WF(s)$ is defined similarly to above (using local coordinates). Then $WF(s)$ is a well-defined conic subset of $T^*(M) \setminus \{0\}$ (the cotangent bundle minus the zero section). The wave front set $WF(s)$ may contain a better piece of information on the singularities of s than $\operatorname{singsupp}(s)$. (As in the Euclidean situation, the latter is the singular support of s, hence a closed subset of M defined by the following: $p \in M$ does *not* lie in $\operatorname{singsupp}(s)$ if there is some smooth function φ with compact support such that $\varphi(p) \neq 0$ and $\varphi \cdot s \in \Gamma(F)$.) Indeed, if $\pi : T^*(M) \setminus \{0\} \to M$ is the canonical projection, then we have $\pi(WF(s)) = \operatorname{singsupp}(s)$. For further properties of wave fronts see [59], [48].

3.5 The Schwartz kernel theorem on manifolds

To simplify matters, in this subsection, all manifolds will be (supposed) closed. For any smooth complex vector bundle we choose a smooth Hermitian metric.

Let E and F be smooth complex vector bundles over a manifold M. Let $P : \Gamma(E) \to \Gamma(F) \subset \mathcal{D}'(F)$ be a continuous linear operator. Its *Schwartz kernel* k_P (see 2.8) is a generalized section, $k_P \in \mathcal{D}'(M \times M, F\hat{\otimes}(E^* \otimes \Omega)) = (\Gamma(F^*\hat{\otimes}E))'$, where $F\hat{\otimes}E$ is the external tensor product of F and E. If $v \in \Gamma(F^*)$ and $u \in \Gamma(E)$, then the value of k_P on the section $v \otimes u$ is given by

$$\langle k_P, v \otimes u\rangle = \langle v, P(u)\rangle.$$

Let $L(\Gamma(E), \mathcal{D}'(F))$ denote the space of continuous linear operators $\Gamma(E) \to \mathcal{D}'(F)$ with the topology of uniform convergence on bounded sets. The Schwartz kernel theorem affirms that the map

$$L(\Gamma(E), \mathcal{D}'(F)) \to \mathcal{D}'(M \times M, F\hat{\otimes}(E^* \otimes \Omega)), \quad P \mapsto k_P,$$

is a bijection.

Using the Schwartz kernel, one introduces *smoothing linear operators*: a linear operator $P : \Gamma(E) \to \Gamma(F)$ is said to be *smoothing* if the Schwartz kernel k_P is smooth, hence if $k_P \in \Gamma(M \times M, F\hat{\otimes}(E^* \otimes \Omega))$. Smoothing operators play an important rôle in Atiyah and Singer's proof of their index theorem (on the theorem, see Bleecker's contribution in this Handbook) published in 1968 ([13]). At the same time, the Atiyah-Singer index theorem implies results on vector distributions, and also on immersions of manifolds in Euclidean spaces, as we shall indicate in Sections 4 and 5.

4 Vector distributions or plane fields on manifolds

4.1 Preliminaries

Let M be a manifold. A smooth *vector field* s on M is a smooth section of the tangent bundle $T(M)$, hence $s \in \Gamma(T(M))$. More generally, a smooth vector field on an open subset U of M is an element of $\Gamma(T(M)_{|U})$. Vector fields $v_1, \ldots, v_t$ on U are said to be *independent on* U if their values $v_1(x), \ldots, v_t(x) \in T(M)_x$ are linearly independent for each $x \in U$. In the sequel, all vector fields will be (supposed to be) smooth.

On the dual side, we denote by $E^p(M)$ the set of all differential p-forms (i.e., the set of all smooth sections of $\Lambda^p T^*(M)$, the pth exterior power of the cotangent bundle of M) on the manifold M, and by $E^*(M)$ we denote the set of all differential forms on M. Let $f : M \to N$ be a smooth map, and let $x \in M$. Then we have the differential $df : T(M)_x \to T(N)_{f(x)}$. Its dual, the *codifferential*, we denote now by $(df)^* : T^*(N)_{f(x)} \to T^*(M)_x$. The same notation is used for the induced algebra-homomorphism, $(df)^* : \Lambda T^*(N)_{f(x)} \to \Lambda T^*(M)_x$. If $\omega \in E^*(N)$, then we pull ω back to a form on M defined by $x \mapsto (df)^*(\omega(f(x)))$. So we obtain a map $E^*(N) \to E^*(M)$; again, we denote it by $(df)^*$, so that we have $((df)^*(\omega))(x) = (df)^*(\omega(f(x)))$. If $U \subset M$ is open, then a p-form on U is an element of $\Gamma(U, \Lambda^p T^*(M)_{|U})$; a collection $\omega_1, \ldots, \omega_t$ of 1-forms on U is called *independent on* U if $\omega_1(x), \ldots, \omega_t(x) \in T^*(M)_x$ are linearly independent for each $x \in M$. In the sequel, all differential forms will be (supposed to be) smooth.

We say ([136, 2.28]) that an ideal $\mathcal{I} \subset E^*(M)$ is *locally generated by* t *independent 1-forms* if for each $x \in M$ there exists an open neighbourhood N of x and a collection of independent 1-forms $\nu_1, \ldots, \nu_t$ on N satisfying the following two conditions:

(i) If $\omega \in \mathcal{I}$, then its restriction to N, $\omega_{|N}$, belongs to the ideal in $E^*(N)$ generated by $\nu_1, \ldots, \nu_t$.

(ii) If $\omega \in E^*(M)$, and if there is a cover of M by sets N (as above) such that for each N in the cover, $\omega_{|N}$ belongs to the ideal generated by $\nu_1, \ldots, \nu_t$, then $\omega \in \mathcal{I}$.

4.2 Vector distributions and codistributions

We say that M has a *vector distribution* (in other words: *plane field*) η of dimension k $(0 < k < n)$ if, for each $x \in M$, $\eta(x)$ is a k-dimensional subspace (briefly: k-plane) in the tangent *vector* space $T(M)_x$. Instead of the long expression "vector distribution of dimension k" (or "plane field of dimension k"), we shall mostly just say "k-distribution" (or "k-plane field"). Note that some authors use the term "*differential system of rank* k" to refer to a k-distribution.

The k-distribution η on M is smooth if for each $x \in M$ there is a neighbourhood $U \subset M$ with vector fields $v_1, \ldots, v_k$ on U such that $v_1(y), \ldots, v_k(y)$ form a basis of $\eta(y)$ for each $y \in U$. In the sequel, all k-distributions (k-plane fields) are to be smooth. More precisely, k-distributions in our sense are sometimes called *regular* distributions.

Given any vector bundle α^a over M and given a positive integer b, $b < a$, we have the associated *Grassmann fibre bundle* $(G_b(\alpha), p, M)$. The fibre of $G_b(\alpha)$ over a point $x \in M$ is the *Grassmann manifold* $G_b(\alpha_x) \cong G_b(\mathbb{R}^a)$ of all b-dimensional vector subspaces in $\alpha_x \cong \mathbb{R}^a$. By definition, an element $g \in G_b(\alpha)$ is a b-dimensional subspace of $\alpha_{p(g)} = (p^*(\alpha))_g$. So we have, over the manifold $G_b(\alpha)$, two canonical vector bundles: $\gamma \subset p^*(\alpha)$ of dimension b, and its complementary $(a-b)$-dimensional vector bundle $\gamma^\perp = p^*(\alpha)/\gamma$. We note a useful fact: the vector bundle along the fibres ([17, §7]) for the fibre bundle $(G_b(\alpha), p, M)$ can be identified with $\operatorname{Hom}(\gamma, \gamma^\perp)$, and so for the tangent bundle of the total space we have

$$T(G_b(\alpha)) \cong p^*(T(M)) \oplus \operatorname{Hom}(\gamma, \gamma^\perp).$$

A k-distribution η on a manifold M is nothing but a (smooth) section of the Grassmann bundle $G_k(T(M))$, hence a (smooth) map $\eta : M \to G_k(T(M))$ such that $\eta(x) \in G_k(T(M)_x)$ for each $x \in M$.

A k-distribution η on M defines a k-dimensional vector subbundle $\tilde{\eta}$ of $T(M)$ such that the fibre, $\tilde{\eta}_x$, for each $x \in M$, is $\eta(x)$. And vice versa, so that there is no need in distinguishing between k-distributions and k-dimensional vector subbundles of $T(M)$. If $T(M)$ has a trivial k-dimensional subbundle, we call the latter a *trivial k-distribution* on M. We shall show in 4.16-4.19 that the property of having a trivial distribution on M is very strong.

It is natural to say that η^k over M^n $(k < n)$ *defines* or *determines* a k-distribution on M if η is isomorphic to a vector subbundle of $T(M)$ or, equivalently, if there is a vector bundle monomorphism (hence a smooth vector bundle morphism whose restriction to each fibre is a linear map of rank k) $\eta \to T(M)$. Of course, each k-distribution is defined by a vector bundle monomorphism.

For any k-distribution η on M^n there is a complementary $(n-k)$-distribution κ, such that $\eta \oplus \kappa = T(M)$. The bundle κ is called the *normal distribution* to the k-distribution η, and we can say that $T(M)$ *splits*, as a Whitney sum of η and κ. In general, we say that a vector bundle *splits* if it can be expressed as a Whitney sum of a finite number of its subbundles, in a nontrivial way.

For example, if $p : M^n \to N^t$ is a smooth fibre bundle, then $T(M) = p^*(T(N)) \oplus \eta^{n-t}$, where η^{n-t} is the vector bundle along the fibres. So then η is an $(n-t)$-distribution on M, and $p^*(T(N))$ is the corresponding normal distribution. It is easy to give nontrivial specific examples of this type. Indeed, for fixed positive integers $n_1, \ldots, n_q$ $(q \geq 2)$, a *flag of type* $(n_1, \ldots, n_q)$ is defined to be a q-tuple $(S_1, \ldots, S_q)$ of mutually orthogonal subspaces in $\mathbb{R}^n$, where $n = n_1 + \cdots + n_q$ and $\dim(S_i) = n_i$. The set $F(n_1, \ldots, n_q)$ of all the flags of type $(n_1, \ldots, n_q)$ may be identified with a quotient space of the orthogonal group, $O(n)/O(n_1) \times \cdots \times O(n_q)$. This makes $F(n_1, \ldots, n_q)$ into a closed manifold known as the *flag manifold* of type $(n_1, \ldots, n_q)$. In particular, $F(n_1, n_2)$ is (up to the obvious diffeomorphism) the Grassmann manifold $G_{n_1}(\mathbb{R}^{n_1+n_2})$. By sending $(S_1, \ldots, S_q) \in F(n_1, \ldots, n_q)$ to $(S_1, \ldots, S_t, S_{t+1} \oplus \ldots \oplus S_q) \in F(n_1, \ldots, n_t, n_{t+1} + \ldots + n_q)$, for a fixed t, we obtain a smooth fibre bundle projection $F(n_1, \ldots, n_q) \to F(n_1, \ldots, n_t, n_{t+1} + \ldots + n_q)$; the fibre is $F(n_{t+1}, \ldots, n_q)$. By varying the value of t, following the general pattern described before, we can produce various distributions on the flag manifold $F(n_1, \ldots, n_q)$.

To give an example of another type, let us suppose that a manifold M^n admits k vector fields, say $v_1, \ldots, v_k$, independent on M. Whenever needed, this assumption can be replaced (thanks to the Gram-Schmidt orthonormalization) with the requirement that the set $\{v_1, \ldots, v_k\}$ be orthonormal, or that the ordered set $(v_1, \ldots, v_k)$ be an orthonormal k-frame in the sense that $(v_1(x), \ldots, v_k(x))$ should be an orthonormal k-frame in $T(M)_x$ for each $x \in M$. Of course, $v_1, \ldots, v_k$ span a trivial k-plane subbundle of $T(M)$, hence they define a trivial k-distribution on M. The question of when M admits k vector fields independent on M (for a given number k) is one of the possible formulations of the *vector field problem* (we shall focus on it in 4.16-4.19). For instance, if M is open (i.e., noncompact), then (using obstruction theory, [122]) one readily proves the existence of a nowhere vanishing vector field on M, so we have then a trivial 1-distribution on M. For closed manifolds, H. Hopf proved, around 1925, that such a manifold M has a nowhere vanishing vector field precisely when its Euler-Poincaré characteristic $\chi(M)$ is zero.

Of course, in general, not all distributions on a manifold are trivial. E.g., the Klein bottle K (recall that $\chi(K) = 0$) has a trivial 1-distribution, but the corresponding normal 1-distribution is nontrivial ($T(K)$ cannot be trivial, because K is nonorientable).

Dually to distributions, we can consider codistributions. We say that M has a *vector codistribution* (or *covector distribution*) ξ of dimension k if, for each $x \in M$, $\xi(x)$ is a k-dimensional subspace in the cotangent *vector* space $T^*(M)_x$. Instead of the long expression "vector codistribution of dimension k", we shall mostly say just "k-codistribution". A k-codistribution ξ on M is smooth if for each $x \in M$ there is a neighbourhood $U \subset M$ with 1-forms $\omega_1, \ldots, \omega_k$ on U such that $\omega_1(y), \ldots, \omega_k(y)$ form a basis of $\xi(y)$ for each $y \in U$. We consider only smooth k-codistributions. It is clear that a k-codistribution is nothing but a smooth section of the Grassmann bundle $G_k(T^*(M^n))$.

Let η^k be a vector bundle over M^n defining a k-distribution on M^n. So there is a vector bundle monomorphism $i : \eta \to T(M)$ or, equivalently, there is a vector bundle epimorphism between the dual bundles, $i^* : T^*(M) \to \eta^*$. The kernel, $\mathrm{Ker}(i^*)$, is then an $(n-k)$-dimensional vector subbundle of the cotangent bundle $T^*(M)$. We call $\mathrm{Ker}(i^*)$ the $(n-k)$-*codistribution* on M, *associated to the k-distribution* η. There is a bijective correspondence between k-distributions and associated $(n-k)$-codistributions on M.

4.3 Integrability of distributions and the Deahna-Clebsch-Frobenius theorem

Let N be a submanifold in M realized by an injective immersion $\iota : N \to M$; or briefly, let (N, ι) be a submanifold in M. If x is a point in M, we say that the submanifold (N, ι) *passes through* x if $x = \iota(n)$ for some $n \in N$. Let η be a k-distribution on M; to avoid ambiguity, from now on (if not specified differently) we shall understand η as a subbundle $\eta = (E, p, M)$ of $T(M)$; let $i : E \to T(M)$ be the inclusion, $i(v) = v$. We say that η is *involutive* if for each two sections s, $t \in \Gamma(\eta)$ (in other words, for every pair of *vector fields* s, t *lying in* η), their Lie bracket $[s, t]$ also is in $\Gamma(\eta)$. In addition to this, we define (N, ι) to be an *integral manifold* of the k-distribution η if $d\iota(T(N)_n) = \eta_{\iota(n)}$ for each $n \in N$. A k-distribution η on M is defined to be *completely integrable* if an integral manifold of η passes through each point of M.

The following is a necessary and sufficient condition for complete integrability (e.g., [136]); we call it the *Deahna-Clebsch-Frobenius theorem.* A k-distribution η on M^n is completely integrable if and only if η is involutive. If η is involutive, then for each $x \in M$ there exists a (cubic) local coordinate system (U, φ) centred at x, with coordinate functions $X_1, \ldots, X_n$ such that the slices $X_i = \text{constant}$ $(i = k+1, \ldots, n)$ are integral manifolds of η. In addition to this, if (N, ι) is a connected integral manifold of η such that $\iota(N) \subset U$, then $\iota(N)$ lies in one of these slices.

This theorem is frequently attributed exclusively to F. Frobenius; he published a proof of the core of it (in a Euclidean version) in 1877. But as remarked by J. Milnor in his MIT-lectures "Foliations and foliated vector bundles" in 1969, "as Frobenius himself pointed out, the theorem in question had been proved a decade earlier by A.Clebsch. In fact a recognizable version had been proved already in 1840, by F. Deahna".

A *maximal integral manifold* (N, ι) of a distribution η on a manifold M is defined to be a connected integral manifold of η such that there does not exist a connected integral manifold (N_1, ι_1) of η with $\iota(N)$ being a proper subset of $\iota_1(N_1)$.

A completely integrable distribution is also called a *foliation* and the maximal connected integral manifolds are called *leaves*. The leaves of a foliation (if it exists) on M give a partition of M. More precisely, we have the following theorem which we refer to as the *Frobenius theorem*; see [136, 1.64]. Let η be an involutive k-distribution on M^n. Let $x \in M$. Then through x there passes a unique maximal connected integral manifold of η, and every connected integral manifold of η passing through x is contained in *the* maximal one.

As already mentioned in the Introduction, we are not going to deal with foliations in a systematic way, although they present a highly interesting and important topic; we refer to [15] and to the references cited therein. Nevertheless, occasionally, we also shall mention facts related to foliations.

4.4 Vector distributions vs. ideals locally generated by independent 1-forms

Following Warner [136], we give yet another version of the Deahna-Clebsch-Frobenius theorem and also of the Frobenius theorem, in terms of differential forms and differential ideals, in the spirit of É. Cartan. We need some preparation.

For a differential q-form $\omega \in E^q(M)$ and $x \in M$, $\omega(x) \in \Lambda^q T^*(M)_x$ can be considered ([136, 2.18]) as an alternating multilinear function on $T(M)_x$. Given a k-distribution η on M^n, $\omega \in E^q(M)$ is said to *annihilate* η if for each $x \in M$ we have $\omega(x)(v_1, \dots, v_q) = 0$ for each $v_1, \dots, v_q \in \eta_x$. A differential form $\omega \in E^*(M)$ is defined to annihilate η if each of the homogeneous components of ω annihilates η. We put

$$\Phi(\eta) = \{\omega \in E^*(M); \omega \text{ annihilates } \eta\}.$$

The assignment $\eta \mapsto \Phi(\eta)$ defines a bijective correspondence between k-distributions on M^n and ideals in $E^*(M)$ locally generated by $n-k$ independent 1-forms. More precisely, one can prove the following theorem (see [136, 2.28]). Let η be a k-distribution on M^n. Then $\Phi(\eta)$ is an ideal in $E^*(M)$ locally generated by $n-k$ independent 1-forms and, conversely, if $\mathcal{I}$ is an ideal in $E^*(M)$ locally generated by $n-k$ independent 1-forms, then there exists a unique k-distribution η on M for which $\Phi(\eta) = \mathcal{I}$.

We shall have a closer look at the second part of this theorem. Let $\mathcal{I} \subset E^*(M)$ be an ideal locally generated by $n-k$ independent 1-forms. Let $x \in M$, and let the independent 1-forms $\nu_1, \dots, \nu_{n-k}$ generate $\mathcal{I}$ on a neighbourhood N of x. We define

$$\eta_x = \{v \in T(M)_x; \nu_1(x)(v) = 0, \dots, \nu_{n-k}(x)(v) = 0\},$$

hence the annihilator of η_x is the subspace $[\nu_1(x), \dots, \nu_{n-k}(x)]$ in $T^*(M)_x$ spanned by $\nu_1(x), \dots, \nu_{n-k}(x)$. Then $\bigcup_{x \in M} \eta_x$ is the (total space of the) desired k-distribution η on M such that $\Phi(\eta) = \mathcal{I}$.

It is quite instructive to look at

$$\{v \in T(M)_x; \nu_1(x)(v) = 0, \dots, \nu_{n-k}(x)(v) = 0\}$$

from a different point of view. Let $i : \eta \to T(M)$ be the inclusion realizing η as a subbundle of $T(M)$. Then clearly the subspace $[\nu_1(x), \dots, \nu_{n-k}(x)]$ in $T^*(M)_x$ is nothing but $\mathrm{Ker}(i^*)_x$, hence $\nu_1(x), \dots, \nu_{n-k}(x)$ span the fibre (over x) of the $(n-k)$-codistribution associated to the k-distribution η on M.

4.5 Pfaffian systems and historical remarks

In the literature on differential equations, a system of equations

$$\nu_1 = 0, \dots, \nu_{n-k} = 0,$$

where $\nu_1, \dots, \nu_{n-k}$ are 1-forms defined and independent on an open subset U in $\mathbb{R}^n$ (or in an n-dimensional manifold), is sometimes called a *system of total differential equations*; it is also called a *Pfaffian system*. A solution to such a system is a k-dimensional submanifold $S \subset U$ such that (in our terminology) the forms $\nu_1, \dots, \nu_{n-k}$ annihilate $T(S)$. J. Pfaff started to study systems now bearing his name as early as 1814.

We have the following version of the Deahna-Clebsch-Frobenius theorem [136, 2.30]): A distribution η on M is involutive if and only if the ideal $\Phi(\eta)$ is a *differential ideal* (hence if $d(\Phi(\eta)) \subset \Phi(\eta)$, where d denotes exterior differentiation).

One can define integral manifolds for ideals of differential forms. An *integral manifold* of an ideal $\mathcal{I} \subset E^*(M)$ is defined to be a submanifold (N, ι) in M such that $(d\iota)^*(\omega) = 0$ for every $\omega \in \mathcal{I}$. A *maximal integral manifold* (N, ι) of an ideal $\mathcal{I}$ is defined to be a connected integral manifold of $\mathcal{I}$ such that there does not exist a connected integral manifold (N_1, ι_1) of $\mathcal{I}$ with $\iota(N)$ being a proper subset of $\iota_1(N_1)$.

Now the Frobenius theorem, in terms of differential ideals, is the following (see [136, 2.32]). Let $\mathcal{I} \subset E^*(M^n)$ be a differential ideal locally generated by $n-k$ independent 1-forms. Let $x \in M$. Then there exists a unique maximal, connected, integral manifold of $\mathcal{I}$ passing through x, and this integral manifold is of dimension k.

In view of what we said above, it is clear that, for any $x \in M$, the maximal connected integral manifold of $\mathcal{I}$ passing through x is defined by a Pfaffian system. The problem of describing the maximal connected integral manifolds for all $x \in M$ is known as the *Pfaff problem*.

The term "distribution" was introduced by C. Chevalley in his book Theory of Lie groups, Princeton Univ. Press 1946. In the 1940's further influential works of C. Ehresmann and G. Reeb appear, and the theory of foliations emerges on the scene. Nevertheless, studies of various problems on distributions, under other names, are much older: they date back to the 19th century. Among them, the studies by Deahna, É. Cartan, Clebsch, Frobenius, Grassmann and others of Pfaffian systems in Euclidean spaces were very important. From this line of development it becomes clear why k-distributions are also called k-dimensional *Pfaffian structures*.

Vector distributions appear in various areas of mathematics. For some of their applications, the reader may consult, e.g., R. Abraham, J. Marsden, and T. Ratiu's [1], P. Griffiths's [43], V. Gershkovich and A. Vershik's [36], O. Krupková's [82], R. Bryant, P. Griffiths, and D. Grossman's [20]. For more applications, see e.g. works by J. Adachi, M. Zhitomirskij, W. Respondek, W. Pasillas-Lépine, I. Zelenko, D. Krupka, B. Dubrov, B. Komrakov, O. Gil-Medrano, J. C. González-Dávila, L. Vanhecke, P. Mormul, W. Krynski, B. Jakubczyk, G. Cairns, P. Molino, M. de León, J. Marín-Solano, J. C. Marrero, M. Castrillón López, J. Muñoz Masqué, A. Weinstein, R. Montgomery. We shall mainly concentrate on the existence question for vector distributions.

4.6 On the existence question for vector distributions

For a given k, one naturally wishes to have some necessary and some sufficient (or better, if possible, some necessary and sufficient) conditions for manifolds to possess k-distributions. For instance, a sufficient condition for a manifold to possess a completely integrable 1-distribution (hence to possess a 1-dimensional smooth foliation) is that the manifold be open. This is implied by obstruction theory and the fact that each open n-dimensional manifold is homotopy equivalent to a CW-complex of dimension $n-1$. From now on, we concentrate on closed manifolds.

Recall (e.g., [94]) that two smooth closed n-dimensional manifolds M and N are *cobordant* if there is a smooth compact $(n+1)$-dimensional manifold such that its boundary is (diffeomorphic to) the disjoint union of M and N. For a vector bundle α over M, let $w_i(\alpha) \in H^i(M;\mathbb{Z}_2)$ be the ith Stiefel-Whitney characteristic class of α, and for a manifold M, let $w_i(M)$ be $w_i(T(M))$. For a manifold M^n and non-negative integers $r_1,\ldots,r_n$ such that $r_1+2r_2+\cdots+nr_n=n$, the value

$$\langle w_1(M)^{r_1}\cdots w_n(M)^{r_n},[M]\rangle,$$

where $[M]\in H_n(M;\mathbb{Z}_2)$ is the mod 2 fundamental class of M, is the Stiefel-Whitney number $w_1^{r_1}\cdots w_n^{r_n}[M]$ of M. Cobordism is an equivalence relation, the cobordism classes of n-dimensional closed manifolds form an additive group, the unoriented cobordism group $\mathcal{N}_n$. As is well known, M^n and N^n represent the same class in $\mathcal{N}_n$ if and only if all their Stiefel-Whitney numbers coincide. So we can speak about the Stiefel-Whitney numbers $w_1^{r_1}\cdots w_n^{r_n}(a)$ of a class $a\in\mathcal{N}_n$.

When we look at the existence question for vector distributions *up to cobordism*, the answer is quite simple. Indeed, by R. Stong [123]: A class $a\in\mathcal{N}_n$ is represented by a manifold M^n having a k-distribution $(k\le n)$ if and only if either

(a) k is even, or

(b) k is odd and $w_n(a)$ is zero.

In contrast to this, the original question (without considering manifolds up to cobordism or some other nontrivial equivalence relation) has no complete general answer. In the sequel, we shall have in mind precisely the original question. In attempts to solve it, several approaches appeared; we shall mention some of them later, presenting also some of the results achieved. Nevertheless, already now, before entering a more detailed exposition, we should mention an approach in which the Schwartzian distributions come in an interplay with the vector distributions. Indeed, in the 1960's-1970's, M. Atiyah and his collaborators developed an approach to the existence question of k-distributions based on the index theory of elliptic differential operators (see Bleecker's contribution in this Handbook). Basically, they observed that certain invariants of a manifold (e.g., the Euler-Poincaré characteristic) are indices of elliptic operators, and the existence of k-distributions implies that these operators have some properties, implying corresponding results for the indices. So, for instance Atiyah ([11]) proved the following necessary condition for oriented 2-distributions on oriented manifolds. Let M^n be a smooth closed connected oriented manifold, and let $n\equiv 0 \pmod 4$. Let $\sigma(M)$ denote the signature of M (see, e.g., [94, §19]). If M has an oriented 2-distribution, then $\chi(M)$ is even and $\chi(M)\equiv\sigma(M) \pmod 4$. For the question of the existence of trivial k-distributions, Atiyah and his collaborators were able to derive still deeper results by employing the Atiyah-Singer index theorem (recall that in one of its proofs Atiyah and Singer made use of the Schwartzian distributions). We come to such results in some detail later. Now we shall consider the existence question for various values of k.

4.7 The existence of 1-distributions

In 1973, U. Koschorke proved the following criterion ([76, 2.1, 2.2]). Let M^n be a closed and connected manifold, and let ξ be a line bundle over M. Then ξ is isomorphic to a subbundle of $T(M)$ if and only if the Euler-Poincaré characteristic $\chi(M)$ of M vanishes, when n is even, resp. if and only if

$$\langle\sum_{i=0}^{n} w_1(\xi)^{n-i} w_i(M), [M]\rangle = 0,$$

when n is odd.

As a consequence, we readily see that a closed connected manifold M has a 1-distribution if and only if $\chi(M) = 0$. We also can say that M has a 1-dimensional foliation if and only if $\chi(M) = 0$; it is true that each 1-distribution is completely integrable.

4.8 Generalized k-distributions with $k \geq 2$

Passing to k-distributions with $k \geq 2$, we take a more general point of view (cf. [128], [79]). Let ξ^q and η^k be (real) vector bundles over a CW-complex X, with $k < q$. We say that η is a *k-distribution in ξ* if η is isomorphic to a subbundle of ξ or equivalently, if there is a (smooth) vector bundle monomorphism $\eta \to \xi$. In particular, to say that η is a k-distribution in $T(M)$ is the same as to say (in our earlier terminology) that η defines a k-distribution on M. Another particular case: since the vector bundle $T(M) \oplus \varepsilon^t$ $(t \geq 1)$ is the *stable tangent bundle* of M, it is natural to say that η^k *defines* a *stable k-distribution* on M if $\eta \oplus \varepsilon^t$ is a $(k + t)$-distribution in $T(M) \oplus \varepsilon^t$ (that is, if there is a vector bundle monomorphism $\eta \oplus \varepsilon^t \to T(M) \oplus \varepsilon^t$). Note (consult, e.g., [60, Part II, Chap. 8, Theorem 1.5] if needed) that this definition does not depend on t if $t \geq 1$; this explains the word “stable”. Let M be a manifold. If a k-distribution η is defined at all but the points of some subset $S \subset M$, then S is called the *singularity* of η. If S is finite (resp. infinite), then we say that η is a *k-distribution with a finite* (resp. *infinite*) *singularity on M*. We observe that if η is a k-dimensional vector bundle over a closed, smooth, connected manifold M^n such that $\eta \oplus \varepsilon^1$ is a subbundle in $T(M) \oplus \varepsilon^1$, then $\eta_{|M_{(n-1)}}$ (where $M_{(t)}$ denotes the t-skeleton of M) is a subbundle of $T(M)_{|M_{(n-1)}}$ and, as a consequence, we obviously obtain a k-distribution with a finite singularity on M. In other words, if there exists a stable k-distribution on M, then there is a k-distribution with a finite singularity on M.

4.9 Obstructions to removing finite singularities and obstructions to liftings

There are two main approaches to the study of the existence question for k-distributions with $k \geq 2$. In the first, one takes a k-distribution with a singularity as a “starting point”; it is then needed to characterize (calculable) obstructions to removing the singularity. In the second, one tries to solve certain lifting problems (in fibrations) which are equivalent to the existence question for k-distributions; it is then needed to characterize (calculable) obstructions. But there is no sharp border line between the two points of view; on the contrary, there are instances where they naturally complement each other.

We first restrict ourselves to *oriented* k-distributions on oriented manifolds. Of course, one has an oriented k-distribution on an oriented manifold M^n $(0 < k < n)$ if and only if there exists a section of the associated fibre bundle $\tilde{G}_k(T(M))$ over M.

The fibre of $\tilde{G}_k(T(M))$ over a point $x \in M$ is the *oriented Grassmann manifold* $\tilde{G}_k(T(M)_x) \cong \tilde{G}_k(\mathbb{R}^n)$ of oriented k-dimensional vector subspaces in the oriented tangent space $T(M)_x$. Suppose that we have an oriented k-distribution η with a finite singularity $S = \{p_1, \ldots, p_t\}$ on M. Then we take a triangulation on M^n such that each singular point p_i lies in the interior of an n-simplex σ_i, and each σ_i contains at most one point of S. Since each n-simplex σ_i is contractible, the restriction $TM|_{\sigma_i}$ is trivial and we can, for each i, choose an orientation preserving trivialization isomorphism $TM|_{\sigma_i} \cong \sigma_i \times \mathbb{R}^n$; this isomorphism can be chosen so that it is compatible with the standard metric on $\sigma_i \times \mathbb{R}^n$. The boundary $\dot{\sigma}_i$ is an oriented $(n-1)$-sphere. Clearly $x \mapsto \eta_x$, for any $x \in S^{n-1} \cong \dot{\sigma}_i$, defines an element o_i of the homotopy group $\pi_{n-1}(\tilde{G}_k(\mathbb{R}^n))$. The element o_i is the *local obstruction* to eliminating the singularity at $p_i \in S$ (see [120, Ch. 1, §3, Theorem 12]). One can then define a *global obstruction* to deforming η into a k-distribution on M as the sum

$$\sum_{i=1}^{t} o_i \in \pi_{n-1}(\tilde{G}_k(\mathbb{R}^n));$$

we denote it by $\mathcal{O}(\eta)$. [We remark that $\mathcal{O}(\eta)$ is also called (see, e.g., [131]) the *index* of η; but we do not call it so here, to avoid confusion with the index of an elliptic operator.] One readily verifies that, by Poincaré duality, $\mathcal{O}(\eta)$ corresponds to the Steenrod obstruction class ([122, Ch. III]) in $H^n(M; \pi_{n-1}(\tilde{G}_k(\mathbb{R}^n)))$. Therefore $\mathcal{O}(\eta)$ is zero precisely when there is a k-distribution on M (without any singularity) which agrees with η on the $(n-2)$-skeleton of M.

The global obstruction $\mathcal{O}(\eta)$ is independent of the triangulation and other choices made above. But, in general, it can be changed, if we reverse the orientation on M. Of course, the information hidden in it becomes "visible" and useful only when we are able to find a "readable" expression for it, in terms of computable invariants (characteristic classes etc.).

The obstruction $\mathcal{O}(\eta)$ is straightforwardly associated with the question: Does M^n admit an oriented k-distribution? Yet it is useful to modify the *direct* existence question slightly, to the following *indirect* one ([128], [131]): Given an oriented vector bundle η^k over an oriented (smooth, closed, connected) manifold M^n, is η a k-distribution in $T(M)$?

More generally, let ξ^q and η^k be oriented vector bundles over a CW-complex X, with $k < q$. Let $BSO(m)$ $(m \geq 1)$ denote the classifying space for oriented m-dimensional vector bundles, and let γ^m denote the canonical classifying m-plane bundle over $BSO(m)$. Obviously, ξ and η (more precisely, their classifying maps) determine a map $(\xi, \eta) : X \to BSO(q) \times BSO(k)$. Additionally, the pair of vector bundles $(\gamma^{q-k} \times \gamma^k, \varepsilon^0 \times \gamma^k)$ defines a map

$$\pi_{q,k} : BSO(q-k) \times BSO(k) \to BSO(q) \times BSO(k),$$

where ε^0 denotes the trivial 0-dimensional vector bundle over $BSO(q-k)$. One readily verifies that η is a k-distribution in ξ if and only if there is a map $\zeta : X \to BSO(q-k)$ such that (ζ, η) lifts the map (ξ, η), up to homotopy, to $BSO(q-k) \times BSO(k)$ or, in other words, makes the following diagram homotopy-commutative:

$$\begin{array}{ccc} & & BSO(q-k) \times BSO(k) \\ & \overset{(\zeta,\eta)}{\nearrow} & \downarrow{\scriptstyle \pi_{q,k}} \\ X & \xrightarrow{(\xi,\eta)} & BSO(q) \times BSO(k). \end{array}$$

As is well known (see, e.g., [120, Chap. II, §8, Theorem 9]), $(BSO(q-k) \times BSO(k), \pi_{q,k}, BSO(q) \times BSO(k))$ can be regarded as a fibration, having as fibre ([128]) the *Stiefel manifold* $V_{q,k} = SO(q)/SO(q-k)$ of orthonormal k-frames in $\mathbb{R}^q$. In principle, to see whether such a "lift" (ζ, η) exists, one can use the Postnikov resolution of $\pi_{q,k}$ or some suitable modification.

The manifold $V_{q,k}$ is $(q-k-1)$-connected. So a standard procedure (apply, e.g., [122, 29.2] and [120, Chap. 7, §6, Theorem 17]) leads to a map $(\zeta', \eta') : X_{(q-k)} \to BSO(q-k) \times BSO(k)$ (where $X_{(t)}$ denotes the t-dimensional skeleton of X) such that the following diagram commutes up to homotopy:

$$\begin{array}{ccc} & & BSO(q-k) \times BSO(k) \\ & \overset{(\zeta',\eta')}{\nearrow} & \Big\downarrow \pi_{q,k} \\ X_{(q-k)} & \xrightarrow{(\xi,\eta)} & BSO(q) \times BSO(k). \end{array}$$

This means that $\eta_{|X_{(q-k)}}$ *is* a k-distribution in $\xi_{|X_{(q-k)}}$. Then (see [128]) $\eta_{|X_{(q-k+1)}}$ *is* a k-distribution in $\xi_{|X_{(q-k+1)}}$ if and only if

$$w_{q-k+1}(\xi - \eta) = 0, \text{ for } q-k \text{ odd},$$

$$\beta^* w_{q-k}(\xi - \eta) = 0, \text{ for } q-k \text{ even}.$$

Here $w_i(\xi - \eta) \in H^i(X; \mathbb{Z}_2)$ denotes the ith Stiefel-Whitney class of the virtual vector bundle $\xi - \eta$, and $\beta^* : H^i(X; \mathbb{Z}_2) \to H^{i+1}(X; \mathbb{Z})$ is the Bockstein homomorphism associated with the short exact sequence

$$0 \to \mathbb{Z} \to \mathbb{Z} \to \mathbb{Z}_2 \to 0.$$

In particular, we have the following theorem (Thomas [128]). Let M^n be an orientable smooth closed connected manifold and let η^k be an oriented vector bundle over M, $1 < k < n$. Then η is a k-distribution over the $(n-k+1)$-skeleton of M if and only if

$$w_{n-k+1}(T(M) - \eta) = 0, \text{ for } n-k \text{ odd},$$

$$\beta^* w_{n-k}(T(M) - \eta) = 0, \text{ for } n-k \text{ even}.$$

4.10 Oriented 2-distributions on orientable even-dimensional manifolds

By Thomas [131], if M^n $(n \geq 4)$ is an orientable smooth closed connected manifold with n even, and if η is an oriented 2-plane bundle over M, then $\beta^* w_{n-2}(T(M) - \eta) = 0$. By the above theorem, η gives a 2-distribution on $M_{(n-1)}$, and obviously extends to a 2-distribution with a finite singularity on M. So with each oriented 2-dimensional vector bundle η over an oriented even-dimensional smooth, closed, connected manifold M^{2t} $(t \geq 2)$ we can associate the global obstruction

$$\mathcal{O}(\eta) \in \pi_{2t-1}(\tilde{G}_2(\mathbb{R}^{2t})) \cong \pi_{2t-1}(V_{2t,2}).$$

The knowledge of the homotopy groups of the Stiefel manifolds implies that $\mathcal{O}(\eta)$ is an element of the group $\mathbb{Z} \oplus \mathbb{Z}$ if $t = 2$, and it is an element of $\mathbb{Z} \oplus \mathbb{Z}_2$ if $t \geq 3$. Hence in the latter case, one can speak of the components of $\mathcal{O}(\eta)$ as the $\mathbb{Z}$-obstruction and the $\mathbb{Z}_2$-obstruction.

In 1958, F. Hirzebruch and H. Hopf ([58]) found all possible pairs of integers which can occur as the global obstruction for some oriented 2-distribution with a finite singularity on an orientable 4-manifold. Following Y. Matsushita [89], we describe a refined version of the corresponding necessary and sufficient condition for the existence of oriented 2-distributions. Let M^4 be an oriented smooth closed manifold, and let H denote the free abelian group

$$H^2(M;\mathbb{Z})/\,\text{Torsion subgroup}\,.$$

The orientation of M gives an isomorphism $H^4(M;\mathbb{Z}) \cong \mathbb{Z}$, and thanks to Poincaré duality we have a symmetric nonsingular bilinear form $\mu_M : H \otimes H \to \mathbb{Z}$, called the *intersection form* of M. [Recall that the signature of μ_M is nothing but the signature $\sigma(M)$.] The following theorem in [89] the author attributes to O. Saeki (the 0-form is considered as a special case of the positive definite forms). Let M^4 be an oriented closed manifold. If the intersection form μ_M is indefinite, then M admits an oriented 2-distribution if and only if $\sigma(M) + \chi(M) \equiv 0 \pmod 4$ and $\sigma(M) - \chi(M) \equiv 0 \pmod 4$. If μ_M is definite (either positive or negative), then M admits an oriented 2-distribution if and only if $\sigma(M) + \chi(M) \equiv 0 \pmod 4$, $\sigma(M) - \chi(M) \equiv 0 \pmod 4$, and $|\sigma(M)| + \chi(M) \geq 0$.

For orientable M^{2t} with $t \geq 3$, the known results on global obstructions of oriented 2-distributions with finite singularities are not so complete as for orientable 4-dimensional manifolds. Note that an oriented vector bundle η^2 over M is completely determined by its Euler class $e(\eta) \in H^2(M;\mathbb{Z})$. Following Thomas [129], for an oriented manifold M^{2t} and $u \in H^2(M;\mathbb{Z})$, we define

$$\theta(u) = \sum_{i,j\geq 0;\, i+j=t-1} w_{2j}(M)\cdot u^i \in H^{2t-2}(M;\mathbb{Z}_2).$$

For any $u \in H^p(M;\mathbb{Z})$ and $v \in H^q(M;\mathbb{Z})$ such that $p+q=2t$, we define $\Gamma(u,v) \in \mathbb{Z}$ by

$$\Gamma(u,v) = \langle u\cdot v, [M]\rangle,$$

where $[M] \in H_{2t}(M;\mathbb{Z})$ is the orientation class of M. Then Thomas (using suitable Postnikov resolutions) characterized in ([129, 1.2]) those integers that can arise as the $\mathbb{Z}$-obstruction of an oriented 2-distribution with a finite singularity. Let M^{2t} $(t \geq 3)$ be a smooth closed connected oriented manifold, and let $u \in H^2(M;\mathbb{Z})$. Then the following integers, and only these, occur as the $\mathbb{Z}$-obstruction of oriented 2-distributions on M with finite singularities, and with Euler class u:

$$\chi(M) - \Gamma(u,v),$$

where v runs over all classes in $H^{2t-2}(M;\mathbb{Z})$ such that $v \mod 2 = \theta(u)$.

In general, E. Thomas did not calculate the $\mathbb{Z}_2$-obstruction of an oriented 2-distribution with finite singularities on M^{2t}. However, he has shown in [129] that if $\dim(M) \equiv 2 \pmod 4$, and if $e(\eta) \mod 2 = 0$, then the $\mathbb{Z}_2$-obstruction for η is zero, and he has proved the following result. Let M^{4t+2} $(t \geq 1)$ be a smooth closed connected oriented manifold, and let $u \in H^2(M;\mathbb{Z})$. Then there exists an oriented 2-distribution on M, with Euler class

$2u$, if and only if there exists a cohomology class $v \in H^{4t}(M;\mathbb{Z})$ such that $v \mod 2 = w_{4t}(M)$, and $2\Gamma(u,v) = \chi(M)$. For $\dim(M) \equiv 0 \pmod 4$, the $\mathbb{Z}_2$-obstruction was computed later by Atiyah and Dupont [12, 6.1]; they applied the index theory for elliptic operators and *Real K-theory* (also called KR-theory; it is defined for locally compact Hausdorff spaces with involution, also called *Real spaces*). More precisely, writing n in the form $4t-s$, they define homomorphisms $\theta^s : \pi_{n-1}(V_{n,k}) \to KR^s(\mathbb{R}P^{k+s-1}, \mathbb{R}P^{s-1})$ which turn out to be isomorphisms for $k \leq 3 \leq n-k$ (here and elsewhere, $\mathbb{R}P^q$ is real projective q-space). They prove: Let M^n be a smooth closed connected oriented manifold, and let $n \geq 8$, $n = 4t$. Then for an oriented 2-distribution with a finite singularity on M the $\mathbb{Z}_2$-obstruction is $\frac{1}{2}(\chi(M) - (-1)^t\sigma(M)) \mod 2$.

Crabb and Steer in [28], by a detailed study of the image under θ of global obstructions, extended some results of Atiyah and Dupont [12]. In particular, they proved: Let M^n be a smooth closed connected oriented manifold with n ($n \geq 6$) even. Then M admits a *spin* 2*-distribution* (that is, an orientable 2-plane subbundle of $T(M)$ such that its second Stiefel-Whitney class is zero) if and only if the Bockstein operation on $w_{n-2}(M)$ vanishes, $\chi(M) \equiv 0 \pmod q$ and, when $m \equiv 0 \pmod 4$, $\sigma(M) \equiv \chi(M) \pmod 4$. Here q is the integer:

0 if $\operatorname{rank}(H^2(M;\mathbb{Z})) = 0$,

4 if $\operatorname{rank}(H^2(M;\mathbb{Z})) \neq 0$ and $w_{n-2}(M) \in H^{n-2}(M;\mathbb{Z}_2)$ is the reduction mod 2 of a torsion element in $H^{n-2}(M;\mathbb{Z})$,

2 otherwise.

For oriented 2-distributions on 8-dimensional spin manifolds, see also [21].

4.11 2-distributions on nonorientable even-dimensional manifolds

Let M be a closed connected manifold. Each element x of the cohomology group $H^1(M;\mathbb{Z}_2)$ (which can be identified with the set of homotopy classes of maps $M \to \mathbb{R}P^\infty$) induces a homomorphism of fundamental groups, $\pi_1(x) : \pi_1(M) \to \mathbb{Z}_2 = \{-1,1\} = \operatorname{Aut}(\mathbb{Z})$. In the sequel, $\mathbb{Z}_x$ will denote the *x-twisted* (or *local*) *integer coefficients* given by $\pi_1(x)$ (if needed, consult, e.g., Sec. 3.H in A. Hatcher, Algebraic topology, Cambridge University Press 2002).

Using obstruction theory (Postnikov resolutions) with twisted integer coefficients $\mathbb{Z}_{w_1(M)}$ (see also [101]), M.H.P.L. Mello in [91] proves the following. Suppose that M^n is a smooth closed connected nonorientable manifold and η is a 2-dimensional vector bundle over M.

(i) When n ($n \geq 6$) is even and η is oriented, then η defines a 2-distribution on M^n if and only if the twisted Bockstein operation on $w_{n-2}(T(M)-\eta)$ vanishes and there is a class $x \in H^{n-2}(M;\mathbb{Z}_{w_1(M)})$ such that the cup product of x and the Euler class of η equals the twisted Euler class of $T(M)$ and $w_{n-2}(T(M)-\eta) = x \mod 2$.

(ii) When $n \equiv 0 \pmod 4$ ($n \geq 8$) and η is nonorientable such that $w_1(\eta) = w_1(M)$, then η defines a 2-distribution on M if and only if $\beta^* w_{n-2}(T(M)-\eta) = 0$ and there is a cohomology class $x \in H^{n-2}(M;\mathbb{Z})$ such that the cup product of x and the $w_1(M)$-twisted Euler class of η equals the twisted Euler class of $T(M)$ and $w_{n-2}(T(M)-\eta) = x \mod 2$.

(iii) When $n \equiv 2 \pmod 4$ $(n \geq 6)$ and η is nonorientable such that $w_1(\eta) = w_1(M)$, then η defines a 2-distribution on M if the following conditions are satisfied: $\beta^* w_{n-2}(T(M) - \eta) = 0$, there exists a class $x \in H^{n-2}(M; \mathbb{Z})$ such that the cup product of x and the $w_1(M)$-twisted Euler class of η equals the twisted Euler class of $T(M)$, $w_{n-2}(T(M) - \eta) = x \mod 2$ and, in addition to this, there exists a cohomology class $v \in H^{n-2}(M; \mathbb{Z})$ such that the cup product of the mod 2 reduction of v with $w_1^2(M) + w_2(\eta)$ is nonzero.

4.12 2-distributions on odd-dimensional manifolds

If $n \equiv 3 \pmod 4$ and M^n is a smooth closed orientable connected manifold, then there are three vector fields independent on M^n, hence there is a trivial oriented 3-distribution – clearly also a trivial oriented 2-distribution – on M. In addition to this, Thomas in [128] proves (using Postnikov decompositions) that if $u \in H^2(M; \mathbb{Z})$, then M has an oriented 2-distribution with Euler class u if and only if $\sum_{i,j \geq 0; 2i+j=n-1} w_j(M) \cdot u^i \in H^{n-1}(M; \mathbb{Z}_2)$ vanishes.

So we are left with the case $n \equiv 1 \pmod 4$. For any odd-dimensional manifold M^{2k+1}, one has its Kervaire mod 2 semi-characteristic

$$\hat{\chi}(M) = (\sum_{i=0}^{k} \dim(H_i(M; \mathbb{Z}_2))) \pmod 2.$$

Thomas [128] proves: Let M^n be an orientable smooth closed connected manifold such that $n \equiv 1 \pmod 4$ and $n \geq 5$. Suppose that $w_2(M) = 0$ (as a consequence, M admits a spin structure). Then M has an oriented 2-distribution with Euler class $2v$ for each $v \in H^2(M; \mathbb{Z})$ if and only if $w_{n-1}(M) = 0$ and $\hat{\chi}(M) = 0$.

For a smooth closed connected manifold M of odd dimension we define the *real* Kervaire semi-characteristic $R(M)$ by $R(M) = \sum_j \dim_{\mathbb{R}} H^{2j}(M; \mathbb{R}) \pmod 2$. In [14], Atiyah and Singer show that $R(M)$ has an analytical interpretation when $\dim(M) \equiv 1 \pmod 4$: $R(M)$ is then the dimension modulo 2 of the kernel of a certain *real* elliptic skew-adjoint operator. Using this and Real K-theory, then Atiyah and Dupont [12, Theorem 7.2] calculated the global obstruction for any oriented 2-distribution with a finite singularity on any orientable manifold M^n with $n \equiv 1 \pmod 4$, $n \geq 5$. Their result is the following. Let M^n be a smooth closed connected oriented manifold with $n \equiv 1 \pmod 4$, $n \geq 5$. Then for any oriented 2-distribution with a finite singularity on M the global obstruction is the semi-characteristic $R(M)$.

Now we pass to 2-distributions on nonorientable odd-dimensional manifolds. Atiyah and Dupont also derived results on 2-distributions on nonorientable manifolds M^n such that $n \equiv 1 \pmod 4$ and $w_1^2(M) = 0$; see [12, §7] for details. Mello in [91] proves the following sufficient condition. Suppose that M^n is a smooth closed connected nonorientable manifold with n $(n \geq 5)$ odd and η is a 2-plane bundle over M. Suppose that $w_1^2(M) + w_1^2(\eta) + w_1(\eta)w_1(M) + w_2(\eta) \neq 0$ for $n \equiv 3 \pmod 4$ or $w_1^2(M) + w_1^2(\eta) + w_1(\eta)w_1(M) \neq 0$ for $n \equiv 1 \pmod 4$. Then η defines a 2-distribution on M if and only if $w_{n-1}(T(M) - \eta) = 0$.

4.13 k-distributions with $k \geq 3$

Results available for 3-distributions cover just some of the possible situations; they are certainly much less complete than those for $k \leq 2$.

For orientable even-dimensional manifolds, one can transform the existence problem for oriented k-distributions into another (in general difficult) problem. More precisely, one says that a stable vector bundle (α) has *geometric dimension* $\leq t$ if there is a t-dimensional vector bundle stably equivalent to α. For n even and $k < n$ odd, by Thomas [128, 5.1], an oriented vector bundle η defines an oriented k-distribution on an oriented manifold M^n if and only if $\chi(M) = 0$ and the geometric dimension of the virtual bundle $T(M) - \eta$ is at most $n-k$. Using this, Thomas derived [128, Theorem 1.5]: Let M^n with $n \equiv 2 \pmod 4$ $(n \geq 6)$ be an orientable manifold such that $w_{n-2}(M) = 0$ and $\chi(M) = 0$. Then every 3-dimensional spin vector bundle over M defines a 3-distribution on M.

Atiyah in [11, Theorem 4.1] gives a sufficient condition: Let M^{4q+1} $(q \geq 1)$ be a smooth closed oriented connected manifold. If M admits a k-distribution with $k \equiv 2 \pmod 4$, then $R(M) = 0$.

Also Crabb and Steer in [28] have several results on (stable) k-distributions with $k = 3$. Among other papers discussing conditions for k-distributions, we mention at least [81].

4.14 Obstructions to removing infinite singularities

We now do not require either k-distributions or manifolds to be orientable. Let M^n be a smooth, closed, connected manifold, and let η be a k-plane bundle over M. We shall approach the question of whether η defines a k-distribution on M, hence whether there exists a vector bundle monomorphism $\eta \to T(M)$, using a *singularity approach*, developed in the 1970's, mainly by U. Koschorke and H. Salomonsen. Some related considerations and results can also be found in papers by J.-P. Dax, A. Hatcher and F. Quinn.

Let η^k and β^b be smooth real vector bundles over a manifold M^n (possibly with boundary). Let $u : \eta \to \beta$ be a smooth vector bundle homomorphism, and let s_u denote the corresponding section of the homomorphism bundle $\mathrm{Hom}(\eta, \beta)$. Following Koschorke [77, Definition 1.4], we call u a $(t-1)$*-morphism* if for all $x \in M$ the rank of $u_x : \eta_x \to \beta_x$ is at least $t-1$. Equivalently, u is a $(t-1)$-morphism if the section s_u goes into

$$W^{t-1}(\eta, \beta) = \bigcup_{x \in M} W^{t-1}(\eta_x, \beta_x),$$

where $W^{t-1}(\eta_x, \beta_x)$ denotes the set of all linear maps $\eta_x \to \beta_x$ of rank at least $t-1$; $W^{t-1}(\eta, \beta)$ is an open subset in $\mathrm{Hom}(\eta, \beta)$. A $(t-1)$-morphism u is called a *non-degenerate* $(t-1)$*-morphism* if $s_u : M \to \mathrm{Hom}(\eta, \beta)$ is transverse to

$$A^{t-1}(\eta, \beta) = \bigcup_{x \in M} A^{t-1}(\eta_x, \beta_x),$$

where $A^{t-1}(\eta_x, \beta_x)$ denotes the set of all linear maps $\eta_x \to \beta_x$ of rank $t-1$; $A^{t-1}(\eta, \beta)$ is a closed smooth submanifold in $W^{t-1}(\eta, \beta)$.

Standard results on transversality imply the following density property. If we have a $(t-1)$-morphism $u : \eta \to \beta$, a closed subspace $L \subset M$ where s_u is already transverse to $A^{t-1}(\eta, \beta)$, and a neighbourhood V of $s_u(M)$ in $W^{t-1}(\eta, \beta)$, then there exists a non-degenerate $(t-1)$-morphism u' with $u'_{|L} = u_{|L}$ and $s_{u'}(M) \subset V$, and such that u and

u' are (linearly) homotopic through $(t-1)$-morphisms. There exists a non-degenerate 0-morphism $u : \eta \to \beta$, for any η and β, but in general it is *not* true that for any η and β (over any M^n) there exists a k-morphism if $k \geq 1$. Indeed (see [71]), over the Grassmann manifold $G_k(\mathbb{R}^{k+m})$ there is no 1-morphism $\gamma \to \gamma^\perp$ if k or m is even; and if both k and m are odd, then there is a 1-morphism, but no 2-morphism.

Now let M^n be a smooth closed connected manifold. If we have a monomorphism $u : \eta^k \to T(M)$, then we have the induced monomorphism $p^*(u) : p^*(\eta) \to p^*(T(M))$ over the (total space of the) *projectification* (also called *projectivization*) *bundle* $P(\eta) = G_1(\eta)$ (see 4.2). Indeed, for each $g \in P(\eta)$, we have $p^*(u)_g = u_{p(g)}$. Of course, since the canonical line bundle γ over $P(\eta)$ is contained in $p^*(\eta)$, we also have a monomorphism $\gamma \to p^*(T(M))$.

Naturally, we would like to find an obstruction to the existence of monomorphisms $\gamma \to p^*(T(M))$, which also is an obstruction to the existence of monomorphisms $u : \eta^k \to T(M^n)$. Let us take a non-degenerate 0-morphism $v : \gamma \to p^*(T(M))$. So $s_v(x) : \gamma_x \to p^*(T(M))_x$ is a monomorphism for all $x \in P(\eta) \setminus S$, where

$$S = \{x \in P(\eta); \operatorname{rank}(v_x) = 0\} = s_v^{-1}(A^0(\gamma, p^*(T(M)))).$$

Of course, from another point of view, $v(\gamma)$ defines a 1-distribution in $p^*(T(M))$ with the singularity S. It turns out (recall that we have now $k \geq 2$) that if the singularity S is nonempty (which we shall suppose because otherwise v simply would be a monomorphism), then it is "huge". Indeed, by elementary transversality theory (see, e.g., [57, Ch. 1, Theorem 3.3]), the codimension of the closed submanifold S in the $(n+k-1)$-dimensional manifold $P(\eta)$ is the same as the codimension of the $(n+k-1)$-dimensional submanifold $A^0(\gamma, p^*(T(M)))$ in the $(n+k-1+n)$-dimensional manifold $\operatorname{Hom}(\gamma, p^*(T(M)))$, hence we have $\dim(S) = k-1$. So S may be a union of circles, or a surface etc., depending on how big is k. Of course, v is a monomorphism over $P(\eta) \setminus S$.

The apparent disadvantage – that the singularity S is far from finite – converts into an advantage. Indeed, let $g : S \hookrightarrow P(\eta)$ be the inclusion. The normal bundle of S in $P(\eta)$ (see [77, 1.6]) is canonically isomorphic to the pullback $g^*(\operatorname{Hom}(\gamma, p^*(T(M))))$, and so we have an isomorphism

$$T(S) \oplus g^*(\operatorname{Hom}(\gamma, p^*(T(M)))) \cong g^*(T(P(\eta))),$$

hence

$$T(S) \oplus g^*(\operatorname{Hom}(\gamma, p^*(T(M)))) \cong g^*(p^*(T(M)) \oplus \operatorname{Hom}(\gamma, \gamma^\perp)).$$

By adding the trivial line bundle $\operatorname{Hom}(\gamma, \gamma) \cong \varepsilon^1$ to both sides, using the fact that $\gamma \oplus \gamma^\perp = p^*(\eta)$, we obtain the isomorphism

$$\bar{g} : T(S) \oplus g^*(\operatorname{Hom}(\gamma, p^*(T(M)))) \oplus \varepsilon^1 \cong g^*(\operatorname{Hom}(\gamma, p^*(\eta)) \oplus p^*(T(M))).$$

In other words, the stable normal bundle of S in $P(\eta)$ can be expressed as a pullback of the virtual vector bundle

$$\Phi = \operatorname{Hom}(\gamma, p^*(T(M))) - \operatorname{Hom}(\gamma, p^*(\eta)) - p^*(T(M))$$

over the manifold $P(\eta)$.

As a result, we obtain a well defined (i.e., independent of the choice of the non-degenerate 0-morphism $\gamma \to p^*(T(M))$) class $\omega(\eta, T(M)) = [P(\eta), g, \bar{g}]$ in the *normal bordism group* $\Omega_{k-1}(P(\eta); \Phi)$. The element $\omega(\eta, T(M))$ is an obstruction to the existence of vector bundle monomorphisms $\gamma \to p^*(T(M))$ and, consequently, also vector bundle monomorphisms $\eta \to T(M)$ or, in other words, to η defining a k-distribution on M. Similarly (see [79]), one can obtain an obstruction, $\omega^{\mathrm{st}}(\eta, T(M)) \in \Omega_{k-1}(\mathbb{R}P^\infty \times M; \Phi)$ (where $\mathbb{R}P^\infty$ = infinite dimensional real projective space) to η defining a stable k-distribution on M.

4.15 An example in the metastable range of dimensions

In the metastable range of dimensions, the obstructions ω and ω^{st} contain complete information. More precisely, Koschorke ([77, 3.7], [79, Theorem 1]) proved the following theorem. Let η be a k-dimensional vector bundle over a smooth, closed, connected manifold M^n. Suppose that $2k < n$. Then η defines a k-distribution or a stable k-distribution on M if and only if $\omega(\eta, T(M)) = 0$ or $\omega^{\mathrm{st}}(\eta, T(M)) = 0$, respectively.

Of course, even if $2k < n$, one highly nontrivial requirement remains: to be able to decide, in any particular case, whether the obstruction $\omega(\eta, T(M))$ (or $\omega^{\mathrm{st}}(\eta, T(M))$) vanishes. A substantial step towards this end is expressing the obstruction in terms of better known, and more easily calculable, invariants. As shown by Koschorke [79], this is, at least sometimes, manageable. Indeed, it is possible to pass from the obstruction $\omega(\eta, T(M))$, e.g., to the following result (a particular case of [79, Corollary 9]). Let M^n be a smooth closed connected manifold, with $6 < n \equiv 2 \pmod 4$, and let η^3 be a vector bundle over M. Assume that the cohomology class $w_1(M)w_1^2(\eta) + w_1(M)w_2(\eta) + w_1(\eta)w_2(\eta) + w_3(\eta) \in H^3(M; \mathbb{Z}_2)$ is nonzero. Then a vector bundle η defines a 3-distribution on M if and only if the Stiefel-Whitney classes $w_n(T(M) - \eta)$ and $w_{n-2}(T(M) - \eta)$, and the twisted Euler class $e(M) \in H^n(M; \mathbb{Z}_{w_1(M)})$ all vanish. For 3-distributions, see also [92].

We remark that, given a vector bundle η^k over a manifold M, the singularity approach can also be effective in counting the number of homotopy classes of vector bundle monomorphisms $\eta^k \to T(M)$ when such monomorphisms exist; examples can be found in [79].

4.16 The vector field problem, alias the existence question for trivial vector distributions, and its historical background

Now we specialize to the question of when a manifold admits a trivial vector distribution. As we shall see it is an important (and also very difficult) problem, not only as part of the general existence question for vector distributions, but also in its own right.

Let M^n be a smooth closed connected manifold, and let α^q be a real (smooth) vector bundle over M. Of course, without loss of generality we may (and we shall) assume that α is equipped with a Euclidean metric. Given an open subset $U \subset M$, we say that sections $s_1, \dots, s_t \in \Gamma(\alpha)$ are *independent on* U if their values $s_1(x), \dots, s_t(x) \in \alpha_x$ are linearly independent for each $x \in U$. [This extends the definition given for vector fields in the beginning of 4.1.] As is well known, there are t continuous sections of $\alpha_{|U}$ independent on U if and only if there are t smooth sections of $\alpha_{|U}$ independent on U.

Clearly, t sections of $\alpha_{|U}$ that are independent on U *span* a trivial t-dimensional subbundle, hence they determine a trivial t-distribution in $\alpha_{|U}$. Then the *span* of α, denoted by $\operatorname{span}(\alpha)$, is naturally defined to be the largest number of sections of α independent on M. The presence of the Euclidean metric on α implies that we have $\operatorname{span}(\alpha) \geq k$ if an only if $\alpha = \varepsilon^k \oplus \eta$ for some vector bundle η. In other words, we have $\operatorname{span}(\alpha) \geq k$ if and only if α admits a trivial k-distribution ε^k (and also its complementary $(q-k)$-distribution κ^{q-k}), hence if and only if there is a vector bundle monomorphism $\varepsilon^k \to \alpha$ (and the complementary monomorphism $\kappa^{q-k} \to \alpha$).

In particular, we define $\operatorname{span}(M) = \operatorname{span}(T(M))$. So to have $\operatorname{span}(M) \geq k$ is the same as to have a trivial k-distribution and the corresponding, in general nontrivial, normal $(n-k)$-distribution on M. Of course, equivalently, we have $\operatorname{span}(M) \geq k$ if and only if there exists a vector bundle monomorphism $\varepsilon^k \to T(M)$.

Recall that the *Lyusternik-Shnirel'man category*, cat(X), of a topological space X is (according to one of the two most used conventions), the least integer k (or ∞) such that X can be covered by k open subsets which are contractible in X. So, e.g., for the sphere S^n we have cat$(S^n) = 2$. For a vector bundle α over a closed manifold M (in particular, for the tangent bundle $T(M)$), the span of α (in particular, the span of M) and the Lyusternik-Shnirel'man category of M impose certain limitations on splittings of α (in particular, on vector distributions on M). Indeed, by Korbaš and Szűcs [73, Theorem 1.1], if α is a vector bundle over a closed manifold M, $k \geq$ cat(M), and α splits nontrivially as $\alpha \cong \alpha_1 \oplus \cdots \oplus \alpha_k$, then we have

$$1 \leq \min\{\dim(\alpha_i);\ i = 1, \ldots, k\} \leq \operatorname{span}(\alpha).$$

Other formulations of the vector field problem (already briefly mentioned in 4.2) are, e.g., the following.

(VFP1) Find $\operatorname{span}(M^n)$ for a given manifold M. In other words, find the maximum k $(0 \leq k \leq n)$ such that M admits a trivial k-distribution. Equivalently, find the maximum k such that there exists a vector bundle monomorphism $\varepsilon^k \to T(M)$.

(VFP2) For a given k, characterize in terms of computable invariants (characteristic classes, characteristic numbers, etc.) all those manifolds M, for which $\operatorname{span}(M) \geq k$, hence those M admitting a trivial k-distribution.

If $\operatorname{span}(M^n) \geq k$ for some $k \geq 1$, hence if we have a trivial k-distribution on M then, of course, we have a positive answer to the existence question of p-distributions (and also $(n-p)$-distributions) with $p \leq k$ on M. This is also interesting in the framework of foliation theory. Indeed, for instance, as proved by W. Thurston ([133, Corollary 1]) in 1974, every trivial k-distribution $(k \geq 2)$ on a closed manifold M is homotopic to the normal distribution of a smooth foliation of M.

Results on the vector field problem also turn out to be useful in the theory of singularities of smooth maps. Indeed, as proved by Y. Ando [3], if $\operatorname{span}(M^n) \geq p-1$ for some p such that $n \geq p \geq 2$, then there exists a fold map $M \to \mathbb{R}^p$ (recall that a *fold map* is a smooth map having only fold singularities). In addition to this, O. Saeki showed ([109]) that if $n-p+1$ is odd and there exists a fold map $M \to \mathbb{R}^p$, then $\operatorname{span}(T(M) \oplus \varepsilon^1) \geq p$ (whence M admits a stable trivial $(p-1)$-distribution).

Recall ([69, IV.4.11-12], [72]) that a vector bundle α is a p-*Clifford module* if there exist vector bundle automorphisms $e_i : \alpha \longrightarrow \alpha$, $i = 1, \ldots, p$, such that $e_i^2 = -\mathrm{id}$ and $e_i e_j + e_j e_i = 0$ for $i \neq j$. The vector field problem is also related to the question of whether or not the tangent bundle of a given orientable even-dimensional closed manifold M^n is a *p-Clifford module* for some $p \geq 1$. Indeed, to ask the question for $p = 1$ is the same as to ask if M^n admits an *almost complex structure*. It is clear that if the answer is positive and $\operatorname{span}(M) \geq 1$, then we also have $\operatorname{span}(M) \geq 2$ (so if $\operatorname{span}(M) = 1$, then we immediately know that M does not admit any almost complex structure). But it is also true (as an obvious special case of the author's [72, Theorem 3.2]) that if the tangent bundle of M is a p-Clifford module and $\operatorname{span}(M) \geq 1$, then $\operatorname{span}(M) \geq p + 1$.

For basic information on relations between the *vector field problem* and *generalized vector field problem* (find $\operatorname{span}(t\xi_p)$ for any t, where ξ_p is the Hopf line bundle over $\mathbb{R}P^p$), we refer to [75, pp. 92-93]. Since the 1960's, when the generalized vector field problem was posed by M. Atiyah, R. Bott, and A. Shapiro (in their seminal paper on Clifford modules, published in Topology in 1964), many authors published results on it. We name at least some: S. Gitler, K. Y. Lam, D. Davis, M. Mahowald, S. Feder, W. Iberkleid, T. Yoshida, T. Kobayashi.

The span is also an interesting *geometric* characteristic of a manifold. In particular, if $\operatorname{span}(M^n) = n$, then M admits a global parallel motion (in the obvious sense), and is therefore called *parallelizable*. For example, each Lie group is parallelizable (use the right multiplication to move a basis chosen for the tangent space at the identity element to obtain a basis for the tangent space at any point). Relations of the vector field problem to further areas (immersions, the Lyusternik-Shnirel'man category) will be shown in 5.10.

The vector field problem has a long history (see, e.g., [74], [75], [131]; here we give just a brief outline). Indeed, already in the 1880's H. Poincaré, in his works on curves defined by differential equations, studied singularities of vector fields. Among other results, he found that each vector field on the 2-dimensional sphere S^2 has somewhere a zero-point. Then, about 1910, L. E. J. Brouwer and J. Hadamard (independently) showed that $\operatorname{span}(S^{2t}) = 0$ and $\operatorname{span}(S^{2t+1}) \geq 1$ for any $t \geq 0$. Around 1920, by independent works of A. Hurwitz and J. Radon, it became known that, expressing any n as $(2a + 1) \cdot 2^{c+4d}$, with $a, c, d \geq 0, c \leq 3$, we have $\operatorname{span}(S^{n-1}) \geq \varrho(n) - 1$, where $\varrho(n) = 2^c + 8d$. The integer $\varrho(n)$ is often referred to as the *Hurwitz-Radon number*. Around 1925 H. Hopf answered the question of when a closed smooth connected manifold M has $\operatorname{span}(M) \geq 1$, namely if and only if $\chi(M) = 0$ (as we already have mentioned in 4.2). For an orientable manifold M^n, we have $\operatorname{span}(M) \geq n - 1$ if and only if $\operatorname{span}(M) \geq n$, and of course no non-orientable manifold can be parallelizable. So for closed manifolds of dimension ≤ 2, the vector field problem had a clear solution. About 10 years after the Hopf theorem, E. Stiefel proved that any orientable 3-dimensional manifold is parallelizable.

In 1958, M. Kervaire and J. Milnor, using results of R. Bott on the stable homotopy of the classical groups, solved the parallelizability question for spheres: only S^1, S^3, and S^7 are parallelizable. This immediately also solves the parallelizability question for real projective spaces: only $\mathbb{R}P^1$, $\mathbb{R}P^3$, and $\mathbb{R}P^7$ are parallelizable. Finally, in the early 1960's, the vector field problem for spheres was also completely solved: J. F. Adams in [2], using operations in K-theory, showed that the Hurwitz-Radon lower bound is also an upper bound. As a consequence, $\operatorname{span}(S^{n-1}) = \varrho(n) - 1$ for each $n \geq 2$. One can then show that also $\operatorname{span}(\mathbb{R}P^{n-1}) = \varrho(n) - 1$ for each $n \geq 2$. So, by the latter, we have a trivial $(\varrho(n) - 1)$-distribution on $\mathbb{R}P^{n-1}$. As proved by H. Glover, W. Homer, and R. Stong [41], its complementary (normal) distribution does not split, and if we have any distributions $\eta_1, \dots, \eta_t$ on $\mathbb{R}P^{n-1}$ such that $T\mathbb{R}P^{n-1} = \eta_1 \oplus \cdots \oplus \eta_t$, then $\max\{\dim(\eta_i)\} \geq n - \varrho(n)$.

Then in 1964, W. Sutherland proved that the Stiefel manifold $V_{n,r}$ for $r \geq 2$ is always parallelizable; later on others, e.g., D. Handel, K. Y. Lam, L. Smith (for $r \geq 3$) or P. Zvengrowski (for $r \geq 2$) gave other, more elementary proofs.

The spheres have many interesting properties. One of them is that $T(S^n) \oplus \varepsilon^1 = \varepsilon^{n+1}$. Manifolds M^n such that $TM^n \oplus \varepsilon^1 = \varepsilon^{n+1}$ are called *stably parallelizable* (or also π-*manifolds*). For these, the vector field problem was solved by G. Bredon and A. Kosinski in 1966. We remark that a great deal of interesting information on parallelizable or stably parallelizable manifolds can be found in M. Gromov's book [46].

Now let us look at the situation for those manifolds which are not necessarily stably parallelizable. If we are interested in concrete manifolds, then briefly said: apart from the trivial cases with $\chi(M) \neq 0$ (hence $\operatorname{span}(M) = 0$) and apart from what we already mentioned, in the 1960's and in the early 1970's remarkable results were achieved on *spherical space forms* S^n/G, where G is a finite subgroup of $O(n+1)$, acting freely on S^n. They can be found in papers by J. Becker, T. Yoshida, or also in the book by N. Mahammed, R. Piccinini, and U. Suter, Some applications of topological K-theory, North-Holland 1980.

The main research activity in the late 1960's and the early 1970's was concentrated on attempts to find theorems similar to the Hopf theorem, this time for two or more everywhere independent vector fields (i.e., (VFP2) for $k \geq 2$). In this context, several methods were developed which can be roughly characterized as adjustments or refinements of the methods used for studies of the general existence question for vector distributions.

4.17 Results on the vector field problem obtained by removing finite singularities

From what we have said above and the fact that the Gram-Schmidt orthonormalization process is continuous, it is clear that we have a trivial smooth k-distribution (spanned by an orthonormal k-frame $(v_1, \ldots, v_k)$ of smooth vector fields) if and only if there exists a smooth section of the associated Stiefel fibre bundle $V_k(T(M))$ over M. The fibre $V_k(T(M))_x$ over $x \in M$ is the Stiefel manifold $V_k(T(M)_x) \cong V_{n,k}$. To simplify matters, let us suppose that M^n is oriented and we have a trivial smooth k-distribution κ with a finite singularity $S = \{p_1, \ldots, p_t\}$ on M. Hence, equivalently, we have an orthonormal k-frame $(v_1, \ldots, v_k)$ of smooth vector fields on $M \setminus S$. Similarly to 4.9, we now take a triangulation on M^n such that each p_i lies in the interior of an n-simplex σ_i, and each σ_i contains at most one point of S. For each i, we choose an orientation preserving trivialization $T(M)|_{\sigma_i} \cong \sigma_i \times \mathbb{R}^n$ in such a way that it is compatible with the standard metric on $\sigma_i \times \mathbb{R}^n$. The assignment $x \mapsto (v_1(x), \ldots, v_k(x)) \in V_{n,k}$, for any $x \in S^{n-1} \cong \dot{\sigma}_i$, defines an element $\bar{o}_i \in \pi_{n-1}(V_{n,k})$, called the index of κ, or of $(v_1, \ldots, v_k)$, at $p_i \in S$. As we know, $\bar{o}_i$ is zero precisely when it is possible, in a small neighbourhood of p_i, to deform κ, or the fields $v_1, \ldots, v_k$, so that the singularity disappears. Then as a global obstruction one takes $\bar{\mathcal{O}}(\kappa) = \bar{\mathcal{O}}(v_1, \ldots, v_k) = \sum_{i=1}^{t} \bar{o}_i \in \pi_{n-1}(V_{n,k})$. This is independent of the triangulation but, in general, it can be changed, if we reverse the orientation on M. Obstruction theory ($\bar{\mathcal{O}}(v_1, \ldots, v_k)$ corresponds to the Steenrod obstruction cohomology class, by Poincaré duality) implies that $\bar{\mathcal{O}}(\kappa) = \bar{\mathcal{O}}(v_1, \ldots, v_k) = 0$ if and only if there are k vector fields $\widetilde{v}_1, \ldots, \widetilde{v}_k$ independent on M so that $(\widetilde{v}_1, \ldots, \widetilde{v}_k)_{|M_{(n-2)}} = (v_1, \ldots, v_k)_{|M_{(n-2)}}$ hence if and only if there exists an oriented trivial k-distribution $\widetilde{\kappa}$ on M such that $\widetilde{\kappa}_{|M_{(n-2)}} = \kappa_{|M_{(n-2)}}$.

For example, if M is $(k-1)$-connected, then (by obstruction theory; recall that $V_{n,k}$

is $(n-k-1)$-connected) there exists an orthonormal k-frame of vector fields with a finite singularity on M; let us denote it by $(v_1, \ldots, v_k)$. In this situation, $\bar{\mathcal{O}}(v_1, \ldots, v_k)$ is a *primary* obstruction, hence independent of the choice of $(v_1, \ldots, v_k)$. So in this case the global obstruction depends only on the manifold M. In such a case, we can define $\bar{\mathcal{O}}_k(M) = \bar{\mathcal{O}}(v_1, \ldots, v_k)$, and then we have

$$\mathrm{span}(M) \geq k \text{ if and only if } \bar{\mathcal{O}}_k(M) = 0.$$

In particular, this applies when $k = 1$ and M is connected. Then it turns out that $\bar{\mathcal{O}}_1(M) = \chi(M)$, and we obtain the Hopf theorem already mentioned above.

In general, if M is not $(k-1)$-connected, but still has a k-field $(v_1, \ldots, v_k)$ with a finite singularity, then $\bar{\mathcal{O}}(v_1, \ldots, v_k)$ can depend on $(v_1, \ldots, v_k)$, and so our interest in global obstructions can lead us to a dead end. Nevertheless, even in such situations, the dead end is not always inevitable. Indeed, for instance (see [6]), for an oriented manifold M and an orthonormal 2-frame (v_1, v_2) of vector fields with a finite singularity on M, we *a priori* do not know if $\bar{\mathcal{O}}(v_1, v_2)$ depends or does not depend on (v_1, v_2). However, in 1958, Hirzebruch and Hopf [58](see also [89]) for 4-dimensional and, in the 1960's, M. Atiyah, D. Frank, and E. Thomas, for higher dimensional oriented manifolds, have shown that this global obstruction depends really only on M (for orientable manifolds of dimension congruent to 0 mod 4, a definite choice of orientation is made once and for all).

In calculations of global obstructions, another interpretation of the vector field problem frequently plays a rôle. To outline it, we recall that we have a natural map (which can be considered as a fibration with fibre $V_{n,k}$) $\pi : BO(n-k) \to BO(n)$, such that $\pi^*(\gamma^n) \cong \gamma^{n-k} \oplus \varepsilon^k$, where $BO(t)$ is the classifying space for t-dimensional vector bundles and γ^t is the classifying vector bundle. Let $f : M \to BO(n)$ be a classifying map of the tangent bundle $T(M)$. Then basic properties of classifying spaces and classifying maps of vector bundles readily imply that the existence of a trivial k-distribution on M^n (hence $\mathrm{span}(M) \geq k$) is equivalent to the existence of a map $\bar{f} : M \to BO(n-k)$ (called a *homotopy lift* of f) such that the maps $\pi \circ \bar{f}$ and f are homotopic ($\pi \circ \bar{f} \simeq f$). Of course, if M is orientable, we can use here $BSO(t)$ instead of $BO(t)$. So the vector field problem can be understood also as a lifting problem (up to homotopy), and can be attacked, e.g., by means of modified Postnikov resolutions of the fibration $\pi : BO(n-k) \to BO(n)$ (or $\tilde{\pi} : BSO(n-k) \to BSO(n)$, if M is orientable). The modified Postnikov resolutions (or towers; see [87]) have been used in many successful calculations of global obstructions for trivial k-distributions with finite singularities. From the rich literature, we mention at least [131], [103], [96], [22].

As a combination of various results, we obtain the following table (Table 1) of necessary and sufficient conditions for $\mathrm{span}(M) \geq 2$ if M^n $(n \geq 4)$ is closed and oriented. In this and in other similar tables, conditions are stated in the form "condition (reference to proof)".

In addition to this ([58]), an oriented closed smooth manifold M^4 has span at least three (equivalently: M^4 is parallelizable) if and only if $\chi(M) = 0$, $w_2(M) = 0$ and $\sigma(M) = 0$.

Also for oriented closed n-dimensional manifolds with $n = 5, 6, 7$, in addition to the information contained in the table, quite complete results on the vector field problem are known. E.g., we have the following (see mainly E. Thomas, Vector fields on low dimensional manifolds, Math. Z., 1968).

$n = 5$**:** Suppose that M^5 is oriented, $H^4(M^5; \mathbb{Z})$ has no element of order 2, $w_4(M) = 0$,

M oriented, $\dim(M) = n$	Necessary and sufficient conditions for $\operatorname{span}(M) \geq 2$
$n = 4t+1, t \geq 1$	$w_{n-1}(M) = 0$ and $R(M) = 0$ ([12])
$n = 4t+2, t \geq 2$	$\chi(M) = 0$ ([129])
$n = 4t+3, t \geq 1$	none ([131])
$n = 4t, t \geq 1$	$\chi(M) = 0$ and $\sigma(M) \equiv 0 \pmod 4$ ([12], for $t = 1$ [89])

Table 1: M oriented, $\operatorname{span}(M) \geq 2$

and $R(M) = 0$. Then $\operatorname{span}(M) \geq 3$ if and only if there is a class $x \in H^2(M;\mathbb{Z})$ such that $x^2 = p_1(M)$ and $x \mod 2 = w_2(M)$. Suppose that for an oriented manifold N^5 the group $H^4(N;\mathbb{Z})$ has no element of order 2. Then N is parallelizable if and only if $w_2(N) = 0$, $p_1(N) = 0$, and $\hat{\chi}(N) = 0$.

$n = 6$**:** Suppose that M^6 is oriented and $\chi(M^6) = 0$. Then $\operatorname{span}(M) \geq 3$ if and only if $w_2^2(M) = 0$. If $H^4(M;\mathbb{Z})$ has no element of order 2, then $\operatorname{span}(M) \geq 4$ if and only if there is some $x \in H^2(M;\mathbb{Z})$ such that $x^2 = p_1(M)$ (the first Pontrjagin class) and $x \mod 2 = w_2(M)$. Suppose that N^6 is an oriented manifold such that $\chi(N) = 0$ and $H^4(N;\mathbb{Z})$ has no element of order 2. Then N is parallelizable if and only if $w_2(N) = 0$ and $p_1(N) = 0$.

$n = 7$**:** Suppose that M^7 is oriented. Then $\operatorname{span}(M) \geq 3$. If $H^4(M;\mathbb{Z})$ has no element of order 2 and $w_2(M) = 0$, then $\operatorname{span}(M) \geq 4$. If $H^4(M;\mathbb{Z})$ has no element of order 2, then $\operatorname{span}(M) \geq 5$ if and only if there is some $x \in H^2(M;\mathbb{Z})$ such that $x^2 = p_1(M)$ and $x \mod 2 = w_2(M)$. If $H^4(M;\mathbb{Z})$ has no element of order 2, then M is parallelizable if and only if $w_2(M) = 0$ and $p_1(M) = 0$.

For trivial 3-distributions with finite singularities, nearly definitive results were achieved by Atiyah and Dupont ([12], [34]) and by Crabb and Steer [28]. The global invariants of manifolds which appear in their generalizations of the Hopf theorem are combinations of the Euler characteristic χ, the signature σ and the real Kervaire semi-characteristic R. But $\chi(M)$ is the analytical index of an elliptic differential operator on the manifold M, $\sigma(M)$ has a similar interpretation, and $R(M)$ is a certain mod 2 index of a skew-adjoint elliptic operator. The basic idea is to pass from elliptic differential operators to their symbols which lie in certain K-groups. Then the Atiyah-Singer index theorem makes it possible to express χ, σ or R in terms of K-theory from these symbols. If we have a trivial k-distribution η with a finite singularity $S = \{p_1, \dots, p_m\}$ on a closed manifold M^n, then the symbols of the operators mentioned above are divisible by some 2^ν on $M \setminus S$. So one may expect to obtain relative K-theory characteristic classes modulo 2^ν for the pair $(X, X - \cup_i\{p_i\})$. Then for each i we get a local characteristic number modulo 2^ν, which is some function of the local obstruction. At the same time, the sum of these local characteristic numbers modulo 2^ν gives the global index of the original operator, and this can then imply a result on the global obstruction for η. Of course, K-theory information is converted into information on global obstructions using the homomorphisms $\theta^s : \pi_{n-1}(V_{n,k}) \to KR^s(\mathbb{R}P^{k+s-1}, \mathbb{R}P^{s-1})$ already mentioned in 4.10 (recall that these are isomorphisms if $k \leq 3 \leq n - k$).

Dupont in [34] extended the results achieved in [12]. More precisely, he studied obstructions for trivial 3-distributions, using means similar to those used by Atiyah and himself for trivial 2-distributions in [12]. Combining various results (see mainly [34]), one obtains the following table (Table 2) of necessary and sufficient conditions for $\mathrm{span}(M^n) \geq 3$ ($n \geq 7$). In this, the closed manifold M^n is supposed to be oriented. For M^n with $n = 4t+1$ it is additionally assumed that the homology group $H_1(M;\mathbb{Z})$ has no 2-torsion; for such manifolds, $L_t(p_1(M),\ldots,p_t(M)) \in H^{4t}(M;\mathbb{Q})$ denotes the tth Hirzebruch polynomial in the Pontrjagin classes $p_1(M),\ldots,p_t(M)$.

M oriented, $\dim(M) = n$	Necessary and sufficient conditions for $\mathrm{span}(M) \geq 3$
$n = 4t+1$, $t \geq 2$	Assume that $H_1(M;\mathbb{Z})$ has no 2-torsion. $\beta^* w_{n-3}(M) = 0$, $R(M) = 0$, and $L_t(p_1(M),\ldots,p_t(M)) \mod 4 = 0$ ([34])
$n = 4t+2$, $t \geq 2$	$\chi(M) = 0$ and $w_{n-2}(M) = 0$ ([12])
$n = 4t+3$, $t \geq 1$	none ([12])
$n = 4t$, $t \geq 2$	$\chi(M) = 0$, $w_{n-2}(M) = 0$, and $\sigma(M) \equiv 0 \pmod 8$ ([12])

Table 2: M oriented, $\mathrm{span}(M) \geq 3$

We remark that a similar (slightly more extensive) table was compiled by D. Randall in [105].

Also, for trivial k-distributions with finite singularities on a nonorientable manifold, one can study the question of when they can be deformed to k-distributions (for details, see [77, §16], [105], [91]). From recent results based on the singularity approach, we mention at least the paper [22]. N. S. Cardim, M.H.P.L. Mello, D. Randall, and M.O.M. Silva use the intrinsic join product to express the global obstruction for a trivial k-distribution with a finite singularity defined on the total space of a smooth fibre bundle as the product of the obstructions for trivial k-distributions with finite singularities given on the fibre and the base space. They calculate global obstructions for trivial k-distributions with finite singularities on a smooth closed manifold M^n in terms of generators of the homotopy groups $\pi_{n-1}(V_{n,k})$ for $2 \leq k \leq 4$ in situations where the obstructions depend only on the oriented homotopy type of M^n. As applications, the authors prove several sufficient conditions for the existence of trivial 2-distributions or 4-distributions.

4.18 Results on the vector field problem achieved by other approaches

In [33, II.6], J.-P. Dax defined and studied for any vector bundle α^r over a finite CW-complex X a *normal cobordism Euler class* of α^r, $\mathrm{cobeul}(\alpha^r) \in \Omega^r(X;(\alpha))$. Here $\Omega^r(X;(\alpha))$ is a normal cobordism group (this abelian group can be identified with a limit of certain homotopy groups; see [33, II.1]) of the space X, in dimension r, with coefficients in the stable vector bundle (α) represented by α^r. The class $\mathrm{cobeul}(\alpha^r)$ is the only obstruction to $\mathrm{span}(\alpha^r) \geq 1$ if $\dim(X) \leq 2r-3$. Similarly, obstructions to the existence of a nowhere vanishing section of a vector bundle were also studied, e.g., by Hatcher and Quinn [53, 4.3] or Crabb [26, 2.4], [27].

Now let ξ_p be the Hopf line bundle over $\mathbb{R}P^p$. Combining (particular cases of) results due to Dax (see [74, p. 6]), we obtain the following criterion: Let us suppose that $1 \leq$

$k < \frac{1}{2}n$. Then a smooth closed connected manifold M^n has its span at least k if and only if $\mathrm{cobeul}(T(M)\widehat{\otimes}\xi_{k-1}) = 0 \in \Omega^n(M \times \mathbb{R}P^{k-1}; (T(M)\hat{\otimes}\xi_{k-1}))$. So, in the metastable range $1 \leq k < \frac{1}{2}n$, the vector field problem can also be interpreted as a task to find necessary and sufficient conditions for $\mathrm{cobeul}(T(M)\widehat{\otimes}\xi_{k-1}) = 0$, in terms of computable invariants. It seems that Dax did not work on finding such invariants. But recall (see 4.15) that Koschorke's $\omega(\varepsilon^k, T(M))$, in the range $1 \leq k < \frac{1}{2}n$, also is a complete obstruction to $\mathrm{span}(M) \geq k$. Koschorke's approach, based on a thorough study of $\omega(\varepsilon^k, T(M))$, really leads to expressing it in terms of better known invariants, at least for $k \leq 4$.

Besides obtaining new theorems on the vector field problem, this approach frequently enables one to prove again or to improve results obtained earlier in a different way. For instance, here is one of Koschorke's theorems ([77, 14.12]) which - under the assumption of orientability – was previously derived by Atiyah and Dupont in [12]: Let M^n be a smooth closed connected manifold (orientable or not), $n > 6$, $n \equiv 2 \pmod 4$. Then $\mathrm{span}(M) \geq 3$ if and only if $\chi(M) = 0$ and $w_{n-2}(M) = 0$.

From Randall's [104] and [105], we have the following tables (Tables 3 and 4) of necessary and sufficient conditions for $\mathrm{span}(M^n) \geq 2$ and $\mathrm{span}(M^n) \geq 3$ if M^n is closed and nonorientable.

There are also several results on trivial 4-distributions. We give just one of them, as an example ([77, 15.18]). Let M^n be a closed connected orientable smooth manifold, $n = 4t+1$, $t \geq 2$. Assume that the Steenrod square homomorphism $\mathrm{Sq}^1 : H^1(M;\mathbb{Z}_2) \to H^2(M;\mathbb{Z}_2)$ is injective and its image does not contain $w_2(M)$. Supposing, in addition, that $w_{n-1}(M) = 0$, $w_{n-3}(M) = 0$ and $\mathrm{Sq}^2 : H^2(M;\mathbb{Z}_2) \to H^4(M;\mathbb{Z}_2)$ is injective, then we have $\mathrm{span}(M) \geq 4$ if and only if the real Kervaire semicharacteristic $R(M)$ vanishes.

Only few theorems on the existence of trivial k-distributions for $k \geq 5$ have been found up to now. For instance, there are several results for $k = 7, 8, 9$, mostly achieved by T. B. Ng, mainly by using modified Postnikov resolutions. Roughly, they apply to manifolds satisfying some connectivity conditions, with dimension restricted by certain congruences (see, e.g., Ng's paper published in Topology Appl. in 1994).

H. Glover and G. Mislin in [40] observe that the work by W. Sutherland (Proc. London Math. Soc. in 1965), R. Benlian and J. Wagoner (C. R. Acad. Sci. Paris Sér. A-B in 1967), and J. Dupont (Math. Scandinavica in 1970) implies that if M^n and N^n are homotopy equivalent closed connected smooth manifolds and if $k < \frac{n}{2}$, then M admits a trivial k-distribution if and only if N does. In [40], they generalize this result using the techniques of localization, at the prime 2, of homotopy types. The existence of trivial k-distributions

M nonorientable, $\dim(M) = n$	Necessary and sufficient conditions for $\mathrm{span}(M) \geq 2$
$n = 3$	$w_1^2(M) = 0$ ([104])
$n = 4t+1, t \geq 1$	$w_{n-1}(M) = 0$ if $w_1^2(M) \neq 0$ ([77])
$n = 4t+1, t \geq 1$	$w_{n-1}(M) = 0$ and $R(M) = 0$ if $w_1^2(M) = 0$ ([12])
$n = 4t+3, t \geq 1$	$w_{n-1}(M) = 0$ if $w_1^2(M) \neq 0$ ([77])
$n = 4t+3, t \geq 1$	none if $w_1^2(M) = 0$ ([104])
$n = 2t, t \geq 2$	$\chi(M) = 0$ and $\beta^* w_{n-2}(M) = 0$ ([77], [101], [104])

Table 3: M nonorientable, $\mathrm{span}(M) \geq 2$

M nonorientable, $\dim(M)=n$	Necessary and sufficient conditions for $\operatorname{span}(M)\geq 3$
$n=4$	$\chi(M)=0$ and $w_2(M)=0$ ([105])
$n=5$	Assume that $H_1(M;\mathbb{Z}_{w_1(M)})$ has no 2-torsion and $w_1^2(M)=0$. There is a class $z\in H^2(M;\mathbb{Z}_{w_1(M)})$ such that $z^2=p_1(M)$ and $z \mod 2=w_2(M)$, and $R(M)=0$ ([105])
$n=4t+2$, $t\geq 2$	$\chi(M)=0$ and $w_{n-2}(M)=0$ ([77])
$n=4t+3$, $t\geq 1$	$\beta^*w_{n-3}(M)=0$ ([105])
$n=4t$, $t\geq 2$	Assume $w_1^2(M)$ generates the kernel of $G: w_1(M)H^1(M;\mathbb{Z}_2)\to H^3(M;\mathbb{Z}_2)$ given by $G(w_1(M)y)=w_1^2(M)y+w_1(M)y^2$. $\chi(M)=0$ and $w_{n-2}(M)=0$ ([105])
$n=4t+1$, $t\geq 2$	Assume that $w_1^2(M)=0$ and $w_1(M)$ generates the kernel of $g: H^1(M;\mathbb{Z}_2)\to H^2(M;\mathbb{Z}_2)$ given by $g(y)=y^2+w_1(M)y$. $\beta^*w_{n-3}(M)=0$, $w_{n-1}(M)=0$, and $R(M)=0$ ([77])
$n=4t+1$, $t\geq 2$	Assume $w_1(M)$ generates the kernel of g given above and $w_1^2(M)$ is not in the image of g. $\beta^*w_{n-3}(M)=0$ and $w_{n-1}(M)=0$ ([77])

Table 4: M nonorientable, $\operatorname{span}(M)\geq 3$

on M^n has also been studied using cobordism theory and the twisted signatures, e.g., by M. Bendersky (Math. Z. in 1989). For more references, see [74].

We have already considered parallelizability versus stable parallelizability. More generally, besides the span of a manifold one can define its *stable span* (in such a way that if a manifold M^n has its stable span n, then it is stably parallelizable). We define

$$\text{stable span}(M)=\operatorname{span}(TM\oplus\varepsilon^r)-r=\operatorname{span}(TM\oplus\varepsilon^1)-1\ (r\geq 1).$$

Of course, we have $\text{stable span}(M^n)\geq k$ if and only if the manifold M admits a trivial stable k-distribution, hence if and only if the geometric dimension of $T(M)$ does not exceed $n-k$ (we remark that there are still other equivalent characterizations).

There can be a big difference between $\operatorname{span}(M)$ and $\text{stable span}(M)$ (e.g., $\operatorname{span}(S^2)=0$, while $\text{stable span}(S^2)=2$), but these two numbers also can coincide. For instance, if M^n is a smooth closed connected manifold with n even and $\chi(M)=0$, then $\text{stable span}(M)=\operatorname{span}(M)$ (while $\operatorname{span}(M)=0$ if $\chi(M)\neq 0$). For more results of this type, see [77, §20]. So another way of attacking the vector field problem is to look for the stable span and then to try to show that it is in fact the span.

One can try to understand the relation between stable span and span also using a homotopy method. Let us suppose that the manifold M is odd-dimensional (if $\dim\ M$ is even, the matter is clear). Then by a special case of a theorem of I. James and E. Thomas [63], there are one or two isomorphism classes of n-dimensional vector bundles over M, which are stably isomorphic to the tangent bundle $T(M)$. If there are two, W. Sutherland [125] defines for any n-plane bundle α stably equivalent to $T(M)$ a number $b_\beta(\alpha)\in\mathbb{Z}_2$, called the *Browder-Dupont invariant*; b_β distinguishes between those two classes of n-plane bundles stably isomorphic to $T(M)$, and $b_\beta(T(M)) = \hat{\chi}_2(M)$. We observe that the task of computing $b_\beta(\alpha)$ is in general very difficult.

A basic fact, needed when one starts to think about the problem of stable span, is ([74, 2.2]) that $\text{stable span}(M)$ is positive if and only if the Euler-Poincaré characteristic $\chi(M)$ is even. Of course, we always have the inequality

$$\text{stable span}(M) \geq \text{span}(M).$$

There are various sources of upper bounds for the stable span, hence also for the span. Among them historically the first (but certainly not old-fashioned) are characteristic classes (see [94]). For instance, if the Stiefel-Whitney class $w_t(M^n) \in H^t(M;\mathbb{Z}_2)$ does not vanish, then we know that $\text{stable span}(M) \leq n - t$. For applications of this see, e.g., [70], [74], [75] (and several papers cited therein).

The following theorem (a possible source of upper bounds for the stable span) is due to K. H. Mayer [90] (see also [74]); he derived it using the Atiyah-Singer index formula. Let M^n be an oriented closed connected smooth manifold with $n \equiv 0 \pmod 4$. If $\text{stable span}(M) \geq r$, then the signature $\sigma(M)$ is divisible by b_r, where $b_{r+8} = 16 \cdot b_r$ and b_r is given by the following table.

r	1	2	3	4	5	6	7	8
b_r	2	4	8	16	16	16	16	32

For other general results on the existence of (stable) distributions, implied by the Atiyah-Singer index theorem see, e.g., [11] or B. Lawson and M.-L. Michelsohn's [85], [86]. Interesting theorems on (stable) distributions can also be found (or derived from) [28], [27]. In [66], B. Junod and U. Suter use the Atiyah γ-operations in KU-theory to prove results on the vector field problem for product manifolds. There are also results on the vector field problem in equivariant contexts (e.g., K. Komiya, Proc. Japan Acad., Ser. A in 1978, U. Namboodiri, Trans. Amer. Math. Soc. in 1983, M. Obiedat, Topology Appl. in 2006).

4.19 On the vector field problem on specific manifolds after 1975

After 1975, much work was dedicated to showing that parallelizability (or even stable parallelizability) is very rare among the frequently occurring families of manifolds in geometry and topology. We now briefly comment on some of this work.

As proved by T. Yoshida in 1975 ([141]) among the real Grassmann manifolds $G_k(\mathbb{R}^n)$, only the following are parallelizable: $G_2(\mathbb{R}^1) = \mathbb{R}P^1$, $\mathbb{R}P^3 = G_1(\mathbb{R}^4) = G_3(\mathbb{R}^4)$, and $\mathbb{R}P^7 = G_1(\mathbb{R}^8) = G_7(\mathbb{R}^8)$. None of the others is even stably parallelizable. Somewhat later (see [74], [75]) different proofs were given by Hiller and Stong, Bartík and Korbaš, Trew and Zvengrowski.

Among the oriented Grassmann manifolds, only $S^1 = \tilde{G}_1(\mathbb{R}^2)$, $S^3 = \tilde{G}_1(\mathbb{R}^4) = \tilde{G}_3(\mathbb{R}^4)$, $S^7 = \tilde{G}_1(\mathbb{R}^8) = \tilde{G}_7(\mathbb{R}^8)$, and $\tilde{G}_3(\mathbb{R}^6)$ are parallelizable. Of the remaining ones only $\tilde{G}_2(\mathbb{R}^4)$ is stably parallelizable. This theorem was first stated and partially proved by I. Miatello and R. Miatello [93], and completely proved in the dissertation of P. Sankaran (University of Calgary, 1985). In [93], the authors also proved the following wider result: For $q > 2$ the *oriented flag manifold* $\tilde{F}(n_1, \dots, n_q) = SO(n_1 + \cdots + n_q)/SO(n_1) \times \cdots \times$

$SO(n_q)$, where without loss of generality we assume $n_1 \geq n_2 \geq \dots$, is parallelizable if and only if $n_1 = n_2 = \dots = 1$, or $n_1 = n_2 = \dots = 3$, or $\{n_1, n_2, \dots\} = \{3, 1\}$, or $\{n_1, n_2, \dots\} = \{2, 1\}$ with at least two n_i equal to one in the latter case. Of the rest only $\tilde{F}(2, \dots, 2)$ and $\tilde{F}(2, \dots, 2, 1)$ are stably parallelizable.

For results on the parallelizability question for the *partially oriented flag manifolds* $O(n_1 + \dots + n_s)/SO(n_1) \times \dots \times SO(n_r) \times O(n_{r+1}) \times \dots \times O(n_s)$, see Sankaran and Zvengowski's [111].

The (stable) parallelizability problem for the flag manifolds $F(n_1, \dots, n_q)$ (see 4.1) was first solved by Korbaš ([70]). Somewhat later, another proof was given by Sankaran and Zvengrowski. The result is that for $q > 2$ only the flag manifold $F(1, 1, \dots, 1)$ is parallelizable, and the other manifolds $F(n_1, \dots, n_q)$ are not stably parallelizable.

In 1986, E. Antoniano, S. Gitler, J. Ucci, and P. Zvengrowski in [5] almost completely (with only one exception) solved the parallelizability question for the projective Stiefel manifold $X_{n,k}$, obtained from the ordinary Stiefel manifold $V_{n,k}$ $(k < n)$ by identifying each $(v_1, \dots, v_k) \in V_{n,k}$ with $(-v_1, \dots, -v_k)$. Their result is that $X_{n,k}$ is parallelizable if (n, k) equals $(n, n-1),\ (2m, 2m-2),\ (16, 8)$, and if $n = 2,\ 4$, or 8, and none of the remaining ones are stably parallelizable, with the possible exception of the undecided $X_{12,8}$.

The complex projective Stiefel manifold $PW_{n,k}$ is obtained from the complex Stiefel manifold $W_{n,k}$ of orthonormal k-frames in $\mathbb{C}^n$ by identifying any frame $(v_1, \dots, v_k)$ with the frame $(zv_1, \dots, zv_k)$ for any z in the circle $U(1)$. In 2000, L. Astey, S. Gitler, E. Micha, and G. Pastor ([8]) settled the question of parallelizability of the manifolds $PW_{n,k}$ in the following way: If $k < n - 1$, then $PW_{n,k}$ is not stably parallelizable; $PW_{n,n-1}$ is parallelizable, except $PW_{2,1} = S^2$; and $PW_{n,n}$ (that is, the projective unitary group) is parallelizable. In 2003, Astey, Micha, and Pastor also solved the parallelizability question for generalized complex projective Stiefel manifolds (for the result, see [7]). Parallelizability is a rare phenomenon also for many other families of manifolds; see W. Singhof [116], Singhof and Wemmer [117].

For further results on the span of concrete frequently used (families of) manifolds, we refer to [74], [75], [142]. Recent papers by D. Ajayi and S. Ilori (2002) and the author (2004) bring new estimates of the span or complete solutions of the span-problem for some families of the flag manifolds $O(n)/O(1) \times O(1) \times O(n-2)$. Many authors also have achieved interesting results on the geometric dimension of vector bundles over concrete manifolds, mainly over real projective spaces (e.g., S. Gitler, M. Mahowald, J. F. Adams, D. Davis, K. Y. Lam, D. Randall, M. Bendersky).

4.20 Remarks on complex vector distributions

In the above, we have restricted our exposition to real subbundles of tangent bundles or more general vector bundles. But it is also of much interest to study complex vector distributions, hence complex subbundles of complex vector bundles. For instance, in 1967, Thomas [130] defined the *complex span* of a complex vector bundle to be its maximum number of everywhere complex linearly independent sections of it. In particular, the *complex span* of an almost complex manifold M^{2n} is defined to be the complex span of the complex n-dimensional vector bundle whose underlying real bundle is (isomorphic to) the tangent bundle $T(M)$. Of course, then the real span of M is at least twice the complex span. As an example, we quote [130, Corollary 1.6]: Let M^{2n} $(n \geq 6)$ be a smooth closed connected almost complex manifold, with an almost complex structure ω, and suppose that M is 3-connected. Then the complex span of M is at least three if and only if $\chi(M) = 0$ and the Chern class $c_{n-2}(\omega) \in H^{2n-4}(M;\mathbb{Z})$ is zero. In 1999, Koschorke in [80], for two complex vector bundles α and β over a closed connected smooth manifold M^n considered the problem of comparing complex and real monomorphisms from α to β. Two such monomorphisms are considered as equal if they are regularly homotopic. Using the singularity approach, he shows that the existence and classification results in the complex and in the real setting are related by transition homomorphisms of normal bordism groups; these homomorphisms fit into a long exact sequence of Gysin type. Under suitable conditions, explicit calculations, giving results in terms of characteristic classes, are possible. For example, if $\dim_{\mathbb{R}}(\alpha) = 2$, $\dim_{\mathbb{R}}(\beta) = n$, and M is nonorientable, then Koschorke proves that the following statements are equivalent: (a) there exists a complex monomorphism $u : \alpha \to \beta$; (b) there exists a real monomorphism $u : \alpha \to \beta$; and (c) the Stiefel-Whitney class $w_n(\beta - \alpha) = \sum_{i\geq 0} w_2(\alpha)^i w_{n-2i}(\beta)$ vanishes. Recently H. Jacobowitz and G. Mendoza (see [61], [62]) studied smooth complex subbundles of the complexification of the tangent bundle of a smooth manifold. They mainly concentrate on properties of interest in the theory of partial differential equations.

5 Immersions and embeddings of manifolds in Euclidean spaces

5.1 Preliminaries

A smooth map $f : N^t \to M^n$ $(t \leq n)$ is an *immersion*, if the map f is *regular*, i.e., if the differential at y, $df : N_y \to M_{f(y)}$, is injective for each $y \in N$. For any immersion $f : N^t \to M^n$ (where N is closed, M is supposed to be a Riemannian manifold and it need not be closed), the pullback of the tangent bundle $f^*(T(M))$ splits as a Whitney sum of two subbundles. More precisely, we have $f^*(T(M)) \cong T(N) \oplus \nu_f$; the vector bundle ν_f is called the *normal bundle* of the immersion f. In particular, for an immersion $g : N^t \to \mathbb{R}^n$ we have $\varepsilon^n \cong T(N) \oplus \nu_g$. So in the reduced KO-group ([69]), ν_g represents the (additive) inverse to the element represented by $T(N)$. As a consequence, the *stable equivalence class* of the normal bundle ν_g, denoted by (ν_g), depends just on N, and is the same for all immersions $N^t \to \mathbb{R}^n$. We speak about this stable equivalence class, but sometimes also about any of its representatives, as *the* stable normal bundle of (any) immersion $N^t \to \mathbb{R}^n$.

Of course, basic properties of the Stiefel-Whitney characteristic classes (which also depend just on stable equivalence classes of vector bundles) immediately imply that a necessary condition for the existence of an immersion of N^t in $\mathbb{R}^n$ is that the *dual* (also called *normal*) *Stiefel-Whitney classes* $\bar{w}_i(N)$ vanish for all $i > n - t$. Another necessary condition is provided by the Atiyah γ-operations in KO-theory ([10]).

An immersion is allowed to have self-intersections; at the same time, any immersion is locally injective. If f is an injective immersion which maps N homeomorphically onto $f(N)$ (with the topology on $f(N)$ induced from M), then f is called an *embedding*. Note that if N is compact, then each injective immersion $N \to M$ is an embedding. If there is an embedding of a closed manifold N^t in $\mathbb{R}^n$, then also a tubular neighbourhood of N can be embedded in $\mathbb{R}^n$. Using this, one can show that a necessary condition for the existence of an embedding of N^t in $\mathbb{R}^n$ is that the dual Stiefel-Whitney classes $\bar{w}_i(N)$ vanish for all $i \geq n - t$ (e.g., [60, Ch. 17, 10.2]).

A smooth manifold N is a *submanifold* of a smooth manifold M if there is an injective immersion $\iota : N \to M$. Formally, one speaks about the submanifold (N, ι); if ι is an embedding, then we call N (more precisely, the pair (N, ι)) an *embedded submanifold* in M. A well known example: there is an immersion of the Klein bottle in $\mathbb{R}^3$, but there is no embedding of this surface in $\mathbb{R}^3$, hence it cannot be realized as a submanifold in $\mathbb{R}^3$.

Roughly speaking, it is easier to study immersions, characterized by just a local property (of having the differential injective at each point), than to study embeddings, characterized not only by a local property (the same as for immersions), but also by a global property (of being homeomorphisms onto the image).

5.2 First results on local isometric immersions : Janet, Cartan

If (N^t, h) and (M^n, g) are smooth Riemannian manifolds (with h a Riemannian metric on N and g a Riemannian metric on M), then a smooth map $f : N \to M$ is defined to be *isometric* in $y \in N$ if $h(u, v) = g(df(u), df(v))$ for each $u, v \in T(N)_y$. One readily verifies that then $df : T(N)_y \to T(M)_{f(y)}$ is injective. So a map $f : N \to M$ that is isometric in each point of N is an immersion, and we call it an *isometric immersion* of N in M. If $f : N \to M$ is an injective isometric immersion which is also a homeomorphism onto $f(N)$ (with the subspace topology), then f is called an *isometric embedding*. If a small neighbourhood of each point of N can be isometrically immersed (alternatively: embedded) in M, then we say that N can be *locally isometrically immersed* (alternatively: *embedded*) in M.

The interest in immersions and embeddings of manifolds in Euclidean spaces seems to have appeared first in Riemannian geometry. Indeed, already in the 1870's, in attempts to better understand the intrinsic geometry of Riemannian manifolds, L. Schläfli conjectured that any n-dimensional Riemannian manifold (M, g) with g of class C^ω can be locally isometrically C^ω-embedded in Euclidean $\frac{n(n+1)}{2}$-space.

In 1926, M. Janet published (in Ann. Sci. Polon. Math.) a proof of Schläfli's conjecture. It turned out later that the proof had gaps; it was rectified by C. Burstin (Mat. Sbornik in 1931). In his reaction to Janet's paper, É. Cartan in 1927 (in Ann. Sci. Polon. Math.) presented another proof of the conjecture, then transformed into the first general embedding theorem in Riemannian geometry (note that in some special situations, there were some older results, e.g., those proved by D. Hilbert).

In the proof Cartan used his theory of differential forms and Pfaffian systems (see 4.5). One can see here an interplay between vector distributions and local isometric immersions (or embeddings). The idea behind this is to obtain maps by looking for their graphs. This is based on the following result (see [136, 2.34]). Let N^t and M^n be smooth manifolds. Let $\pi_1 : N \times M \to N$ and $\pi_2 : N \times M \to M$ be the canonical projections. Suppose that there exists a basis $\{\omega_1, \ldots, \omega_n\}$ for the 1-forms on M (i.e., $\{\omega_1(x), \ldots, \omega_n(x)\}$ is a basis of the cotangent space, $T^*(M)_x$, for each $x \in M$).

(a) If $f : N \to M$ is a smooth map, then its graph (i.e., the submanifold (N, g) of $N \times M$, where $g(y) = (y, f(y))$ for each $y \in N$) is an integral manifold of the ideal in $E^*(N \times M)$ generated by the 1-forms

$$(d\pi_1)^*(df)^*(\omega_1) - (d\pi_2)^*(\omega_1), \ldots, (d\pi_1)^*(df)^*(\omega_n) - (d\pi_2)^*(\omega_n).$$

(b) Suppose that $\alpha_1, \ldots, \alpha_n$ are 1-forms on N such that the ideal in $E^*(N \times M)$ generated by the forms

$$(d\pi_1)^*(\alpha_1) - (d\pi_2)^*(\omega_1), \ldots, (d\pi_1)^*(\alpha_n) - (d\pi_2)^*(\omega_n)$$

is a differential ideal. Then, for any $(y_0, x_0) \in N \times M$, there exists a neighbourhood U of y_0 and a smooth map $g : U \to M$ such that $g(y_0) = x_0$ and such that

$$(dg)^*(\omega_1) = \alpha_{1|U}, \ldots, (dg)^*(\omega_n) = \alpha_{n|U}.$$

In addition to this, if U is any connected open set containing y_0 for which there exists a smooth map $g : U \to M$ such that $g(y_0) = x_0$ and such that

$$(dg)^*(\omega_1) = \alpha_{1|U}, \ldots, (dg)^*(\omega_n) = \alpha_{n|U},$$

then there exists precisely one such map on U.

For instance, when $M = \mathbb{R}^n$, then it is easy to find a basis $\{\omega_1, \ldots, \omega_n\}$ for the 1-forms on M. If $M = \mathbb{R}^n$ and all the requirements from (b) can be satisfied, for some forms $\alpha_1, \ldots, \alpha_n$, in such a way that $\alpha_1(y), \ldots, \alpha_n(y)$ generate $T^*(N)_y$ for each $y \in U$, then of course $(dg)^* : T^*(M)_{g(y)} \to T^*(U)_y$ is an epimorphism, hence $g : U \to \mathbb{R}^n$ is an immersion.

5.3 Whitney's global immersions and embeddings

In 1936, H. Whitney, in the introduction to [138] writes (we use our notation but closely follow his words) that a differentiable manifold is generally defined in one of two ways: as a point set with neighbourhoods homeomorphic with Euclidean space $\mathbb{R}^n$, coordinates in overlapping neighbourhoods being related by a differentiable transformation, or as a subset of $\mathbb{R}^n$, defined near each point by expressing some of the coordinates in terms of the others by differentiable functions. Thanks to this paper by Whitney, it became known that the first definition (apparently more abstract, appearing since the 1930's in works by O. Veblen and J. H. C. Whitehead, P. S. Alexandroff and H. Hopf, J. Alexander) is no more general than the second (apparently more "palpable", introduced in 1895 by H. Poincaré in Analysis situs; see J. Dieudonné's book A history of algebraic and differential topology, Birkhäuser 1989). Indeed, Whitney proves there that any smooth Hausdorff manifold M^n with a countable base for its topology can be considered as an embedded submanifold in $\mathbb{R}^{2n+1}$; in addition to this, there is a smooth immersion $M \to \mathbb{R}^{2n}$. We restrict ourselves here to the C^∞-context, but Whitney's result is more general; he also proves there a theorem about approximating smooth maps $M \to \mathbb{R}^{2n+1}$ by embeddings. Later, in [139], he presented still better general embedding and immersion dimensions by proving that every smooth Hausdorff manifold M^n with a countable basis embeds in $\mathbb{R}^{2n}$ if $n \geq 1$ (but the approximation property mentioned above cannot be transferred to this case), and immerses in $\mathbb{R}^{2n-1}$ if $n \geq 2$. In general, these dimensions cannot be improved, because, e.g., $\mathbb{R}P^{2^r}$ cannot be immersed in $\mathbb{R}^{2^{r+1}-2}$ and cannot be embedded in $\mathbb{R}^{2^{r+1}-1}$ (for immersions, see [94, 4.8]; for embeddings, take the Klein bottle). But with added information or further assumptions, e.g., on n or on properties of manifolds, one can try to find better (that is lower) or even optimal (in the obvious sense) immersion or embedding dimensions, hence some Whitney-like immersion or embedding theorems, under more specific conditions. This really became the aim of much effort in the subsequent years.

5.4 Global isometric immersions: Nash, Kuiper et al.

J. Nash in 1954 for $q \geq n+2$ and N. Kuiper in 1955 for $q \geq n+1$ proved that any smooth immersion $f_0 : M^n \to \mathbb{R}^q$ admits a C^1 homotopy of immersions $f_t : M \to \mathbb{R}^q$ ($t \in I$, where $I = [0,1]$ is the closed unit interval between 0 and 1) to an isometric immersion $f_1 : M \to \mathbb{R}^q$. The first result on global isometric embeddings, showing that also in the realm of smooth Riemannian manifolds, "abstractly" defined manifolds are not more general than those defined as submanifolds of Euclidean spaces, came quite a few years after the local result of Janet and Cartan or after Whitney's results. Indeed, in 1956, J. Nash ([95]) proved that every compact Riemannian C^t-manifold $(3 \leq t \leq \infty)$ of dimension n can be isometrically C^t-embedded in Euclidean $\frac{n(3n+11)}{2}$-space and every non-compact Riemannian C^t-manifold $(3 \leq t \leq \infty)$ of dimension n can be isometrically C^t-embedded in Euclidean $\frac{n(n+1)(3n+11)}{2}$-space. One of the main steps in Nash's proof was a certain perturbation problem. To solve it, he invented an unprecedented procedure, based on a hard implicit function theorem, which later, so to say, started to live its own life. Indeed, in the 1960-1980's, the procedure was improved by several authors (among them, J. Moser, L. Hörmander, L. Nirenberg, R. Hamilton, H. Jacobowitz, J. Schwartz, E. Zehnder) and developed into what is known as Nash-Moser's theory in the context of differential operators and non-linear systems of partial differential equations (see [46]).

Nash's embedding dimension mentioned above has subsequently been improved. M. Gromov (see [46, 3.1.1]) proves that every (closed or open) Riemannian C^t-manifold with $2 < t \leq \infty$ admits an isometric C^t-embedding in $\mathbb{R}^{n^2+10n+3}$, and for $t > 4$, the dimension of the ambient space can be reduced to $\frac{(n+2)(n+3)}{2}$. In 1987, quite surprisingly, M. Günther [49] shows that Nash's complicated procedure was in fact not necessary for solving the perturbation problem mentioned above. Günther then (Math. Nachr. 144 (1989)), using his method, also lowered Nash's embedding dimension: he proved that any C^∞ Riemannian metric on M^n can be induced by a C^∞-embedding of the manifold M into $\mathbb{R}^q$ with

$$q = \max(\frac{n(n+5)}{2},\, 5 + \frac{n(n+3)}{2}).$$

There are many questions (freeness of immersions; classes of differentiability of manifolds versus classes of differentiability of isometric immersions etc.) which we could not touch here. For them, the reader may wish to consult, besides Gromov's book and other works already cited in this subsection, also, e.g., [45], [4], [18], [29], [115]. Several results relating isometric immersions to vector distributions (and other interesting results) are presented by Z. Z. Tang in [126]. For instance, he proves that if an oriented smooth Riemannian manifold M^n with positive scalar curvature can be isometrically C^∞-immersed in $\mathbb{R}^{n+2}$, then M is stably parallelizable. From many others working recently on isometric immersions or embeddings, we mention, e.g., Yu. A. Aminov, J. Cieslinski, M. Dajczer, R. Tojeiro, H. Tanabe, M. Pakzad, A. Savo, U. Lumiste, S. Mardare.

5.5 Hirsch and Smale's immersion theory

As an additional reading to the material of this subsection, we recommend, mainly for the developments in immersion and embedding theory until 1973, S. Gitler's [37] and D. Spring's [121].

In 1959, the papers by S. Smale [118] and M. Hirsch [56] (the second building on the results of the first) brought a new point of view to the study of immersions. In these papers, Smale for the immersions of spheres in Euclidean spaces and Hirsch for immersions of closed smooth manifolds in smooth manifolds, reduced the existence and classification questions for immersions of manifolds to problems in homotopy theory. To explain what is meant by this, we first introduce some notations and definitions.

Two immersions $f, g : N^t \to M^n$ are defined to be *regularly homotopic* (recall that immersions are *regular* maps, hence the name) if, roughly speaking, there is a homotopy $F : N \times I \to M$ (called then a *regular homotopy*) from f to g through (smooth) immersions. More precisely, it is required here that $F(x, 0) = f(x)$ and $F(x, 1) = g(x)$ for all $x \in N$, that the map $F_t : N \to M$ defined by $F_t(x) = F(x, t)$ be an immersion for each $t \in I$, and that the map $\tilde{F} : T(N) \times I \to T(M)$ defined by $\tilde{F}((x, v), t) = (F_t(x), d(F_t)_x(v))$ be continuous (here $x \in N$, $v \in T(N)_x$, and $d(F_t) : T(N) \to T(M)$ is the differential of the smooth map F_t). Let $\mathrm{Imm}(N, M)$ be the space of immersions from N to M (with the C^∞ topology; see [57, Ch. 2]), and let $\mathrm{Mono}(T(N), T(M))$ be the space of vector bundle monomorphisms $T(N) \to T(M)$ with the compact-open topology. Recall that a (continuous) map $f : X \to Y$ between the topological spaces X and Y is a *weak homotopy equivalence* if f induces a one-to-one correspondence between the path components of X and those of Y, and also an isomorphism between the ith homotopy group of X and the ith homotopy group of Y for each $i \geq 1$. Note that a point in the space $\mathrm{Imm}(N, M)$ is nothing but an immersion $N \to M$, and a path in $\mathrm{Imm}(N, M)$ is a regular homotopy.

Now we can state a fundamental theorem, which we shall call the Hirsch-Smale theorem: If $t < n$, then the map

$$D : \mathrm{Imm}(N^t, M^n) \to \mathrm{Mono}(T(N), T(M)),$$

mapping each $g \in \mathrm{Imm}(N, M)$ to its differential, is a weak homotopy equivalence.

We remark that this is not the original formulation from [56]; it was found later by Hirsch and Palais (see Smale's survey [119, 3.12]). An analogous result was proved by A. Phillips in [100] for submersions $U^p \to W^q$ $(p \geq q)$ if U^p is an open manifold (hence this extends the Hirsch-Smale theorem also to the case $t = n$ if U is open; the latter case was also proved by Hirsch in 1961).

As a consequence of the Hirsch-Smale theorem, the existence of an immersion $N^t \to M^n$ $(t < n)$ is equivalent to the existence of a vector bundle monomorphism $T(N) \to T(M)$. In addition, two immersions $N^t \to M^n$ $(t < n)$ are regularly homotopic if and only if their differentials are homotopic through vector bundle monomorphisms from $T(N)$ to $T(M)$.

Now let us take $M^n = \mathbb{R}^n$; of course, the tangent bundle of $\mathbb{R}^n$ is trivial. Then the Hirsch-Smale theorem readily implies the following vector bundle criterion for the existence of immersions of a smooth closed manifold N^t in Euclidean spaces: N^t immerses in $\mathbb{R}^n$ $(t < n)$ if and only if there exists a vector bundle κ^{n-t} such that the Whitney sum $T(N) \oplus \kappa^{n-t}$ is a trivial vector bundle. By this theorem, e.g., we immediately know that any stably parallelizable (closed) manifold of dimension n immerses in R^{n+1} (we note that every open n-dimensional parallelizable manifold can be immersed in $\mathbb{R}^n$, as shown by Hirsch in 1959). A surprising consequence of the Hirsch-Smale theorem and the fact that the second homotopy group $\pi_2(V_{3,2}) = \pi_2(SO(3))$ is trivial is that the sphere S^2 can be turned inside out in the sense that the usual embedding of S^2 in $\mathbb{R}^3$, defined by $x \mapsto x$, is regularly homotopic to the embedding defined by $x \mapsto -x$. This was found by Smale; for a generalization, see U. Kaiser [67].

Supposing that $n > t$, if a manifold N^t can be immersed in $\mathbb{R}^n$, then of course, it can also be immersed in $\mathbb{R}^{n+s}$, and the geometric dimension (see 4.13) of the stable normal bundle of the latter immersion does not exceed $n-t$. By what we said above (using stability properties of vector bundles; see [60, Part II, Ch. 8, Theorem 1.5]), one sees that also the converse is true. Indeed, supposing that $n > t$, if there is an immersion $N^t \to \mathbb{R}^{n+s}$ and if the corresponding stable normal bundle can be represented by an $(n-t)$-dimensional vector bundle, then N immerses in $\mathbb{R}^n$. In particular, if $n > t$ and we have an immersion $N^t \to \mathbb{R}^{n+s}$ with the normal bundle ν^{n+s-t} having a trivial s-distribution, hence such that $\operatorname{span}(\nu^{n+s-t}) \geq s$ (or, equivalently, if the associated Stiefel fibre bundle $V_s(\nu)$, with fibre isomorphic to $V_{n+s-t,s}$, has a section), then we have an immersion $N^t \to \mathbb{R}^n$. Here we see a close relation between immersions of manifolds in Euclidean spaces on the one hand, and vector distributions (in vector bundles) on the other.

Summarizing, we can look at immersions of manifolds in Euclidean spaces in the following way. Given a (closed) manifold N^t, we know from Whitney's results, that it immerses in $\mathbb{R}^{2t-1}$ $(t \geq 2)$; let (ν) be the stable normal bundle of such immersions. The question of whether or not N^t also immerses in some $\mathbb{R}^n$ with $t < n < 2t-1$ reduces, thanks to Hirsch-Smale theory, to the question of whether or not the geometric dimension of (ν) is less than or equal to $n-t$. As we have seen in Sec. 4, the latter question can be attacked by obstruction theory in its various forms, including the theory of (modified) Postnikov resolutions (see, e.g., [87]) or the theory of obstructions obtained by the singularity approach (see, e.g., [77] and the references cited therein), and in general is very difficult. From the same point of view it is clear that results on homotopy classification of sections of Stiefel bundles associated to vector bundles (like $V_s(\nu)$ above) may imply results on classification of immersions under the equivalence relation of regular homotopy. Indeed, some results in this direction can be found, e.g., in [64], [65]. There are also other types of classification results for immersions; see, e.g., Haefliger and Hirsch's paper published in 1962 in the Annals of Mathematics, where among other important results they show that the classification of immersions $N^t \to M^n$ in the stable range of dimensions, $2n > 3t+1$, is independent of the differentiable structures on the closed manifolds N and M. Somewhat later, in 1974, H. Glover and G. Mislin [38] proved for simple manifolds (hence such that the fundamental group operates trivially on all the homotopy groups) the following strong result in the same direction. Let M and N be smooth, closed, connected, orientable, simple manifolds of dimension n whose 2-localizations are homotopy equivalent. Suppose that M immerses in $\mathbb{R}^{n+k}$ for some $k \geq [n/2]+1$. Then N immerses in $\mathbb{R}^{n+2[k/2]+1}$.

5.6 Homotopy methods in embedding theory: Haefliger et al.

Once Hirsch-Smale theory of immersions was created, the theory of embeddings did not have to wait too long for its own "homotopization", although less complete than for immersions. Indeed, already in the early 1960's, A. Haefliger, extending results of A. Shapiro and W. T. Wu, succeeded, in the stable range, in reducing also the existence and classification questions for embeddings to homotopy problems. We need some definitions. An *isotopy* is defined to be a regular homotopy which at each "time" t $(t \in I)$ is an embedding. Two embeddings $f,\ g : N^t \to M$ are called *isotopic* if there is an isotopy $F : N \times I \to M$ from

$f = F_0$ to $g = F_1$. Haefliger, in [54], [55], presented several new results on embedding smooth manifolds in other smooth manifolds (then in later papers he added more). In particular, he reduced the classification of smooth embeddings $N^t \to \mathbb{R}^n$ to the classification of sections of a suitable fibre bundle, under the condition $2n > 3t + 3$. More precisely, let $\Delta = \{(x, y) \in N \times N; x = y\}$ be the diagonal for the closed manifold N^t. The group $\mathbb{Z}_2$ acts on $N \times N \setminus \Delta$ (we have $(-1)(x, y) = (y, x)$) without fixed points. Let N^* be the orbit space of this action. Let $E \to N^*$ denote the fibre bundle, with fibre the sphere S^{n-1}, associated to the principal $\mathbb{Z}_2$-bundle $N \times N - \Delta \to N^*$; here the action of $\mathbb{Z}_2$ on S^{n-1} is defined by the antipodal map $a : S^{n-1} \to S^{n-1}$, $a(x) = -x$. Haefliger's theorem then says that the (smooth) isotopy classes of smooth embeddings $N^t \to \mathbb{R}^n$ are in a bijective correspondence with the homotopy classes of sections of the fibre bundle $E \to N^*$, if $2n > 3t + 3$. [The correspondence is via sending any embedding $f : N^t \to \mathbb{R}^n$ to the $\mathbb{Z}_2$-equivariant map $\tilde{f} : N \times N \setminus \Delta \to S^{n-1}$, $\tilde{f}(x, y) = \frac{f(x)-f(y)}{||f(x)-f(y)||}$; such equivariant maps canonically correspond to sections of $E \to N^*$.]

Let $\pi_k(C^\infty(N^t, M^n), \mathrm{Emb}(N, M), f_0)$ (briefly π_k) $(k \geq 1)$ be the kth relative homotopy group if $k \geq 2$ or just the kth relative homotopy set if $k = 1$, where M and N are smooth manifolds, M is supposed to be closed, $C^\infty(N^t, M^n)$ is the space (with the C^∞-topology) of all smooth maps $N \to M$, $\mathrm{Emb}(N, M)$ is its subspace of embeddings, and $f_0 : N \to M$ is an embedding. Although $\pi_0(C^\infty(N^t, M^n), \mathrm{Emb}(N, M), f_0)$ (briefly π_0) is not defined, we agree to write symbolically $\pi_0(C^\infty(N^t, M^n), \mathrm{Emb}(N, M), f_0) = 0$ (briefly $\pi_0 = 0$) to mean that every path component of $C^\infty(N^t, M^n)$ intersects $\mathrm{Emb}(N, M)$ or, equivalently, that each map $N \to M$ is homotopic to an embedding. About 1972, J.-P. Dax, inspired by Haefliger's results and methods, studies in [33], in a profound and extensive way, π_k for $k \geq 0$, generalizing some previously known theorems (due to Whitney, Haefliger et al.) and also obtaining completely new information for $k \geq 2$. He improved methods for elimination of double points of immersions (previously used, in simpler forms, also by Whitney and Haefliger) and reduced the study of $\pi_k(C^\infty(N^t, M^n), \mathrm{Emb}(N, M), f_0)$ to the study of certain normal bordism or normal cobordism groups. Similar ideas also appear in Hatcher and Quinn's [53].

In the early 1960's, another approach to embeddings has been developed, mainly in the works of W. Browder, J. Levine, and S. P. Novikov. They used surgery for constructing embeddings of one manifold in another; further substantial contributions are due to A. Casson, D. Sullivan, C. T. C. Wall. On the foundations of the theory and on later developments and ramifications, the reader can consult M. Cencelj, D. Repovš, and A. Skopenkov's and T. Goodwillie, J. Klein, and M. Weiss's recent surveys [23] and [42], respectively.

5.7 Gromov's h-principle

About 1969, a major reform in the immersion theory (and also in other areas, e.g., in the theory of submersions, symplectic and contact geometry, non-linear partial differential equations etc.) was initiated by M. Gromov. The main tool of this reform, which enables one to bring apparently unrelated results on the same platform, is called the *homotopy principle*, briefly h-principle.

To explain what is meant by the h-principle, we need some preparation. Following Gromov [46, 1.1.1], consider a smooth fibration $p : X \to N$. Let, for $r \geq 0$, $X^{(r)}$ be the space of r-jets (of germs) of smooth sections of p (we identify $X^{(0)} = X$). In particular, the space $X^{(1)}$ consists of the linear maps between tangent spaces, $L : T(N)_{p(x)} \to$

$T(X)_x$ for all $x \in X$, such that $dp \circ L = \mathrm{id}$. We have natural projections $p^r : X^{(r)} \to N$ and $p^s_r : X^{(s)} \to X^{(r)}$ for $s > r \geq 0$. Let the r-jet of a C^r-section $f : N \to X$ be denoted by $J^r_f : N \to X^{(r)}$. A section $g : N \to X^{(r)}$ is called *holonomic* if $g = J^r_f$ for some C^r-section $f : N \to X$ (if such a section f exists, then it is unique).

Now a differential relation imposed on sections is a generalization (to smooth fibrations) of differential equations or inequalities. More precisely, an rth order *differential relation* imposed on sections $f : N \to X$ is a subset $\mathcal{R} \subset X^{(r)}$. A C^r-section $f : N \to X$ is said to satisfy (or to be a solution of) the relation $\mathcal{R}$ if the r-jet of f, $J^r_f : N \to X^{(r)}$, maps N into $\mathcal{R}$. Hence solutions of $\mathcal{R}$ can be naturally identified with holonomic sections $N \to \mathcal{R} \subset X^{(r)}$. We say that the relation $\mathcal{R}$ *satisfies the h-principle* (or that the h-principle holds for obtaining solutions of $\mathcal{R}$) if every continuous section $N \to \mathcal{R}$ is homotopic through sections $N \to \mathcal{R}$ to a holonomic section $N \to \mathcal{R}$.

In particular, if X is a trivial fibration, $X = N \times M \to N$, $(a, b) \mapsto a$, then sections $N \to X$ correspond to maps $N \to M$. So, quite naturally, instead of the h-principle for sections $N \to X$, we speak about the h-principle for maps $N \to M$.

For example, let us take the trivial fibre bundle $X = N \times M \to N$, $(a, b) \mapsto a$, where N^t and M^n are C^1-manifolds. The jet space $X^{(1)}$ consists then (after the obvious identifications) of all linear maps $T(N)_a \to T(M)_b$ for all $(a, b) \in N \times M$. Further, in the rôle of $\mathcal{R}$, we define the *immersion relation* $\mathcal{I} \subset X^{(1)}$, fibred over $X = N \times M$ by the projection $p^1 : X^{(1)} \to X$, by the requirement that the fibre $\mathcal{I}_x$ over $x = (a, b) \in X = N \times M$ should consist of those linear maps belonging to $X^{(1)}_x$ which are *injective*, hence of injective linear maps $T(N)_a \to T(M)_b$. Now sections $N \to \mathcal{I}$ correspond to vector bundle monomorphisms $T(N) \to T(M)$, and a section $g : N \to \mathcal{I} \subset X^{(1)}$ is holonomic if and only if there exists some $f : N \to M$ such that $g = df$. In other words, holonomic sections $g : N \to \mathcal{I} \subset X^{(1)}$ are differentials $df : T(N) \to T(M)$ of immersions $f : N \to M$.

Now it is clear that the Hirsch-Smale theorem can be stated as follows: If $t < n$, then every vector bundle monomorphism $T(N) \to T(M)$ is homotopic, through vector bundle monomorphisms, to a vector bundle homomorphism which is the differential of an immersion $N \to M$. In other words, if $t < n$, then every section $N \to \mathcal{I}$ is homotopic, through sections $N \to \mathcal{I}$, to a holonomic section $N \to \mathcal{I}$. So the Hirsch-Smale theorem is the same as to say that if $t < n$, then the immersion relation $\mathcal{I}$ satisfies (in other wording: immersions $N \to M$ satisfy) the h-principle.

In 1969, Gromov [44] presented his method of *convex integration of differential relations* (see also [46, 2.4]) as one of several known methods for proving the h-principle. Other such methods are mainly the method of removal of singularities, the covering homotopy method used in Gromov's doctoral dissertation supervised by V. A. Rokhlin, the method of continuous sheaves, and the method of inversions of differential operators. In particular, in this way one obtains new proofs of the Hirsch-Smale theorem, but also, e.g., of the results on isometric immersions obtained by Nash or Kuiper. So the h-principle became really a common platform for many results which before appeared unrelated.

In addition to bringing new proofs of known theorems, the studies of the h-principle have also lead to many new results, not only on immersions, and not only in global analysis or geometry. Details on relevant methods and results can be found in [46], in original articles (by Gromov, Eliashberg, Mishachev, du Plessis, and others), and also in several recent books on the h-principle, e.g., by D. Spring (1998) or by Y. Eliashberg and N.

Mishachev (2002).

In 1974, E. Bierstone [16] derived an equivariant version of Gromov's theory (as presented in [44]). Some of Bierstone's assumptions were then weakened by S. Izumiya (Manuscripta Math. in 1979). In 1998, M. Datta and A. Mukherjee [30] studied the open extension theorem of Gromov ([46, p. 86]) in an equivariant setting (G-manifolds, where G is a compact Lie group, G-invariant relations etc.). As applications, they obtained a generalization of the transversality theorem of Gromov [46, p. 87] and an equivariant version of the Hirsch-Smale immersion theorem.

5.8 A category theory approach

There is also a categorial reformulation of the (extended) Hirsch-Smale theorem. Roughly, supposing that N^t and M^n are smooth manifolds (without boundary) such that $t < n$ (or also $t = n$, but then N must be open), let $\mathcal{O}$ be the poset of open subsets of N, ordered by inclusion. Of course, $\mathcal{O}$ is a category, with just one morphism $U \to V$ if $U \subset V$ and with no morphism if $U \not\subset V$. Then the spaces $\mathrm{Imm}(N^t, M^n)$ and $\mathrm{Mono}(T(N), T(M))$ are just the values $E(N)$ and $F(N)$, respectively, if E and F are contravariant functors from $\mathcal{O}$ to the category of spaces defined by $E(U) = \mathrm{Imm}(U, M)$ and $F(U) = \mathrm{Mono}(T(U), T(M))$, respectively. By identifying any immersion with its differential (which is then a vector bundle monomorphism between the corresponding tangent bundles), we obtain an obvious inclusion $E(U) \subset F(U)$ for any open $U \subset N$. Now a categorial reformulation of the Hirsch-Smale theorem reads: The functors E and F are excisive (or, in Goodwillie's calculus terminology: they are polynomial of degree at most one). For ideas and results in this area see, e.g., [137], [42].

5.9 Rourke and Sanderson's compression theorem

About 1999, C. Rourke and B. Sanderson presented a new approach to immersion theory. It is based on their new *compression theorem*: Let N^t and M^n be smooth manifolds; let N be closed. Suppose that $n > t$ and N is embedded in $M \times \mathbb{R}$ and equipped with a normal vector field. Then the vector field can be made parallel to the given $\mathbb{R}$ direction by an isotopy of M and normal field in $M \times \mathbb{R}$. Hence, N can be isotoped to a position, where it projects by the obvious projection (in other words, it is "compressed") to an immersion in M.

The Rourke-Sanderson compression theorem is related to Gromov's theorem on directed embeddings ([46, 2.4.5 (C')]), implies a constructive proof of the Hirsch-Smale theorem on immersions (helps to "visualize" them), and has also other interesting applications; see [107], [108].

5.10 Various general results on immersions and embeddings

Hirsch and Smale's theory combined with Whitney's theorems and obstruction theory (used for obtaining information on the geometric dimension of the stable normal bundle, as already mentioned above) became the most effective tools for deriving new results on immersions of manifolds in Euclidean spaces. For instance, in 1964, M. Mahowald [87], using modified Postnikov resolutions (or towers) and also results due to W. Massey, F. Peterson, A. Haefliger, and M. Hirsch, proves the following theorem: Let M^n $(n > 4)$ be a closed orientable manifold. If n is even, then M^n immerses in $\mathbb{R}^{2n-2}$ if and only if the cup product of the dual Stiefel-Whitney classes, $\bar{w}_2(M)\bar{w}_{n-2}(M)$, vanishes; if n is odd, then M^n immerses in $\mathbb{R}^{2n-2}$. Or, using higher order cohomology operations for calculating k-invariants in Postnikov resolutions for the stable normal bundle of a manifold, D. Randall proves in [102] that a spin manifold M^n immerses in $\mathbb{R}^{2n-3}$ for $n \equiv 0 \pmod 4$ and n not a power of 2, and several other theorems on immersions of orientable manifolds.

For nonorientable manifolds, it was not rare that neither the classical (Steenrod) obstruction theory nor Postnikov-resolution approaches worked well. Fortunately, in the 1980's the situation changed, thanks to bordism-obstructions which turned out to produce interesting results also on nonorientable manifolds. In [78], Koschorke gave a detailed adjustment of his bordism-obstruction theory ([77]; see also our Sec. 4) to immersion theory, and illustrated it with many concrete calculations. His student C. Olk (Dissertation, University of Siegen 1980), in this framework, derived several results on immersions of a smooth closed manifold M^n, not necessarily orientable, in $\mathbb{R}^{2n-k}$, $k = 2, 3, 4, 5$. A sample result, for $k = 2$: Suppose that M^n is a smooth closed manifold, $n \geq 5$. If $n \not\equiv 0 \pmod 4$, then M immerses in $\mathbb{R}^{2n-2}$. If $n \equiv 0 \pmod 4$ and M is orientable, then Olk's criterion is the same as that of Mahowald (mentioned above). If $n \equiv 0 \pmod 4$ and M is nonorientable, then M immerses in $\mathbb{R}^{2n-2}$ if and only if the dual twisted integer Stiefel-Whitney class $\bar{W}_{n-1}(M)$ vanishes. The bordism-obstructions turned out to be effective also in classification of immersions. Indeed, U. Kaiser and B. H. Li in [68] gave an enumeration of homotopy classes of vector bundle monomorphisms $\alpha^n \to \beta^{2n-2}$ for vector bundles α^n and β^{2n-2} over a smooth closed manifold M^n $(n \geq 6)$. In particular, they enumerated immersions $M^n \to \mathbb{R}^{2n-2}$ for all $n \geq 6$.

Similarly to studying immersions of M^n in $\mathbb{R}^{2n-k}$ with low values of k, there appeared several papers considering immersions (especially up to cobordism) with values of k close to n. The previous efforts in this direction – by A. Liulevicius, R. L. Brown, S. Kikuchi, for $k = n - 1$ or $k = n - 2$ – are summarized, and some of them also improved, by R. Stong in [124]. He also derives new necessary conditions for immersing M^n in $\mathbb{R}^{n+3}$. We cite at least a sample result: If M^n is oriented, immerses in $\mathbb{R}^{n+3}$, and n is even and larger than 4, then M^n is an unoriented boundary. The lowest dimension of Euclidean space in which all n-dimensional orientable manifolds are immersible up to unoriented cobordism was determined by I. Takata in [127].

About 1960, W. Massey proved that, for any smooth closed manifold M^n, the dual Stiefel-Whitney classes $\bar{w}_i(M)$ vanish for $i > n - \alpha(n)$, where $\alpha(n)$ is the number of ones in the dyadic expansion of n. [This result is best possible, because if $n = 2^{j_1} + 2^{j_2} + \cdots + 2^{j_s}$ $(j_1 < j_2 < \cdots < j_s)$, then one readily verifies that $\bar{w}_{n-\alpha(n)}(\mathbb{R}P^{2^{j_1}} \times \cdots \times \mathbb{R}P^{2^{j_s}})$ does not vanish.] So any smooth closed manifold M^n $(n \geq 2)$ fulfils the Stiefel-Whitney necessary condition (see the beginning of Sec. 5) for the existence of an immersion $M \to \mathbb{R}^{2n-\alpha(n)}$. The claim that each smooth closed manifold M^n $(n \geq 2)$ immerses in $\mathbb{R}^{2n-\alpha(n)}$ became known as the *immersion conjecture*. E. Brown and F. Peterson developed, in their papers published in 1964-1979, a program with the aim of proving this conjecture, but they did not complete it. In the meantime, in 1971, R. L. Brown [19] proved the conjecture up to cobordism, which made it even more attractive. More precisely, he proved that every smooth closed manifold M^n is cobordant to a manifold immersing in $\mathbb{R}^{2n-\alpha(n)}$ and embedding in $\mathbb{R}^{2n-\alpha(n)+1}$. Then in 1985, R. Cohen in [24] presented his completion of the Brown-Peterson program.

We say that a closed manifold M^n has *immersion codimension* k if it immerses in $\mathbb{R}^{n+k}$ but does not in $\mathbb{R}^{n+k-1}$. The *embedding codimension* of M^n is defined analogously. Since the 1960's, much effort was devoted to optimizing immersion or embedding dimensions, hence to improving upper bounds for the immersion or embedding codimensions. Much work was also dedicated to the study of cobordism groups of immersions and embeddings (see, e.g., papers by R. Wells, G. Pastor, A. Szűcs), to the study of geometric and topological properties of multiple points of immersions of manifolds in Euclidean spaces (see, e.g., papers by U. Koschorke, F. Ronga, P. Eccles, A. Szűcs, N. Boudriga, S. Zarati), or to the study of various immersion (or embedding) related questions (see, e.g., papers by C. Biasi, W. Motta, O. Saeki, A. Szűcs, M. Takase, K. Sakuma, T. Yasui).

In the 1970's there also appeared several new existence theorems on embeddings. For instance, in 1974, Glover and Homer present an immersion-to-embedding result (obtained using homotopy localization methods) in [39]. In 1976, E. Thomas in [132], summarizes the knowledge on embeddings of a smooth closed manifold M^n in R^{2n-1} as follows. By Whitney, every M^n embeds in R^{2n}; a combination of results due to Haefliger, Hirsch, Massey, Peterson, and Rigdon (for precise data, see [132]) implies that every orientable M^n embeds in $\mathbb{R}^{2n-1}$ $(n > 4)$; if n is not a power of two, then every M^n embeds in $\mathbb{R}^{2n-1}$; finally, for n a power of two $(n > 4)$, a nonorientable M^n embeds in $\mathbb{R}^{2n-1}$ if an only if the dual Stiefel-Whitney class $\bar{w}_{n-1} \in H^{n-1}(M;\mathbb{Z}_2)$ is zero. Thomas then derives, using the embedding theory of Haefliger [55]), two sets of sufficient conditions for embedding an n-dimensional manifold in $\mathbb{R}^{2n-2}$. We cite at least one of them: Let M^n $(n \geq 7)$ be an orientable smooth closed manifold with $\bar{w}_{n-3+i}(M) = 0$ for $i \geq 0$. If either $w_3(M) \neq 0$, or $w_2(M) \neq 0$ and the homology group $H_1(M;\mathbb{Z})$ has no 2-torsion, then M embeds in $\mathbb{R}^{2n-2}$.

For a smooth closed manifold M, let $[M \subset \mathbb{R}^m]$ denote the set of isotopy classes of embeddings $M \to \mathbb{R}^m$. T. Yasui in [140] first summarizes known results on the set $[M^n \subset \mathbb{R}^{2n+i}]$, where $i = 1, 2, 3$ (for instance, by Whitney [138], $[M^n \subset \mathbb{R}^{2n+2}]$ is just a 1-element set if $n \geq 1$). Then he generalizes a result due to Haefliger, by identifying the set $[M^n \subset \mathbb{R}^{2n-1}]$ with a set obtained from the cohomology groups of M using cohomology operations and characteristic classes, under weaker restrictions on M than those posed by Haefliger. Yasui studied such "enumerations of embeddings" in a series of papers, but he also attacked other embedding related questions, e.g., the question of when a map is homotopic to an embedding. For a recent result of this type see [84].

An important rôle in the study of immersions and embeddings has also been played by necessary conditions; they mainly serve as a source of nonimmersion or nonembedding results (essential, when one tries to optimize immersion or embedding dimensions). We already have mentioned some of such necessary conditions (vanishing of certain Stiefel-Whitney classes or Atiyah's γ-operations). Now we shall add more. We first recall some notions (they can be found in [9]; see also [69]).

Let M^{2n} be a smooth closed connected oriented even-dimensional manifold, and let $p_i = p_i(M) \in H^{4i}(M;\mathbb{Z})$ be the ith Pontrjagin class of M. Recall (see, e.g., [17, §23]) that the total Atiyah-Hirzebruch class $\hat{\mathcal{A}}(M)$ is defined to be $\sum_{j=0}^{\infty} \hat{A}_j(p_1,\ldots,p_j)$, where $\{A_j\}_j$ is the multiplicative sequence of polynomials associated with the power series $(\frac{1}{2}\sqrt{t})/\sinh(\frac{1}{2}\sqrt{t})$. For each $d \in H^2(M;\mathbb{Q})$ and each $z \in H^*(M;\mathbb{Q})$, we define $\hat{A}(M,d,z)$ to be the rational number obtained by evaluating the top-dimensional component of $ze^d\hat{\mathcal{A}}(M)$ on the fundamental homology class $[M]$. Further we denote by

$i_* : KO(M) \to K(M)$ the homomorphism from the real K-ring to its complex counterpart defined by complexification. Let $ch : K(M) \to H^{\text{even}}(M;\mathbb{Q})$ be the Chern character (which is a ring homomorphism). We denote the image $ch(K(M))$ by $ch(M)$, and the image $ch(i_*(KO(M))$ by $chO(M)$. Any class in $ch(M)$ can be written as $z = \sum_i z_i$ with $z_i \in H^{2i}(M;\mathbb{Q})$. We define (for any $t \in \mathbb{Q}$) $z^{(t)} = \sum_i t^i z_i$. Then, for each $d \in H^2(M;\mathbb{Z})$ and $z \in ch(M)$, $\hat{A}(M, d/2, z^{(t)})$ is a polynomial in t, with rational coefficients, of degree less than or equal to n, and is called the *Hilbert polynomial* for M (and the fixed classes d and z); we shall denote it by $H(t)$. [For a motivation of the name of this polynomial, see [9].] If $d \bmod 2 = w_2(M)$, then $H(t)$ is an integer for each $t \in \mathbb{Z}$.

Now we can state the following theorem, due to Atiyah, Hirzebruch, and Mayer ([9], [90]), giving some necessary conditions for immersions or embeddings. Let M^{2n} be a smooth, closed, connected, oriented manifold.

(i) If M^{2n} can be immersed in $\mathbb{R}^{2n+2s}$ or in $\mathbb{R}^{2n+2s+1}$, then $2^{n+s}H(\frac{1}{2}) \in \mathbb{Z}$. If M^{2n} can be embedded in $\mathbb{R}^{2n+2s}$, then $2^{n+s-1}H(\frac{1}{2}) \in \mathbb{Z}$.

(ii) Assume that $n = 2m$, $d = 0$, and $z \in chO(M)$. If M^{2n} immerses in $\mathbb{R}^{4m+2s}$ and $2m + s \equiv 1, 2, 3 \pmod 4$ or if M^{2n} immerses in $\mathbb{R}^{4m+2s+1}$ and $2m + s \equiv 1, 2 \pmod 4$, then $2^{2m+s-1}H(\frac{1}{2}) \in \mathbb{Z}$. If M embeds in $\mathbb{R}^{4m+2s}$ and $2m + s \equiv 2 \pmod 4$, then $2^{2m+s-2}H(\frac{1}{2}) \in \mathbb{Z}$.

We remark that Mayer's proof uses the Atiyah-Singer index theorem. Similar necessary conditions, also based on methods involving indices of elliptic operators, were published by H. Lawson and M.-L. Michelsohn ([85], [86]).

In the studies of the corresponding problems, inequalities of Korbaš and Szűcs [73, Proposition 3.8], interrelating immersions of closed manifolds in Euclidean spaces, trivial vector distributions (the vector field problem), and the Lyusternik-Shnirel'man category, may also be useful. We have in mind the following theorem. Let M^n be a smooth closed manifold which is not stably parallelizable, let s be its stable span and k its immersion codimension. Then we have

$$k \leqq (n - s)(\text{cat}(M^n) - 1) \text{ and } n - s \leqq k(\text{cat}(M^n) - 1).$$

We remark that the second inequality cannot be improved in general, because it is an equality for $M = \mathbb{R}P^2$.

5.11 Remarks on immersions and embeddings of specific manifolds

The projective spaces (not only real, but also complex and quaternionic) are a natural family of manifolds for testing general immersion or embedding theorems, but there are also special approaches invented and working just for them. Results in this area are due to, e.g., H. Whitney, H. Hopf, S. S. Chern, W. T. Wu, J. Milnor, D. Epstein, R. Schwarzenberger, M. Ginsburg, J. Levine, S. Feder, B. Sanderson, R. J. Milgram, J. Adem, S. Gitler, I. James, J. Adams, M. Mahowald, R. Bruner, M. Bendersky, B. Steer, P. Baum, W. Browder, F. Nussbaum, D. Davis, A. Berrick, D. Randall, M. Crabb, N. Singh, V. Zelov. In 1998, D. Davis in [31] presented a survey of the embedding problem for real projective spaces, and he also added a new result. For the present state of solving the immersion and embedding problem for projective spaces, we recommend to see [32].

For other families of specific manifolds, some results on immersions and embeddings in Euclidean spaces are also known for Dold manifolds (J. Ucci, T. Fukuda, C. Yoshioka, C. Olk), lens spaces (A. Berrick, J. González, T. Shimkus), real, complex or quaternionic Grassmann manifolds (H. Hiller, R. Stong, S. Hoggar, S. Ilori, V. Oproiu, V. Bartík, J. Korbaš, N. Paryjas, T. Sugawara, M. Markl, Z. Z. Tang, K. H. Mayer, K. Monks), more general real or complex flag manifolds (K. Y. Lam, R. Stong, J. Tornehave, M. Walgenbach, A. Conde, M. Percia Mendes), for real, complex or quaternionic Stiefel manifolds or projective Stiefel manifolds (H. Scheerer, N. Barufatti, D. Hacon, K. Y. Lam, P. Sankaran, P. Zvengrowski) or other homogeneous spaces (E. Rees, H. Scheerer).

Bibliography

[1] R. Abraham, J. Marsden and T. Ratiu: *Manifolds, tensor analysis, and applications* (New York: Springer-Verlag, 1988)

[2] J. F. Adams: Vector fields on spheres *Ann. of Math.* **75** (1962) 603–632

[3] Y. Ando: Existence theorems of fold-maps *Japan. J. Math. (N.S.)* **30** (2004) 29–73

[4] B. Andrews: Notes on the isometric embedding problem and the Nash-Moser implicit function theorem *In: Surveys in analysis and operator theory* (Canberra, 2001, Proc. Centre Math. Appl. Austral. Nat. Univ., 40; Canberra, Austral. Nat. Univ., 2002) 157–208

[5] E. Antoniano, S. Gitler, J. Ucci and P. Zvengrowski: On the K-theory and parallelizability of projective Stiefel manifolds *Bol. Soc. Mat. Mexicana* **31** (1986) 29–46

[6] J. Arraut and D. Randall: Index of tangent fields on compact manifolds *Contemporary Math.* **12** (1982) 31–46

[7] L. Astey, E. Micha and G. Pastor: On the parallelizability of generalized complex projective Stiefel manifolds *Proc. Roy. Soc. Edinburgh Sect. A* **133** (2003) 497–504

[8] L. Astey, S. Gitler, E. Micha, and G. Pastor: Parallelizability of complex projective Stiefel manifolds *Proc. Amer. Math. Soc.* **128** (2000) 1527–1530

[9] M. Atiyah and F. Hirzebruch: Quelques théorèmes de non-plongement pour les variétés différentiables *Bull. Soc. Math. France* **87** (1959) 383–396

[10] M. Atiyah: Immersions and embeddings of manifolds *Topology* **1** (1961) 125–132

[11] M. Atiyah: *Vector fields on manifolds* (Köln und Opladen: Westdeutscher Verlag, 1970)

[12] M. Atiyah and J. Dupont: Vector fields with finite singularities *Acta Math.* **128** (1971) 1–40

[13] M. Atiyah and I. Singer: The index of elliptic operators: I *Ann. of Math.* **87** (1968) 484–530

[14] M. Atiyah and I. Singer: The index of elliptic operators: V *Ann. of Math.* **93** (1971) 139–149

[15] R. Barre and A. El Kacimi Alaoui: Foliations *In: Handbook of differential geometry. Vol. II* (Amsterdam: Elsevier-Science, North-Holland, 2006) 35–77

[16] E. Bierstone: Equivariant Gromov theory *Topology* **13** (1974) 327–345

[17] A. Borel and F. Hirzebruch: Characteristic classes and homogeneous spaces. I; II; III *Amer. J. Math.* **80** (1958) 458–538; *Amer. J. Math.* **81** (1959) 315–382; *Amer. J. Math.* **82** (1960) 491–504

[18] A. A. Borisenko: Izometricheskie pogruzheniya prostranstvennykh form v rimanovy i psevdorimanovy prostranstva postoyannoj krivizny *Uspekhi mat. nauk* **56** (2001) 3–78;

A. A. Borisenko: Isometric immersions of space forms into Riemannian and pseudo-Riemannian spaces of constant curvature *Russian Math. Surveys* **56** (English translation), (2001) 425–497

[19] R. L. Brown: Immersions and embeddings up to cobordism *Canad. J. Math.* **23** (1971) 1102–1115

[20] R. Bryant, P. Griffiths and D. Grossman: *Exterior differential systems and Euler-Lagrange partial differential equations* (Chicago: The Chicago University Press, 2003)

[21] M. Čadek and J. Vanžura: On 2-distributions in 8-dimensional vector bundles over 8-complexes *Colloq. Math.* **70** (1996) 25–40

[22] N. S. Cardim, M. H. P. L. Mello, D. Randall and M. O. M. Silva: Join products and indices of k-fields over differentiable fibre bundles *Topology Appl.* **136** (2004) 275–291

[23] M. Cencelj, D. Repovš and A. Skopenkov: On the Browder-Levin-Novikov embedding theorems *In: Geometric topology and set theory* (Proc. Steklov Inst. Math. 247, 2004) 259–268

[24] R. Cohen: The immersion conjecture for differentiable manifolds *Ann. of Math.* **122** (1985) 237–328

[25] J.-F. Colombeau: *Elementary introduction to new generalized functions* (Amsterdam: North-Holland, 1985)

[26] M. Crabb: $\mathbb{Z}/2$*-homotopy theory* (London Math. Soc. LNS 44, Cambridge: Cambridge Univ. Press, 1980)

[27] M. Crabb: On the $KO_{\mathbb{Z}/2}$-Euler class. I and II *Proc. R. Soc. Edinb., Sect. A* **117** (1991) 115–137; *Proc. R. Soc. Edinb., Sect. A* **117** (1991) 139–154

[28] M. Crabb and B. Steer: Vector-bundle monomorphisms with finite singularities *Proc. London Math. Soc.* **30** (1975) 1–39

[29] M. Dajczer, M. Antonucci, G. Oliveira, P. Lima-Filho and R. Tojeiro: *Submanifolds and isometric immersions* (Houston, Texas: Publish or Perish, 1990)

[30] M. Datta and A. Mukherjee: Parametric homotopy principle of some partial differential relations *Math. Slovaca* **48** (1998) 411–421

[31] D. Davis: Embeddings of real projective spaces *Bol. Soc. Mat. Mexicana* **4** (1998) 115–122

[32] D. Davis: Table of embeddings and immersions of real projective spaces; `http://www.lehigh.edu/~dmd1/immtable`

[33] J.-P. Dax: Étude homotopique des espaces de plongements *Ann. Scient. Éc. Norm. Sup.* **5** (1972) 303–377

[34] J. Dupont: K-theory obstructions to the existence of vector fields *Acta Math.* **133** (1974) 67–80

[35] I. M. Gel'fand, M. I. Graev and N. Ya. Vilenkin: *Obobshchennye funktsii* (5, Moskva: Gosudarstv. izdat. fiz.-mat. lit. (1962); English translation: Generalized functions, 5, New York: Academic Press, 1966)

[36] V. Gershkovich and A. Vershik: Nonholonomic manifolds and nilpotent analysis *J. Geom. Phys.* **5** (1988) 407–452

[37] S. Gitler: Immersion and embedding of manifolds *In: Algebraic topology* (Proc. Sympos. Pure Math., Vol. 22, Univ. Wisconsin, Madison, Wis., 1970, Providence, R.I., 1971) 87–96

[38] H. Glover and G. Mislin: Immersion in the metastable range and 2-localization *Proc. Amer. Math. Soc.* **43** (1974) 443–448

[39] H. Glover and G. Mislin: Metastable embedding and 2-localization *In: Localization in group theory and homotopy theory, and related topics* (Seattle, Wash., 1974, Lecture Notes in Math., 418, Berlin: Springer, 1974) 48–57

[40] H. Glover and G. Mislin: Vector fields on 2-equivalent manifolds *In: Conference on homotopy theory* (Evanston, Ill., 1974, Sociedad Matemática Mexicana, Mexico City, 1975) 29–45

[41] H. Glover, W. Homer and R. Stong: Splitting the tangent bundle of projective space *Indiana Univ. Math. J.* **31** (1982) 161–166

[42] T. Goodwillie, J. Klein and M. Weiss: Spaces of smooth embeddings, disjunction and surgery *In: Surveys on surgery theory. Vol. 2* (Papers dedicated to C. T. C. Wall on the occasion of his 60th birthday, Princeton, NJ: Princeton University Press, 2001) 221–284

[43] P. Griffiths: *Exterior differential systems and the calculus of variations* (Progress in Math. 25, Boston, Mass.: Birkhäuser, 1983)

[44] M. L. Gromov: Stabil'nye otobrazheniya sloenij v mnogoobraziya *Izv. AN SSSR, ser. mat.* **33** (1969) 707–734;
M. L. Gromov: Stable maps of foliations into manifolds *Trans. Math. USSR Izvestia* **3** (English translation), (1969) 671–694

[45] M. L. Gromov and V. A. Rokhlin: Vlozheniya i pogruzheniya v rimanovoj geometrii *Uspekhi matem. nauk* **25** (1970) 3–62;
M. L. Gromov and V. A. Rokhlin: Embeddings and immersions in Riemannian geometry *Russ. Math. Surv.* **25** (English translation), (1970) 1–57

[46] M. Gromov: *Partial differential relations* (Berlin: Springer-Verlag, 1986)

[47] M. Grosser, M. Kunzinger, M. Oberguggenberger and R. Steinbauer: *Geometric theory of generalized functions with applications to general relativity* (Dordrecht: Kluwer, 2001)

[48] V. Guillemin and S. Sternberg: *Geometric asymptotics* (Providence R. I.: American Mathematical Society, 1977)

[49] M. Günther: On the perturbation problem associated to isometric embeddings of Riemannian manifolds *Ann. Global Anal. Geom.* **7** (1989) 69–77

[50] A. Haefliger and M. Hirsch: On the existence and classification of differentiable embeddings *Topology* **2** (1963) 129–135

[51] R. Harvey and J. Zweck: Stiefel-Whitney currents *J. Geometric Analysis* **8** (1998) 809–844

[52] R. Harvey, B. Lawson and J. Zweck: The de Rham-Federer theory of differential characters and character duality *Amer. J. Math.* **125** (2003) 791–847

[53] A. Hatcher and F. Quinn: Bordism invariants of intersections of submanifolds *Trans. Amer. Math. Soc.* **200** (1974) 327–344

[54] A. Haefliger: Differentiable imbeddings *Bull. Amer. Math. Soc.* **67** (1961) 109–112

[55] A. Haefliger: Plongements différentiables dans le domaine stable *Comment. Math. Helvetici* **37** (1962/1963) 155–176

[56] M. Hirsch: Immersions of manifolds *Trans. Amer. Math. Soc.* **93** (1959) 242–276
[57] M. Hirsch: *Differential topology* (Berlin: Springer, 1991)
[58] F. Hirzebruch and H. Hopf: Felder von Flächenelementen in 4-dimensionalen Mannigfaltigkeiten *Math. Annalen* **136** (1958) 156–172
[59] L. Hörmander: *The analysis of linear partial differential operators I. Distribution theory and Fourier Analysis* (Berlin: Springer, 1990)
[60] D. Husemoller: *Fibre bundles* (New York: McGraw Hill, 1966)
[61] H. Jacobowitz and G. Mendoza: Sub-bundles of the complexified tangent bundle *Trans. Amer. Math. Soc.* **355** (2003) 4201–4222
[62] H. Jacobowitz: Generic systems of co-rank one vector distributions *Trans. Amer. Math. Soc.* **358** (2006) 4521–4531
[63] I. James and E. Thomas: An approach to the enumeration problem for non-stable vector bundles *J. Math. Mech.* **14** (1965) 485–506
[64] I. James and E. Thomas: On the enumeration of cross-sections *Topology* **5** (1966) 95–114
[65] I. James and E. Thomas: Submersions and immersions of manifolds *Invent. Math.* **2** (1967) 171–177
[66] B. Junod and U. Suter: On the vector field problem for product manifolds *Hiroshima Math. J.* **28** (1998) 437–461
[67] U. Kaiser: Immersions in codimension 1 up to regular homotopy *Arch. Math. (Basel)* **51** (1988) 371–377
[68] U. Kaiser and B. H. Li: Enumeration of immersions of m-manifolds in $(2m - 2)$-manifolds by the singularity method *In: Differential Topology* (Siegen, 1987, Lecture Notes in Mathematics 1350, Berlin: Springer, 1988) 171–187
[69] M. Karoubi: *K-theory. An introduction* (Berlin: Springer, 1978)
[70] J. Korbaš: Vector fields on real flag manifolds *Ann. Global Anal. Geom.* **3** (1985) 173–184
[71] J. Korbaš: The rank of vector fields on real Grassmann manifolds *Rend. Circ. Mat. Palermo (II)* **41** (1992) 29–30
[72] J. Korbaš: On sectioning multiples of vector bundles and more general homomorphism bundles *Manuscripta Math.* **82** (1994) 67–70
[73] J. Korbaš and A. Szűcs: The Lyusternik-Shnirel'man category, vector bundles, and immersions of manifolds *Manuscripta Math.* **95** (1998) 289–294
[74] J. Korbaš and P. Zvengrowski: The vector field problem: A survey with emphasis on specific manifolds *Expo. Math.* **12** (1994) 3–30
[75] J. Korbaš and P. Zvengrowski: On sectioning tangent bundles and other vector bundles *In: Rend. Circ. Mat. Palermo (II)* (Proceedings of the Winter School on geometry and physics, Srní, Czech Republic, January 1994, Suppl. 39, 1996) 85–104
[76] U. Koschorke: Concordance and bordism of line fields *Inventiones Math.* **24** (1974) 241–268
[77] U. Koschorke: *Vector fields and other vector bundle morphisms – a singularity approach* (Lecture Notes in Mathematics 847, Berlin: Springer-Verlag, 1981)
[78] U. Koschorke: The singularity method and immersions of m-manifolds into manifolds of dimensions $2m - 2$, $2m - 3$ and $2m - 4$ *In: Differential Topology* (Siegen, 1987, Lecture Notes in Mathematics 1350, Berlin: Springer, 1988) 188–212

[79] U. Koschorke: Nonstable and stable monomorphisms of vector bundles *Topology Appl.* **75** (1997) 261–286
[80] U. Koschorke: Complex and real vector bundle monomorphisms *Topology Appl.* **91** (1999) 259–271
[81] L. Kristensen: On splittings of the tangent bundle of a manifold *In: Algebraic topology* (Proc. Sympos. Pure Math., Vol. 22, Univ. Wisconsin, Madison, Wis., 1970, Providence, R.I., 1971) 113–123
[82] O. Krupková: *The geometry of ordinary variational equations* (Lecture Notes in Mathematics 1678, Berlin: Springer-Verlag, 1997)
[83] M. Kunzinger and R. Steinbauer: Generalized pseudo-Riemannian geometry *Trans. Amer. Math. Soc.* **354** (2002) 4179-4199
[84] Y. Kuramoto and T. Yasui: On Haefliger's obstructions to embeddings and transfer maps *Osaka J. Math.* **40** (2003) 69–79
[85] H. Lawson and M.-L. Michelsohn: Clifford bundles, immersions of manifolds and the vector field problem *J. Differential Geom.* **15** (1980) 237–267
[86] H. Lawson and M.-L. Michelsohn: *Spin geometry* (Princeton Mathematical Series, 38, Princeton, N.J.: Princeton University Press, 1989)
[87] M. Mahowald: On obstruction theory in orientable fibre bundles *Trans. Amer. Math. Soc.* **110** (1964) 315–349
[88] J. Marsden: Generalized Hamiltonian mechanics: A mathematical exposition of non-smooth dynamical systems and classical Hamiltonian mechanics *Arch. Rational Mech. Anal.* **28** (1967/1968) 323–361
[89] Y. Matsuhita: Fields of 2-planes and two kinds of almost complex structures on compact 4-dimensional manifolds *Math. Z.* **207** (1991) 281–291
[90] K.-H. Mayer: Elliptische Differentialoperatoren und Ganzzahligkeitssätze für charakteristische Zahlen *Topology* **4** (1965) 295–313
[91] M. H. P. L. Mello: Two-plane sub-bundles of nonorientable real vector-bundles *Manuscripta Math.* **57** (1987) 263–280
[92] M. H. P. L. Mello and M. O. M. da Silva: A note on nonstable monomorphisms of vector bundles *Manuscr. Math.* **101** (2000) 191–198
[93] I. Miatello and R. Miatello: On stable parallelizability of $\tilde{G}_{k,n}$ and related manifolds *Math. Ann.* **259** (1982) 343–350
[94] J. Milnor and J. Stasheff: *Characteristic classes* (Princeton, N.J.: Princeton University Press, 1974)
[95] J. Nash: The imbedding problem for Riemannian manifolds *Ann. of Math.* **63** (1956) 20–63
[96] T. B. Ng: On the geometric dimension of vector bundles, span of a manifold and immersion of manifolds in manifolds *Expo. Math.* **8** (1990) 193–226
[97] M. Oberguggenberger: *Multiplication of distributions and applications to partial differential equations* (Harlow: Longman Scientific & Technical, 1992)
[98] H. Osborn: *Vector bundles. Vol. 1* (New York: Academic Press, 1982)
[99] E. Pap and A. Takaci: Generalized derivatives *In: Handbook of measure theory, Vol. I, II* (Amsterdam: North-Holland, 2002) 1237–1260
[100] A. Phillips: Submersion of open manifolds *Topology* **6** (1966) 171–206
[101] B. Pollina: Tangent 2-fields on even-dimensional nonorientable manifolds *Trans. Amer. Math. Soc.* **271** (1982) 215–224

[102] D. Randall: Some immersion theorems for manifolds *Trans. Amer. Math. Soc.* **156** (1971) 45–58

[103] D. Randall: Tangent frame fields on spin manifolds *Pacific J. Math.* **76** (1978) 157–167

[104] D. Randall: CAT 2-fields on nonorientable CAT manifolds *Quart. J. Math. Oxford* **38** (1987) 355–366

[105] D. Randall: On indices of tangent fields with finite singularities *In: Differential Topology* (Siegen, 1987, Lecture Notes in Mathematics 1350, Berlin: Springer, 1988) 213–240

[106] G. de Rham: *Variétés différentiables. Formes, courants, formes harmoniques* (Paris: Hermann et Cie (1955); English translation: Differentiable manifolds. Forms, currents, harmonic forms, Berlin: Springer, 1984)

[107] C. Rourke and B. Sanderson: The compression theorem I, II *Geometry and Topology* **5** (2001) 399–429; *Geometry and Topology* **5** (2001) 431–440

[108] C. Rourke and B. Sanderson: The compression theorem III *Algebraic and Geometric Topology* **3** (2003) 857–872

[109] O. Saeki: Fold maps on 4-manifolds *Comment. Math. Helvetici* **78** (2003) 627–647

[110] P. Sankaran and P. Zvengrowski: Stable parallelizability of partially oriented flag manifolds *Pacif. J. Math.* **128** (1987) 349–359

[111] P. Sankaran and P. Zvengrowski: Stable parallelizability of partially oriented flag manifolds, II *Canad. J. Math.* **20** (1997) 1-17

[112] L. Schwartz: Généralisation de la notion de fonction, de dérivation, de transformation de Fourier et applications mathématiques et physiques *Annales de l'université de Grenoble* **21** (1945) 57–74

[113] L. Schwartz: *Théorie des distributions, 1-2* (Paris: Hermann, 1950/1951)

[114] L. Schwartz: Sur l'impossibilité de la multiplication des distributions *C. R. Acad. Sci. Paris* **239** (1954) 847–848

[115] J.-M. Schlenker: Des immersions isometriques de surfaces aux varietes hyperboliques a bord convexe *In: Seminaire de Theorie Spectrale et Geometrie* (Vol. 21, Saint-Martin-d'Heres: Univ. Grenoble I, 2003) 165–216

[116] W. Singhof: Parallelizability of homogeneous spaces, I *Math. Ann.* **260** (1982) 101–116

[117] W. Singhof and D. Wemmer: Parallelizability of homogeneous spaces, II *Math. Ann.* **274** (1986) 157–176; *Erratum Math. Ann.* **276** (1987) 699–700

[118] S. Smale: The classification of immersions of spheres in Euclidean spaces *Ann. of Math.* **69** (1959) 327–344

[119] S. Smale: A survey of some recent developments in differential topology *Bull. Amer. Math. Soc.* **69** (1963) 133–145

[120] E. Spanier: *Algebraic topology* (New York: McGraw-Hill, 1966)

[121] D. Spring: The golden age of immersion theory in topology: 1959-1973. A mathematical survey from a historical perspective *Bull. Amer. Math. Soc.* **42** (2005) 163-180

[122] N. Steenrod: *The topology of fibre bundles* (Princeton, N.J.: Princeton Univ. Press, 1951)

[123] R. Stong: Subbundles of the tangent bundle *Trans. Amer. Math. Soc.* **200** (1974) 185–197

[124] R. Stong: Manifolds which immerse in small codimension *Illinois J. Math.* **27** (1983) 182–223

[125] W. Sutherland: The Browder-Dupont invariant *Proc. London. Math. Soc.* **33** (1976) 94–112

[126] Z. Z. Tang: Some existence and nonexistence results of isometric immersions of Riemannian manifolds *Commun. Contemp. Math.* **6** (2004) 867–879

[127] I. Takata: Lowest dimensions for immersions of orientable manifolds up to unoriented cobordism *Hiroshima Math. J.* **28** (1998) 481–499

[128] E. Thomas: Fields of tangent k-planes on manifolds *Inventiones Math.* **3** (1967) 334–347

[129] E. Thomas: Fields of tangent 2-planes on even-dimensional manifolds *Ann. of Math.* **86** (1967) 349–361

[130] E. Thomas: Real and complex vector fields on manifolds *J. Math. Mech.* **16** (1967) 1183–1205

[131] E. Thomas: Vector fields on manifolds *Bull. Amer. Math. Soc.* **75** (1969) 643–683

[132] E. Thomas: Embedding manifolds in Euclidean space *Osaka J. Math.* **13** (1976) 163–186

[133] W. Thurston: The theory of foliations of codimension greater than one *Comment. Math. Helvetici* **49** (1974) 214–231

[134] F. Treves: *Introduction to pseudodifferential and Fourier integral operators, 1-2* (New York: Plenum Press, 1980)

[135] V. S. Vladimirov: *Obobshchennye funktsii v matematicheskoj fizike* (Moskva: Nauka (1979); English translation: Generalized functions in mathematical physics, Moscow: Mir, 1979)

[136] F. Warner: *Foundations of differentiable manifolds and Lie groups* (New York: Springer, 1983)

[137] M. Weiss: Calculus of embeddings *Bull. Amer. Math. Soc.* **33** (1996) 177–187

[138] H. Whitney: Differentiable manifolds *Ann. of Math.* **37** (1936) 645–680

[139] H. Whitney: The selfintersections of a smooth n-manifold in $2n$-space *Ann. of Math.* **45** (1944) 220–246

[140] T. Yasui: Enumerating embeddings of n-manifolds in Euclidean $(2n-1)$-space *J. Math. Soc. Japan* **36** (1984) 555–576

[141] T. Yoshida: Parallelizability of Grassmann manifolds *Hiroshima Math. J.* **31** (1975) 13–15

[142] P. Zvengrowski: Remarks on the span of projective Stiefel manifolds *In: XI Brazilian Topology Meeting* (Rio Claro, 1998, River Edge, N.J.: World Sci. Publishing, 2000) 85–98

JÚLIUS KORBAŠ

Department of Algebra, Geometry, and Mathematical Education,
Faculty of Mathematics, Physics, and Informatics, Comenius University,
Mlynská dolina, SK-842 48 Bratislava 4, Slovakia

Mathematical Institute, Slovak Academy of Sciences,
Štefánikova 49, SK-814 73 Bratislava 1, Slovakia

E-mail: korbas@fmph.uniba.sk

Handbook of Global Analysis
Demeter Krupka and David Saunders

Geometry of differential equations

Boris Kruglikov and Valentin Lychagin

Contents

Introduction

In this paper we present an overview of geometric and algebraic methods in the study of differential equations. The latter are considered as co-filtered submanifolds in the spaces of jets, possibly with singularities. Investigation of singularities is a very subtle question, so that we will be mainly assuming regularity.

Jets are formal substitutions to actual derivatives and certain geometric structure retains this meaning, namely the Cartan distribution. Thus geometry enters differential equations. Differential-geometric methods, in particular connections and curvatures, are our basic tools.

Since differential operators form a module, (differential) algebra is also an essential component in the study of differential equations. This algebra is non-commutative, but the associated graded object is commutative, and so commutative algebra plays a central role in the investigation.

Thus we get the main ingredients and the theory is based on the interplay between them. Our exposition will center around compatibility theory, followed by formal/local (and only eventually global) integrability. So we are mainly interested in the cases, when the number of independent variables is at least two. Therefore we consider systems of partial differential equations (PDEs) and discuss methods of investigation of their compatibility, solvability or integrability.

Part of the theory is trivial for ODEs, but some methods are useful for establishing exact solutions, discovering general solutions and analysis of their singularities also for this case.

The exposition is brief and we don't prove or try to explain the results in details. The reader is referred to the cited papers/books. We have not covered some important topics, like transformation theory, equivalence problem, complete integrability, integro-differential equations etc. However our short panorama of the general theory of differential equations should help in understanding the modern progress and a possible future development.

1 Geometry of jet spaces

1.1 Jet spaces

Let us fix a smooth manifold E of dimension $m+n$ and consider submanifolds in E of the fixed codimension m. We say that two such submanifolds N_1 and N_2 are k-equivalent at a point $a \in N_1 \cap N_2$ if they are tangent (classically "have contact") of order $k \geq 0$ at this point.

Denote by $[N]_a^k$ the k-equivalence class of a submanifold $N \subset E$ at the point $a \in N$. This class is called *k-jet*[1] of N at a. Let $J_a^k(E,m)$ be the space of all k-jets of all submanifolds of codimension m at the point a and let $J^k(E,m) = \cup_{a\in E} J_a^k(E,m)$ be the space of all k-jets.

The reductions k-jets to l-jets $[N]_a^k \mapsto [N]_a^l$ gives rise to the natural projections $\pi_{k,l} : J^k(E,m) \to J^l(E,m)$ for all $k > l \geq 0$. The jet spaces carry a structure of smooth manifolds and the projections $\pi_{k,l}$ are smooth bundles.

For small values of k these bundles have a simple description. Thus $J^0(E,m) = E$ and $J^1(E,m) = \mathrm{Gr_n}(TE)$ is the Grassman bundle over E.

For each submanifold $N \subset E$ of $\operatorname{codim} N = m$ there is the natural embedding $j_k : N \to J^k(E,m)$, $N \ni a \mapsto [N]_a^k \in J^k(E,m)$ and $\pi_{k,l} \circ j_k = j_l$. The submanifolds $j_k(N) \subset J^k(E,m)$ are called the k-jet extensions of N.

Let $\pi : E_\pi \to M$ be a rank m bundle over an n-dimensional manifold. Local sections $s \in C^\infty_{\mathrm{loc}}(\pi)$ are submanifolds of the total space E_π of codimension m that are transversal to the fibres of the projection π. Let $[s]_x^k$ denote k-jet of the submanifold $s(M)$ at the point $a = s(x)$, which is also called k-jet of the section s at x.

Denote by $J_x^k(\pi) \subset \cup_{a\in\pi^{-1}(x)} J_a^k(E_\pi, m)$ the space of all k-jets of the local sections at the point $x \in M$ and by $J^k(\pi) \subset J^k(E_\pi, m)$ the space of all k-jets. $J^k(\pi)$ is an open dense subset of the latter space and thus the projections $\pi_{k,l} : J^k\pi \to J^l\pi$ form smooth fiber bundles for all $k > l$.

Projection to the base will be denoted by $\pi_k : J^k\pi \to M$. Then local sections $s \in C^\infty_{\mathrm{loc}}(M)$ have k-jet extensions $j_k(s) \in C^\infty_{\mathrm{loc}}(\pi_k)$ defined as $j_k(s)(x) = [s]_x^k$. Points of $J^k\pi$ will be also denoted by x_k and then their projections are: $\pi_{k,l}(x_k) = x_l$, $\pi_k(x_k) = x$.

If we assume π is a smooth vector bundle (without loss of generality for our purposes we'll be doing it in the sequel), then $\pi_{k,l}$ are vector bundles and the following sequences are exact:

$$0 \to S^kT^* \otimes \pi \to J^k(\pi) \overset{\pi_{k,k-1}}{\longrightarrow} J^{k-1}(\pi) \to 0,$$

where $T^* = T^*M$ is the cotangent bundle of M.

[1]The notion of jet was introduced by Ehresmann [17], though was essentially in use already in S.Lie's time [61].

Smooth maps $f : M \to N$ can be identified with sections s_f of the trivial bundle $\pi : E = M \times N \to M$ and k-jets of maps $[f]^k_x$ are k-jets of these sections. We denote the space of all k-jets of maps f by $J^k(M, N)$.

For small values of k we have: $J^0(M,N) = M \times N$ and $J^1_{(x,y)}(M,N) = \mathrm{Hom}(T_xM, T_yN)$.

For $N = \mathbb{R}$ we denote $J^k(M,\mathbb{R}) = J^k(M)$. In this case $J^1(M) = T^*M \times \mathbb{R}$. In the dual case $J^1(\mathbb{R}, M) = TM \times \mathbb{R}$. The spaces $J^k(\mathbb{R}, M)$ are manifolds of "higher velocities".

1.2 Differential groups and affine structures

Let us denote by $\mathbf{G}^k_{x,y}$ the subset of k-jets of local diffeomorphisms in $J^k_{(x,y)}(M,M)$. Then composition of diffeomorphisms defines the group structure on $\mathbf{G}^k_{x,x}$. This Lie group is called a complete *differential group of order* k (this construction is a basic example of groupoid and is fundamental for the notion of pseudogroups, see §4.5). The group $\mathbf{G}^1_{x,x}$ is the linear group $\mathrm{GL}(T_xM)$.

From §1.1 we deduce the group epimorphisms $\pi_{k,l} : \mathbf{G}^k_{x,x} \to \mathbf{G}^l_{x,x}$ for $l < k$ and exact sequences of groups for $k \geq 2$:

$$0 \to S^kT^*_x \otimes T_x \to \mathbf{G}^k_{x,x} \xrightarrow{\pi_{k,k-1}} \mathbf{G}^{k-1}_{x,x} \to 1.$$

In other words, groups $\mathbf{G}^k_{x,x}$ are extensions of the general linear group $\mathrm{GL}(T_xM)$ by means of abelian groups $S^kT^*_x \otimes T_x$, $k > 1$.

The differential groups $\mathbf{G}^k_{x,x}(M)$ act naturally on the jet spaces:

$$\mathbf{G}^k_{x,x}(M) \times J^k_x(M,m) \to J^k_x(M,m), \quad [F]^k_x \times [N]^k_x \mapsto [F(N)]^k_x.$$

The kernel $S^kT^*_x \otimes T_x$ of the projection $\mathbf{G}^k_{x,x} \to \mathbf{G}^{k-1}_{x,x}$ is the abelian group, which acts transitively on the fibre $F(x_{k-1}) = \pi^{-1}_{k,k-1}(x_{k-1})$, $x_{k-1} = [N]^{k-1}_x$, of the projection $\pi_{k,k-1} : J^k_x(M,m) \to J^{k-1}_x(M,m)$. Therefore, the fibre $F(x_{k-1})$ is an affine space. The associated vector space for the fibre is

$$S^kT^*_xN \otimes \nu_x(N),$$

where $\nu_x(N) = T_xM/T_xN$ is the *normal space* to N at the point $x \in N$.

Thus, the bundle $\pi_{k,k-1} : J^k(M,m) \to J^{k-1}(M,m)$ has a canonical *affine structure* for $k \geq 2$. Moreover, each local diffeomorphism $F : M \to M$ has the natural lifts to local diffeomorphisms $F^{(k)} : J^k(M,m) \to J^k(M,m)$, $[N]^k_x \mapsto [F(N)]^k_{F(x)}$ preserving the affine structures.

For a vector bundle $\pi : E_\pi \to M$ the affine structure in the fibers of $J^k\pi \to J^{k-1}\pi$ coincides with the structure induced by the vector bundle structure. If π is a fiber bundle, the preceding construction provides the affine structure. This gives rise to the following construction.

Let $M \subset E$ be a submanifold of codimension m and $U \supset M$ be its neighborhood, which is transversally foliated, so that the projection along the fibers $\pi : U \to M$ can be identified with the normal bundle. We can denote $U = E_\pi$. Then the embedding $\varkappa : E_\pi \subset E$ induces the embedding $\varkappa^{(k)} : J^k(\pi) \hookrightarrow J^k(E,m)$ with $M^{(k)}$ being the zero section. The image is an open neighborhood and the affine structure on $J^k(\pi) \to J^{k-1}(\pi)$ induces the affine structure on $J^k(E,m) \to J^{k-1}(E,m)$, so that both projections agree and are denoted by the same symbol $\pi_{k,k-1}$. These neighborhoods $J^k(\pi) \hookrightarrow J^k(E,m)$ together with the maps $\varkappa^{(k)}$ are called *affine charts*.

Remark 1 Usage of affine charts in general jet-spaces is the exact analog of independence condition in exterior differential systems. Most of the theory works for spaces $J^k(E,m)$, though for simplicity we will often restrict to the case of jets of sections $J^k\pi$.

1.3 Cartan distribution

In addition to the affine structure on the co-filtration $\pi_{k,k-1} : J^k(E,m) \to J^{k-1}(E,m)$, the space $J^k(E,m)$ bears an additional structure, which allows to distinguish submanifolds $j_k(N) \subset J^k(E,m)$, $N \subset E$, among all submanifolds in E of dimension $n = \dim N$. To describe it denote

$$L(x_{k+1}) = T_{x_k}[j_k(N)] \subset T_{x_k}J^k(\pi), \qquad x_{k+1} = [N]_x^{k+1}$$

(this subspace does not depend on a particular choice of N, but only on x_{k+1}). Define the *Cartan distribution* on the space $J^k(E,m)$ by the formula:

$$\mathcal{C}_k(x_k) = \operatorname{span}\{L(x_{k+1}) : x_{k+1} \in \pi^{-1}_{k+1,k}(x_k)\} = (d\pi_{k,k-1})^{-1}L(x_k).$$

Submanifolds of the form $N^{(k)}$ are clearly integral manifolds of the Cartan distributions such that $\pi_{k,0} : N^{(k)} \to E$ are embeddings. The inverse is also true: if $W \subset J^k(E,m)$ is an integral submanifold of dimension n of the Cartan distribution such that $\pi_{k,0} : W \to E$ is an embedding, then $W = N^{(k)}$ for the submanifold $N = \pi_{k,0}(W) \subset E$. In other words, the Cartan distribution gives a geometrical description for the jet-extensions.

In a similar way one can construct the Cartan distributions for the jet spaces $J^k(\pi)$. Moreover, any affine chart $\varkappa^{(k)} : J^k(\pi) \to J^k(E,m)$ sends the Cartan distribution on $J^k(\pi)$ to the Cartan distribution on $J^k(E,m)$. By using this observation we can restrict ourselves to Cartan distributions on the jet-spaces of sections.

For the case $J^k\pi$, there is a description of the Cartan distribution in terms of differential forms. Namely, let us denote by $\Omega_0^r(J^k\pi)$ the module of π_k-horizontal forms, that is, such differential r-forms ω that $i_X\omega = 0$ for any π_k-vertical vector field X: $d\pi_k(X) = 0$.

These forms can be clearly identified with non-linear differential operators[2] $\operatorname{diff}_k(\pi, \Lambda^r T^*M)$ acting from sections of π to differential r-forms on the manifold M. Indeed the space of such non-linear operators is nothing else than the space of smooth maps $C^\infty(J^k\pi, \Lambda^r T^*M)$.

The composition with the exterior differential $d : \Omega^r(M) \to \Omega^{r+1}(M)$ generates the *total differential* $\widehat{d} : \Omega_0^r(J^k\pi) \to \Omega_0^{r+1}(J^{k+1}\pi)$. The total differential is a differentiation of degree 1 and it satisfies the property $\widehat{d}^{\,2} = 0$.

Hence any function $f \in C^\infty(J^{k-1}\pi)$ defines two differential forms on the jet-space $J^k(\pi)$: $\widehat{d}f \in \Omega_0^1(J^k\pi)$ and $d(\pi^*_{k,k-1}f) = \pi^*_{k,k-1}(df) \in \Omega^1(J^k\pi)$. Both of them coincide on k-jet prolongations $j_k(s)$. Their difference:

$$U(f) = d(\pi^*_{k,k-1}f) - \widehat{d}f \in \Omega^1(J^k\pi)$$

[2]These will be treated in §2.4. We introduce here only a minor part of the theory.

is called the *Cartan form* associated with function $f \in C^\infty(J^{k-1}\pi)$.

The annihilator of the Cartan distribution on $J^k\pi$ is generated by the Cartan forms: $\operatorname{Ann}\mathcal{C}_k(x_k) = \operatorname{span}\{U(f)_{x_k} : f \in C^\infty(J^{k-1}\pi)\}$.

As an example consider the case $m = 1$, $k = 1$. Then the Cartan distribution on $J^1(E, 1)$ is the classical contact structure on the space of contact elements. It is known that it cannot be defined by one differential 1-form. On the other hand, for the affine chart $J^1(M) = T^*M \times \mathbb{R}$ the Cartan distribution (=the standard contact structure) can be defined by one Cartan form $U(u) = du - p\,dq$, where $u : J^1(M) \to \mathbb{R}$ is the natural projection and $p\,dq$ is the Liouville form on T^*M.

1.4 Lie transformations

Any local diffeomorphism $F : E \to E$ has prolongations $F^{(k)} : J^k(E, m) \to J^k(E, m)$, $[N]^k_x \mapsto [F(N)]^k_{F(x)}$, and they satisfy: $(F \circ G)^{(k)} = F^{(k)} \circ G^{(k)}$, $\pi_{k,k-1} \circ F^{(k)} = F^{(k-1)} \circ \pi_{k,k-1}$. Moreover, by the construction, the diffeomorphisms $F^{(k)}$ are symmetries of the Cartan distribution, i.e. they preserve $\mathcal{C}_k$.

For $m = 1$ the Cartan distribution on the 1-jet space $J^1(E, 1)$ defines the contact structure, and not all contact diffeomorphisms have the form $F^{(1)}$, where $F : E \to E$. Let $\phi : J^1(E, 1) \to J^1(E, 1)$ be a contact local diffeomorphism and let $x_k = [N]^k_x$. We can consider this point as $(k-1)$-jet of an integral manifold $N^{(1)}$ at the point $x_1 = [N]^1_x$. Then $\phi(N^{(1)})$ is an integral manifold of the contact structure, and it has the form $N^{(1)}_\phi$ for some submanifold $N_\phi \subset E$ if $\pi_{1,0} : \phi(N^{(1)}) \to E$ is an embedding.

Denote by $\Sigma_\phi \subset J^1(E, 1)$ the set of points x_1, where the last condition is not satisfied. Then, for the points $x_k \in J^k(E, 1)$, such that projections $x_1 = \pi_{k,1}(x_k)$ belong to the compliment Σ^c_ϕ, we can define the lift $\phi^{(k-1)} : J^k(E, 1) \to J^k(E, 1)$, $[N]^k_x \mapsto [\phi(N^{(1)})]^{(k-1)}_{\phi(x_1)}$. As before we get:

$$(\phi \circ \psi)^{(k-1)} = \phi^{(k-1)} \circ \psi^{(k-1)}, \quad \pi_{k,k-1} \circ \phi^{(k-1)} = \phi^{(k-2)} \circ \pi_{k,k-1}.$$

Diffeomorphisms $F : E \to E$ are also called *point transformations*. So the local diffeomorphisms $F^{(k)}$ and $\phi^{(k-1)}$ are called prolongations of the point transformation F or the contact transformation ϕ respectively.

A local diffeomorphism of $J^k(E, m)$ preserving the Cartan distribution is called a *Lie transformation*. The following theorem is known as Lie-Backlund theorem on prolongations, see [40].

Theorem 1 *Any Lie transformation of $J^k(E, m)$ is the prolongation of*

$m \geq 2$: *Local point transformation $F : E \to E$,*

$m = 1$: *Local contact diffeomorphism $\phi : J^1(E, 1) \to J^1(E, 1)$.*

In the same way one can construct prolongations of vector fields on E and contact vector fields on $J^1(E, 1)$ to $J^k(E, m)$ or $J^k(E, 1)$ respectively and the prolongations preserve the Cartan distribution. A vector field on $J^k(E, m)$, which preserves the Cartan distribution, is called a *Lie vector field*. The Lie-Backlund theorem claims that Lie vector fields are prolongations of vector fields on E if $m \geq 2$ or contact vector fields on $J^1(E, 1)$ if $m = 1$.

The same statements hold for $E = E_\pi$, when the jet-space is $J^k\pi$.

Remark 2 Prolongations $F^{(k)}$ of the point transformations preserve the affine structure for any $k \geq 2$, i.e. starting from the 2nd jets. The prolongations $\phi^{(k)}$ of the contact transformations also preserve the affine structure for $k \geq 2$, but this means starting from the 3rd jets.

Let us briefly introduce here systems of PDEs[3]. Such a system of pure order k is represented as a smooth subbundle $\mathcal{E} \subset J^k(\pi)$. It is possible to use more general jet-spaces $J^k(E,m)$; exteriour differential systems concern the case $k = 1$. Scalar PDEs correspond to the trivial bundle $E_\pi = M \times \mathbb{R}$ and $\mathcal{E} \subset J^k(M)$.

Solutions of $\mathcal{E}$ on an open set $U_M \subset M$ are sections $s \in C^\infty_{(\text{loc})}(\pi)$ such that $j_k(s)(U_M) \subset \mathcal{E}$. Generalized solutions are n-dimensional integral manifolds W^n of the Cartan distribution such that $W \subset \mathcal{E}$ (in this form there's no difference with equations in the general jet-space $J^k(E,m) \supset \mathcal{E}$). If $\pi_{k,0} : W \to M$ is not an embedding, we call such solution multi-valued or singular.

Another description of generalized solutions are n-dimensional integral manifolds of the induced Cartan distribution $\mathcal{C}_\mathcal{E} = \mathcal{C}_k \cap T\mathcal{E}$. Then internal Lie transformations (finite or infinitesimal) are (local) diffeomorphisms of $\mathcal{E}$ that are symmetries of $\mathcal{C}_\mathcal{E}$ (they transform generalized solutions to generalized solutions [62, 63]).

In general there exist higher internal Lie transformations, which are not prolongations from lower-order jets. But for certain type of systems $\mathcal{E}$ we have the exact analog of Lie-Backlund theorem, see [40].

1.5 Calculations

A coordinate system (x^i, u^j) on E_π, subordinated to the bundle structure, induces coordinates (x^i, p^j_σ) on $J^k\pi$, where multiindex $\sigma = (i_1, \dots, i_n)$ has length $|\sigma| = i_1 + \cdots + i_n \leq k$ and $p^j_\sigma\left([s]^k_x\right) = \dfrac{\partial^{|\sigma|} s^j}{\partial x^\sigma}(x)$.

For a vector field $X \in \mathcal{D}(M)$ the operator of total derivative along X is $\mathcal{D}_X = i_X \circ \widehat{d} : C^\infty(J^k\pi) \to C^\infty(J^{k+1}\pi)$ (this is just a post-composition of a differential operator with Lie derivative along X) and it has the following expression. Let $X = \sum \xi^i \partial_{x^i}$. Then $\mathcal{D}_X = \sum \xi^i \mathcal{D}_i$, where the basis total derivation operator $\mathcal{D}_i = \mathcal{D}_{\partial_{x^i}}$ is given by infinite series

$$\mathcal{D}_i = \partial_{x^i} + \sum p_{\sigma+1_i} \partial_{p_\sigma}.$$

If in the above sum we restrict $|\sigma| < k$ we get vector fields $\mathcal{D}^{(k)}_i$ on $J^k\pi$. In terms of them the Cartan distribution on J^kM is given by

$$\mathcal{C}_k = \langle \mathcal{D}^{(k)}_i, \partial_{p_\sigma} \rangle_{1 \leq i \leq n, |\sigma| = k} \,.$$

To write it via differential forms note that the operator of total derivative equals $\widehat{d} = \sum \mathcal{D}_i \otimes dx^i$. Thus for $f \in C^\infty(J^{k-1}\pi)$ we get expression for the Cartan forms $U(f) = \sum_{i,j;|\sigma|<k} \big((\partial_{p^j_\sigma} f)\, dp^j_\sigma + (\partial_{x^i} f - \mathcal{D}_i f)\, dx^i \big)$.

In particular, the differential forms $\omega^j_\sigma = U(p^j_\sigma) = dp^j_\sigma - \sum p^j_{\sigma+1_i} dx^i$ span the annulator of the Cartan distribution, i.e. $\mathcal{C}_k = \operatorname{Ker}\{\omega^j_\sigma\}_{0 \leq |\sigma| < k}$.

[3] Main definitions come only in §2.3,§3.4 after development of algebraic machinery.

Finally let us express in coordinates Lie infinitesimal transformations.

Vector field $X = \sum_i a^i(x,u)\,\partial_{x^i} + \sum_j b^j(x,u)\,\partial_{u^j}$ on $E = J^0\pi$ (point transformation) prolongs to

$$X^{(k)} = \sum_i a^i(x,u)\,\mathcal{D}_i^{(k+1)} + \sum_{j;|\sigma|\le k} \mathcal{D}_\sigma(\varphi^j)\,\partial_{p^j_\sigma}, \tag{1}$$

where $\varphi^j = b^j - \sum_{i=1}^n a^i p^j_i$ are components of the so-called *generating function* $\varphi = (\varphi^1,\dots,\varphi^r)$. Though the coefficients of (1) depend seemingly on the $(k+1)$-jets, the Lie field belongs in fact to $\mathcal{D}(J^k\pi)$.

A contact vector field $X^{(1)} = X_\varphi$ on $J^1\pi$ is determined by generating scalar-valued function $\varphi = \varphi(x^i,u,p_i)$ via the formula

$$X^{(1)} = \sum_i \Bigl[\mathcal{D}_i^{(1)}(\varphi)\,\partial_{p_i} - \partial_{p_i}(\varphi)\,\mathcal{D}_i^{(1)}\Bigr] + \varphi\,\partial_u.$$

The prolongation of this field to $J^k\pi$ is given by the formula similar to (1):

$$X^{(k)} = -\sum_i \partial_{p_i}(\varphi)\,\mathcal{D}_i^{(k+1)} + \sum_{|\sigma|\le k} \mathcal{D}_\sigma(\varphi)\,\partial_{p_\sigma}. \tag{2}$$

Again this is a vector field on $J^k\pi$, coinciding with X_φ for $k=1$.

1.6 Integral Grassmanians

Denote

$$I_0(x_k) = \{L(x_{k+1}) : x_{k+1} \in F(x_k)\} \subset \mathrm{Gr}_n(T_{x_k}J^k)$$

the Grassmanian of all tangent planes to jet-sections through x_k. The letter I indicates that this can be represented as the space of integral elements. Consider for simplicity the space of jets of sections of a vector bundle π.

The map $C^\infty(J^{k-1}\pi) \ni f \mapsto dU(f)|_{\mathcal{C}(x_k)} \in \Lambda^2(\mathcal{C}^*(x_k))$ is a derivation and therefore defines a linear map $\Omega_{x_k} : T^*_{x_{k-1}}(J^{k-1}\pi) \to \Lambda^2(\mathcal{C}^*(x_k))$. Since the latter vanishes on $\mathrm{Im}(d\pi^*_{k-1,k-2})$ it descends to the linear map

$$\Omega_{x_k} : S^{k-1}T_x \otimes \pi^*_x \to \Lambda^2\mathcal{C}(x_k)^*,$$

which is called the *metasymplectic structure* on the Cartan distribution. We treat Ω_{x_k} as a 2-form on $\mathcal{C}(x_k)$ with values in $F(x_{k-2}) \simeq S^{k-1}T^*_x \otimes \pi_x$.

Remark that for the trivial rank 1 bundle $\pi = \mathbf{1}$ and $k=1$ the metasymplectic structure Ω_{x_1} on $J^1(M)$ coincides with the symplectic structure on the Cartan distribution induced by the contact structure.

Call a subspace $L \subset \mathcal{C}(x_k)$ *integral* if $\Omega_{x_k}|_L = 0$. Then $I(x_k)$ consists of all integral n-dimensional spaces for Ω_{x_k} and $I(x_k) \supset I_0(x_k)$. Denote

$$I_l(x_k) = \{L \in I(x_k) : \dim\,(\pi_k)_*(L) = n - l\} = \{L : \dim\,\mathrm{Ker}((\pi_k)_*|_L) = l\}.$$

Elements $I_0\,(x_k)$ are called *regular*; they correspond to tangent spaces of the usual smooth solutions (jet-extensions of sections). The others are tangent spaces of singular (multi-valued) solutions.

The difference $\dim I_0(x_k) - \dim I_l(x_k)$ depends on m, k, l only. We denote this number by $c_{m,k,l}$ and call *formal codimension* of $I_l(x_k)$. Usually this number is negative. The only cases when $m \leq 4$ and $c > 0$ are listed in the following tables:

m = 1

$k \backslash l$	1	2	3	4
1	1	3	6	10
2	1	2	1	**-4**
3	1	1	**-6**	**-29**
4	1	0	**-15**	**-48**
5	1	**-1**	**-26**	**-124**

m = 2

$k \backslash l$	1	2	3	4
1	1	2	3	4
2	1	0	**-7**	**-24**
3	1	**-2**	**-21**	**-74**

m = 3

$k \backslash l$	1	2	3	4
1	1	1	0	**-2**
2	1	**-2**	**-15**	**-44**

m = 4

$k \backslash l$	1	2	3
1	1	0	**-3**
2	1	**-4**	**-23**

These tables show that the regular cell $I_0(x_k)$, as a rule, has smaller dimension than $I(x_k)$. Indeed $c \geq 0$ iff $l = 1$ or $l = 2$ & $km \leq 4$ or $l > 2$ & $k + m \leq 4$. This means that most elements of $I(x_k)$ are not tangent planes to multi-valued jet-extensions $j_k s$ with singularities of projection on the set of measure zero.

To avoid paradoxical integral planes we introduce the notion of R-Grassmanian and R-spaces. By R-Grassmanian $RI(x_k)$ we mean the closure of the regular cell $I_0(x_k)$ in $I(x_k)$. Its elements are called R-spaces.

When $RI(x_k) \neq I(x_k)$ (which is often the case by the above mentioned dimensional reasons), then there are integral manifolds, which represent singular solutions such that "all small deformations in the class of integral manifolds" have singularities too!

This leads to the notion of an R-manifold, which is an n-dimensional integral manifold N of the Cartan distribution with all tangent spaces $T_{x_k}N$, $x_k \in N$, being R-spaces. For a more detailed description of R-spaces and R-manifolds based on the associated Jordan algebra structures consult [65].

Remark 3 This discussion makes important distinction between exterior differential systems and systems of PDEs embedded in jets. While with the first approach a system is given just as a subbundle in a Grassmanian, the second case keeps algebraic structures visible, in particular structure of integral manifolds is graspable and stratification of singularities is prescribed.

Notice that $c_{m,k,l} = l^2 - mk\binom{l+k-1}{k+1}$ ([65]). Since $\binom{l+k-1}{k+1} \sim \frac{1}{(k+1)!}\,l^{k+1}$, we observe that only for $k = 1$, $m \leq 2$ the formal codimensions of I_l are non-negative for all l. These numbers are $c_{1,1,l} = \frac{l(l-1)}{2}$ and $c_{2,1,l} = l$.

For $k = m = 1$ we have Legendrian Grassmanian $I(x_1) \subset \mathrm{Gr}_n(T_{x_1}J^1(E,1))$. Its restriction to the affine chart $\tilde{I}(x_1) \subset \mathrm{Gr}_n(T_{x_1}(T^*M \times \mathbb{R}))$ induces via projection the

Lagrangian Grassmanian $\mathrm{LG}(a) \subset \mathrm{Gr}_n(T_a(T^*M))$, $a = (x, p)$.

The case $k = 1$, $m = 2$ includes complex Grassmanians. All other integral Grassmanians are singular (with regard to the stratum $I_0(x_k)$).

The topological structure of integral Grassmanians is important in investigation of singularities of solutions.

Theorem 2 *[65]. The cohomology ring $H^*(I(x_k), \mathbb{Z}_2)$ is isomorphic to the polynomial ring $\mathbb{Z}_2[w_1^k, \dots, w_n^k]$ up to dimension n, where $w_1^k, \dots, w_n^k$ are Stiefel-Whitney classes of the tautological bundle over $I(x_k)$.*

For a system of differential equations $\mathcal{E}$ of order k its integral Grassmanian is $I\mathcal{E}(x_k) = I(x_k) \cap \mathrm{Gr}_n(T_{x_k}\mathcal{E})$. In other words, if $\mathcal{C}_\mathcal{E}$ is the Cartan structure on $\mathcal{E}$ and $\Omega_\mathcal{E} = \Omega_{x_k}|_{T_{x_k}\mathcal{E}} \in \Lambda^2\mathcal{C}_\mathcal{E}^* \otimes F(x_{k-2})$ is the restriction of the metasymplectic structure, then $I\mathcal{E}(x_k)$ coincides with the space of all integral n-dimensional planes for $\Omega_\mathcal{E}$.

Note that the tangent spaces to solutions of the system are integral spaces. Thus description of integral Grassmanians of systems of PDEs is important for investigation of solutions (remark that fixation of the subbundle $I\mathcal{E}(x_k)$ is essentially the starting point in EDS approach). We shall return to this problem in §3.7.

2 Algebra of differential operators

2.1 Linear differential operators

Denote by $\mathbf{1}$ the trivial one-dimensional bundle over M . Let $\mathcal{A}_k = \mathrm{Diff}_k(\mathbf{1}, \mathbf{1})$ be the $C^\infty(M)$-module of scalar linear differential operators of order $\le k$ and $\mathcal{A} = \cup_k \mathcal{A}_k$ be the corresponding filtered algebra, $\mathcal{A}_k \circ \mathcal{A}_l \subset \mathcal{A}_{k+l}$.

Notice that the associated graded algebra $\mathrm{gr}(\mathcal{A}) = \oplus \mathcal{A}_{k+1}/\mathcal{A}_k$ is the symmetric power of the tangent bundle:

$$\mathrm{gr}(\mathcal{A}) = ST = \oplus_i S^i T, \quad \text{where } T = T_x M.$$

Consider two linear vector bundles π and ν. Denote by $\mathrm{Diff}(\pi, \nu) = \cup_k \mathrm{Diff}_k(\pi, \nu)$ the filtered module of all linear differential operators from $C^\infty(\pi)$ to $C^\infty(\nu)$. We have the natural pairing

$$\mathrm{Diff}_k(\rho, \nu) \times \mathrm{Diff}_l(\pi, \rho) \to \mathrm{Diff}_{k+l}(\pi, \nu)$$

given by the composition of differential operators.

In particular, $\mathrm{Diff}(\pi, \mathbf{1})$ is a filtered left $\mathcal{A}$-module, $\mathrm{Diff}(\mathbf{1}, \pi)$ is a filtered right $\mathcal{A}$-module and they have an $\mathcal{A}$-valued $\mathcal{A}$-linear pairing

$$\Delta \in \mathrm{Diff}_l(\pi, \mathbf{1}),\ \nabla \in \mathrm{Diff}_k(\mathbf{1}, \pi) \mapsto \langle \Delta, \nabla \rangle = \Delta \circ \nabla \in \mathcal{A}_{k+l},$$

with $\langle \theta \circ \Delta, \nabla \rangle = \theta \circ \langle \Delta, \nabla \rangle$, $\langle \Delta, \nabla \circ \theta \rangle = \langle \Delta, \nabla \rangle \circ \theta$ for $\theta \in \mathcal{A}$.

Each linear differential operator $\Delta : C^\infty(\pi) \to C^\infty(\nu)$ of order l induces a right $\mathcal{A}$-homomorphism $\phi_\Delta : \mathrm{Diff}(\mathbf{1}, \pi) \to \mathrm{Diff}(\mathbf{1}, \nu)$ by the formula:

$$\mathrm{Diff}_k(\mathbf{1}, \pi) \ni \nabla \mapsto \Delta \circ \nabla \in \mathrm{Diff}_{k+l}(\mathbf{1}, \nu).$$

Its $\langle\, , \rangle$-dual is a left $\mathcal{A}$-homomorphism $\phi^\Delta : \mathrm{Diff}(\nu, \mathbf{1}) \to \mathrm{Diff}(\pi, \mathbf{1})$ given by

$$\mathrm{Diff}_k(\nu, \mathbf{1}) \ni \square \mapsto \square \circ \Delta \in \mathrm{Diff}_{k+l}(\pi, \mathbf{1}).$$

Denoting by $\mathcal{J}^k(\pi) = C^\infty(\pi_k)$ the space of (non-holonomic) sections of the jet-bundle we have:

$$\mathrm{Diff}_k(\pi,\nu) = \mathrm{Hom}_{C^\infty(M)}(\mathcal{J}^k(\pi), C^\infty(\nu)), \tag{3}$$

and differential operators Δ of order k are in bijective correspondence with morphisms $\psi^\Delta : J^k(\pi) \to \nu$ via the formula $\Delta = \psi^\Delta \circ j_k$, where $j_k : C^\infty(\pi) \to \mathcal{J}^k(\pi)$ is the jet-section operator.

The prolongation $\psi_l^\Delta : J^{k+l}(\pi) \to J^l(\nu)$ of $\psi^\Delta = \psi_0^\Delta$ is conjugated to the $\mathcal{A}$-homomorphism $\phi^\Delta : \mathrm{Diff}_l(\nu, \mathbf{1}) \to \mathrm{Diff}_{k+l}(\pi, \mathbf{1})$ via isomorphism (3). This makes a geometric interpretation of the differential operator Δ as the bundle morphism.

Similarly one can interpret the $\mathcal{A}$-homomorphism $\phi_\Delta : \mathrm{Diff}_l(\mathbf{1}, \pi) \to \mathrm{Diff}_{k+l}(\mathbf{1}, \nu)$, see [40]. More generally lift of the operator Δ, obtained via post-composition, is $\hat{\Delta} : \mathrm{Diff}_l(\xi, \pi) \to \mathrm{Diff}_{k+l}(\xi, \nu)$.

2.2 Prolongations, linear PDEs and formal integrability

A system $\mathcal{E}$ of PDEs of order k associated to an operator $\Delta \in \mathrm{Diff}_k(\pi, \nu)$ is, by definition, the subbundle $\mathcal{E}_k = \mathrm{Ker}(\psi^\Delta) \subset J^k(\pi)$. Its prolongation is $\mathcal{E}_{k+l} = \mathcal{E}_k^{(l)} = \mathrm{Ker}(\psi_l^\Delta) \subset J^{k+l}(\pi)$.

If $\nu = r \cdot \mathbf{1}$ is the trivial bundle of $\mathrm{rank}\,\nu = r$, we can identify $\Delta = (\Delta_1, \ldots, \Delta_r)$ to be a collection of scalar operators. Then the system $\mathcal{E}_{k+l}$ is given by the equations $\mathcal{D}_\sigma \circ \Delta_j\,[u(x)] = 0$, where $1 \le j \le r$, $\sigma = (i_1, \ldots, i_n)$ is a multi-index of length $|\sigma| = \sum i_s \le l$ and $\mathcal{D}_\sigma = \mathcal{D}_1^{i_1} \cdots \mathcal{D}_n^{i_n}$.

Points of $\mathcal{E}_k$ can be identified as k-jet solutions (not k-jets of solutions!) of the system $\Delta = 0$ and the points of $\mathcal{E}_{k+l}$ are $(k+l)$-jet solutions of the l-prolonged system. Formal solutions are points of $\mathcal{E}_\infty = \varprojlim \mathcal{E}_i$.

Not all the points from $\mathcal{E}_k$ can be prolonged to $(k+l)$-jet solutions, but only those from $\pi_{k+l,k}(\mathcal{E}_{k+l}) \subset \mathcal{E}_k$. Investigation of these as well as formal solutions can be carried successively in l and we arrive to

Definition 1 System $\mathcal{E}$ is formally integrable if $\mathcal{E}_i$ are smooth manifolds and the maps $\pi_{i+1,i} : \mathcal{E}_{i+1} \to \mathcal{E}_i$ are vector bundles.

Define the dual $\mathcal{E}^* = \mathrm{Coker}\,\phi^\Delta$ as the collection of spaces $\mathcal{E}_i^*$ given by the exact sequence:

$$\mathrm{Diff}_k(\nu, \mathbf{1}) \xrightarrow{\phi_k^\Delta} \mathrm{Diff}_{k+l}(\pi, \mathbf{1}) \to \mathcal{E}_{k+l}^* \to 0.$$

We endow the dual $\mathcal{E}^*$ with natural maps $\pi_{i+1,i}^* : \mathcal{E}_i^* \to \mathcal{E}_{i+1}^*$. But it becomes an $\mathcal{A}$-module only when these maps are injective.

However in general we can define the inductive limit $\mathcal{E}^\Delta = \varinjlim \mathcal{E}_i^*$. It is a filtered left $\mathcal{A}$-module. Thus we can consider the system as a module over differential operators ($\mathcal{D}$-module).

The dual $\mathcal{E}_\Delta = \mathrm{Ker}(\phi_\Delta) \subset \mathrm{Diff}(\mathbf{1}, \pi)$ is a right $\mathcal{A}$-module and we have the pairing $\mathcal{E}^\Delta \times \mathcal{E}_\Delta \to \mathcal{A}$. For formally-integrable systems this pairing is non-degenerate, as follows from the following statement:

Proposition 3 *A system $\mathcal{E} = \mathrm{Ker}(\psi^\Delta)$ is formally integrable iff $\mathcal{E}_i^*$ are projective $C^\infty(M)$-modules and the maps $\pi_{i+1,i}^* : \mathcal{E}_i^* \to \mathcal{E}_{i+1}^*$ are injective.*

Proof The projectivity condition is equivalent to regularity (constancy of rank of the projection $\pi_{i+1,i}$), while invjectivity of $\pi^*_{i+1,i}$ is equivalent to surjectivity of $\pi_{i+1,i}$. □

If we have several differential operators $\Delta_i \in \mathrm{Diff}(\pi, \nu_i)$ of different orders k_i, $1 \le i \le t$, then their sum is no longer a differential operator of pure order $\Delta = (\Delta_1, \dots, \Delta_t) : C^\infty(\pi) \to C^\infty(\nu)$, $\nu = \oplus \nu_i$. Thus ϕ^Δ is not an $\mathcal{A}$-morphism, unless we put certain weights to the graded components ν_i.

Namely if we introduce weight k_i^{-1} for the operator Δ_i (equivalently to the bundle ν_i), then the operator Δ becomes an $\mathcal{A}$-homomorphism of degree 1. This allows to treat formally the systems of different orders via the same algebraic machinery as for the systems of pure order k. Geometric approach will be explained in the next section.

The prolongation theory wholly transforms for systems of PDEs $\mathcal{E}$ of different orders. In particular for formally integrable systems we have left $\mathcal{A}$-module $\mathcal{E}^*$. It is not a bi-module, but one can investigate sub-algebras $\mathcal{S} \subset \mathcal{A}$, which act on $\mathcal{E}^*$ from the right. It will be clear from §4.4 that they correspond to symmetries of the system $\mathcal{E}$.

2.3 Symbols, characteristics and non-linear PDEs

Consider symbolic analogs of the above modules (we will write sometimes $T = T_x M$ for brevity). Since $ST \otimes \pi^* = \oplus S^i T \otimes \pi^*$ is the graded module associated to the filtrated $C^\infty(M)$-module $\mathrm{Diff}(\pi, \mathbf{1}) = \cup \mathrm{Diff}_i(\pi, \mathbf{1})$, the bundle morphism ϕ^Δ produces the graded homomorphisms, called symbols of our differential operator Δ:

$$\sigma_\Delta : ST \otimes \nu^* \to ST \otimes \pi^*.$$

The value $\sigma_{\Delta,x}$ of σ_Δ at $x \in M$ is a homomorphism of ST-modules.

The ST-module $\mathcal{M}_\Delta = \mathrm{Coker}(\sigma_{\Delta,x})$ is called the *symbolic module* at the point $x \in M$ ([28]). Its annihilator is called the *characteristic ideal* $I(\Delta) = \oplus I_q$, where I_q are homogeneous components. The set of covectors $p \in T^* \setminus \{0\}$ annihilated by $I(\Delta)$ is the *characteristic variety* $\mathrm{Char}_{\mathrm{aff}}(\Delta)$. We will consider projectivization of this conical affine variety $\mathrm{Char}_x(\Delta) \subset \mathbb{P}T^*$.

It is often convenient to work over complex numbers. If we complexify the symbolic module, we get the complex characteristic variety

$$\mathrm{Char}^{\mathbb{C}}_x(\Delta) = \{p \in \mathbb{P}^{\mathbb{C}} T^* \mid f(p^q) = 0 \,\forall f \in I_q, \forall q\}.$$

Proposition 4 [23, 88]. *For $p \in T^*_x M \setminus \{0\}$ let $\mathfrak{m}(p) = \oplus_{i>0} S^i T \subset ST$ be the maximal ideal of homogeneous polynomials vanishing at p. Then covector p is characteristic iff the localization $(\mathcal{M}_\Delta)_{\mathfrak{m}(p)} \neq 0$.*

The set of localizations $(\mathcal{M}_\Delta)_{\mathfrak{m}(p)} \neq 0$ for characteristic covectors p form the *characteristic sheaf* $\mathcal{K}$ over the characteristic variety $\mathrm{Char}^{\mathbb{C}}_x(\Delta)$.

The above definitions work for systems $\mathcal{E}$ of different order PDEs if we impose the weight-convention of the previous section. However it will be convenient to interpret this case geometrically and such approach works well even in non-linear situation.

A system of PDEs of pure order k is represented as a smooth subbundle $\mathcal{E}_k \subset J^k(\pi)$, non-linear case corresponds to fiber-bundles (in regular situation; in general the fibers $\pi^{-1}_{k,k-1}(*) \cap \mathcal{E}_k$ are not diffeomorphic and $\mathcal{E}_k$ is just a submanifold in $J^k\pi$). Prolongations are defined by the formula

$$\mathcal{E}_k^{(1)} = \{x_{k+1} = [s]^{k+1}_x \in J^{k+1}\pi : T_{x_k}[j_k s(M)] \subset T_{x_k}\mathcal{E}\}.$$

Higher prolongations can be defined successively in regular case, but in general $\mathcal{E}_k^{(l)}$ equals the set of all points $x_{k+l} = [s]_x^{k+l} \in J^{k+l}\pi$ with the property that $j_k s(M)$ is tangent to $\mathcal{E}$ at x_k with order $\geq l$.

To cover the case of several equations of different orders we modify the usual definition. By a differential equation/system of maximal order k we mean a sequence $\mathcal{E} = \{\mathcal{E}_i\}_{-1\leq i\leq k}$ of submanifolds $\mathcal{E}_i \subset J^i(\pi)$ with $\mathcal{E}_{-1} = M$, $\mathcal{E}_0 = J^0 M = E_\pi$ such that for all $0 < i \leq k$ the following conditions hold:

(a) $\pi^{\mathcal{E}}_{i,i-1} : \mathcal{E}_i \to \mathcal{E}_{i-1}$ are smooth fiber bundles.

(b) The first prolongations $\mathcal{E}_{i-1}^{(1)}$ are smooth subbundles of π_i and $\mathcal{E}_i \subseteq \mathcal{E}_{i-1}^{(1)}$.

Consider a point $x_k \in \mathcal{E}_k$ with $x_i = \pi_{k,i}(x_k)$ for $i < k$ and $x = x_{-1}$. It determines the collection of symbols $g_i(x_k) \subset S^i T_x^* M \otimes N_{x_0}$, where $N_{x_0} = T_{x_0}\left[\pi^{-1}(x)\right]$, by the formula

$$g_i(x_k) = T_{x_i}\left[\pi^{-1}_{i,i-1}(x_{i-1})\right] \cap T_{x_i}\mathcal{E}_i \subset S^i T_x^* M \otimes N_{x_0} \quad \text{for } i \leq k.$$

For $i > k$ the symbolic spaces g_i are defined as symbols of the prolongations $\mathcal{E}_i = \mathcal{E}_k^{(i-k)}$, and they still depend on the point $x_k \in \mathcal{E}_k$.

In this situation $g^*(x_k) = \oplus g_i^*$ is a graded module over the algebra $\mathcal{R} = ST_x M$ of homogeneous polynomials on the cotangent space $T_x^* M$. It is called the *symbolic module* of $\mathcal{E}$ at the point x_k and for systems of linear PDEs $\mathcal{E} = \mathrm{Ker}(\Delta)$ this coincides with the previously defined module $\mathcal{M}_\Delta$.

The *characteristic ideal* is defined by $I_{x_k}(\mathcal{E}) = \mathrm{ann}(g^*) \subset \mathcal{R}$ (in the symbolic context denoted by $I(g)$). The *characteristic variety* is the (projectivized/complexified) set of non-zero covectors $v \in T^*$ such that for every i there exists a vector $w \in N \setminus \{0\}$ with $v^i \otimes w \in g_i$. If the system is of maximal order k, it is sufficient for this definition to take $i = k$ only. We denote it by $\mathrm{Char}^{\mathbb{C}}_{x_k}(\mathcal{E}) \subset \mathbb{P}^{\mathbb{C}} T_x^* M$ (variants: $\mathrm{Char} \subset \mathbb{P} T_x^* M$, $\mathrm{Char}^{\mathbb{C}}_{\mathrm{aff}} \subset {}^{\mathbb{C}}T_x^* M$ etc).

Denote by $\mathrm{diff}(\pi, \nu)$ the space of all non-linear differential operators (linear included) between sections of bundles π and ν. Let $F \in \mathrm{diff}(\pi, \nu)$ determine the system $\mathcal{E}$. Then its symbol at $x_k \in \mathcal{E}_k$ resolves the symbolic module $g^*(x_k)$:

$$\cdots \to ST_x M \otimes \nu_x^* \xrightarrow{\sigma_F(x_k)} ST_x M \otimes \pi_x^* \to g^*(x_k) \to 0.$$

Here we use the weight convention in order to make the symbol map $\sigma_F(x_k)$ into $\mathcal{R}$-homomorphism. Its value at covector $p \in \mathrm{Char}^{\mathbb{C}}_{x_k}(\mathcal{E})$ is the fiber of the characteristic sheaf: $\mathcal{K}_p = \mathrm{Coker}\left[\sigma_F(x_k)(p)\right]$.

Precise form of the above free resolution in various cases allows to investigate the system $\mathcal{E}$ in details. In particular, results of §3.6 are based on the Buchsbaum-Rim resolution [6].

Working with symbolic modules we inherit various concepts from commutative algebra (consult e.g. [16]). Some of them are of primary importance for PDEs. For instance

$$\dim_{\mathcal{R}} g^* = \dim_{\mathbb{C}} \mathrm{Char}^{\mathbb{C}}_{\mathrm{aff}}(\mathcal{E}) = \dim_{\mathbb{C}} \mathrm{Char}^{\mathbb{C}}(\mathcal{E}) + 1$$

is the Chevalley dimension of $g^* = g^*(x_k)$.

System $\mathcal{E}$ is called a *Cohen-Macaulay system* if the corresponding symbolic module g^* is Cohen-Macaulay, i.e. $\operatorname{depth} g^* = \dim g^*$ (see [5]). Other notions like grade and height turns to be important in applications to differential equations as well ([46]).

Castlnuovo-Mumford regularity of g^* is closely related to the notion of involutivity (we'll give definition via Spencer δ-cohomology in §3.2). It is instructive to notice that, even though quasi-regular sequences are basic for both classes, involutive systems exhibit quite unlike properties compared to Cohen-Macaulay systems (some apparent duality is shown in [51]).

2.4 Non-linear differential operators

Let $\mathfrak{F} = C^\infty(J^\infty\pi)$ be the filtered algebra of smooth functions depending on finite jets of π, i.e. $\mathfrak{F} = \cup_i \mathfrak{F}_i$ with $\mathfrak{F}_i = C^\infty(J^i\pi)$.

Denote $\mathfrak{F}_i^{\mathcal{E}} = C^\infty(\mathcal{E}_i)$. The projections $\pi_{i+1,i} : \mathcal{E}_{i+1} \to \mathcal{E}_i$ induce the maps $\pi_{i+1,i}^* : \mathfrak{F}_i^{\mathcal{E}} \to \mathfrak{F}_{i+1}^{\mathcal{E}}$, so that we can form the space $\mathfrak{F}^{\mathcal{E}} = \cup \mathfrak{F}_i^{\mathcal{E}}$, the points of which are infinite sequences $(f_i, f_{i+1}, \dots)$ with $f_i \in \mathfrak{F}_i^{\mathcal{E}}$ and $\pi_{i+1,i}^*(f_i) = f_{i+1}$. This $\mathfrak{F}^{\mathcal{E}}$ is a $C^\infty(M)$-algebra. If the system $\mathcal{E}$ is not formally integrable, the set of infinite sequences can be void, and the algebra $\mathfrak{F}^{\mathcal{E}}$ can be trivial. To detect formal integrability, we investigate the finite level jets algebras $\mathfrak{F}_i^{\mathcal{E}}$ via the following algebraic approach.

Let $\mathcal{E}$ be defined by a collection $F = (F_1, \dots, F_r) \in \operatorname{diff}(\pi, \nu)$ of non-linear scalar differential operators of orders $k_1, \dots, k_r$ (can be repeated). Post-composition of our differential operator $F : C^\infty(\pi) \to C^\infty(\nu)$ with other non-linear differential operators $\Box$ (composition from the left $\hat{\Box} \circ F$) gives the following exact sequence of $C^\infty(M)$-modules

$$\operatorname{diff}(\nu, \mathbf{1}) \xrightarrow{F} \operatorname{diff}(\pi, \mathbf{1}) \to \mathfrak{F}^{\mathcal{E}} \to 0. \tag{4}$$

Denote $\mathcal{J}_t(F) = \langle \hat{\Box}_i \circ F_i \,\big|\, \operatorname{ord} \Box_i + k_i \le t, 1 \le i \le r \rangle \subset \operatorname{diff}_t(\pi, \mathbf{1})$ the submodule generated by $F_1, \dots, F_r$ and their total derivatives up to order t. Then

$$\mathfrak{F}_i^{\mathcal{E}} = \operatorname{diff}_i(\pi, \mathbf{1}) / \mathcal{J}_i(F_1, \dots, F_r). \tag{5}$$

It is important that the terms of (4) are modules over the algebra of scalar $\mathcal{C}$-differential operators $\mathcal{C}\operatorname{Diff}(\mathbf{1}, \mathbf{1})$, which are total derivative operators and have the following form in local coordinates [40]: $\Delta = \sum f_\sigma \mathcal{D}_\sigma$, with $f_\sigma \in C^\infty(J^\infty(M))$. We can identify $\mathcal{C}\operatorname{Diff}(\mathbf{1}, \mathbf{1}) = \cup \mathfrak{F}_i^{\mathbf{1}} \otimes \operatorname{Diff}_j(\mathbf{1}, \mathbf{1})$ with the twisted tensor product of the algebras $\mathfrak{F}^{\mathbf{1}} = C^\infty(J^\infty(M))$ and $\operatorname{Diff}(\mathbf{1}, \mathbf{1})$ over the action

$$\hat{\Delta} : \mathfrak{F}_i^{\mathbf{1}} \to \mathfrak{F}_{i+j}^{\mathbf{1}} \quad \text{for} \quad \Delta \in \operatorname{Diff}_j(\mathbf{1}, \mathbf{1}).$$

This $\mathcal{C}\operatorname{Diff}(\mathbf{1}, \mathbf{1})$ is a non-commutative $C^\infty(M)$-algebra. We need a more general $\mathfrak{F}$-module of $\mathcal{C}$-differential operators $\mathcal{C}\operatorname{Diff}(\pi, \mathbf{1}) = \cup \mathcal{C}\operatorname{Diff}_i(\pi, \mathbf{1})$, where

$$\mathcal{C}\operatorname{Diff}_i(\pi, \mathbf{1}) = \mathfrak{F}_i \otimes_{C^\infty(M)} \operatorname{Diff}_i(\pi, \mathbf{1}).$$

Remark that $\mathcal{C}\operatorname{Diff}(\pi, \mathbf{1})$ is a filtered $\mathcal{C}\operatorname{Diff}(\mathbf{1}, \mathbf{1})$-module.

Define now the filtered $\mathfrak{F}^{\mathcal{E}}$-module $\mathcal{C}\operatorname{Diff}^{\mathcal{E}}(\pi, \mathbf{1})$ with $\mathcal{C}\operatorname{Diff}_i^{\mathcal{E}}(\pi, \mathbf{1}) = \mathfrak{F}_i^{\mathcal{E}} \otimes \operatorname{Diff}_i(\pi, \mathbf{1})$. Since the module $\operatorname{Diff}(\pi, \mathbf{1})$ is projective and we can identify $\operatorname{diff}(\pi, \mathbf{1})$ with $\mathfrak{F}$, we have from (5) the following exact sequence

$$0 \to \mathcal{J}_i(F) \otimes \operatorname{Diff}_i(\pi, \mathbf{1}) \to \mathcal{C}\operatorname{Diff}_i(\pi, \mathbf{1}) \to \mathcal{C}\operatorname{Diff}_i^{\mathcal{E}}(\pi, \mathbf{1}) \to 0. \tag{6}$$

Similar modules can be defined for the vector bundle ν and they determine the $\mathfrak{F}^{\mathcal{E}}$-module $\mathcal{E}^* = \cup\mathcal{E}_i^*$ by the following sequence:

$$\mathcal{C}\operatorname{Diff}_i^{\mathcal{E}}(\nu, \mathbf{1}) \xrightarrow{\ell_F} \mathcal{C}\operatorname{Diff}_{i+k}^{\mathcal{E}}(\pi, \mathbf{1}) \to \mathcal{E}_{i+k}^* \to 0, \tag{7}$$

where $\ell : \operatorname{diff}(\pi, \nu) \to \mathfrak{F} \otimes_{C^\infty(M)} \operatorname{Diff}(\pi, \nu)$ is the operator of universal linearization [40], $\ell_F = \ell(F)$ (described in the next section).

This sequence is not exact in the usual sense, but it becomes exact in the following one. The space to the left is an $\mathfrak{F}_i^{\mathcal{E}}$-module, the middle term is an $\mathfrak{F}_{i+k}^{\mathcal{E}}$-module. The image $\ell_F(\mathcal{C}\operatorname{Diff}_i^{\mathcal{E}}(\nu, \mathbf{1}))$ is an $\mathfrak{F}_i^{\mathcal{E}}$-module, but we generate by it an $\mathfrak{F}_{i+k}^{\mathcal{E}}$-submodule in the middle term. With this understanding of the image the term $\mathcal{E}_{i+k}^*$ of (7) is an $\mathfrak{F}_{i+k}^{\mathcal{E}}$-module and the sequence is exact. In other words

$$\mathcal{E}_s^* = \mathcal{C}\operatorname{Diff}_s^{\mathcal{E}}(\pi, \mathbf{1}) / (\mathfrak{F}_s^{\mathcal{E}} \cdot \operatorname{Im} \ell_F).$$

Sequences (7) are nested (i.e. their union is filtered) and so we have the sequence

$$\mathcal{E}_{s-1}^* \to \mathcal{E}_s^* \to \mathcal{F}g_s^* \to 0, \tag{8}$$

which becomes exact if we treat the image of the first arrow as the corresponding generated $\mathfrak{F}_s^{\mathcal{E}}$-module. Thus $\mathcal{F}g_s^*$ is an $\mathfrak{F}_s^{\mathcal{E}}$-module with support on $\mathcal{E}_s$ and its value at a point $x_s \in \mathcal{E}_s$ is dual to the s-symbol of the system $\mathcal{E}$:

$$(\mathcal{F}g_s^*)_{x_s} = g_s^*(x_s); \qquad g_s(x_s) = \operatorname{Ker}[T_{x_s}\pi_{s,s-1} : T_{x_s}\mathcal{E}_s \to T_{x_{s-1}}\mathcal{E}_{s-1}].$$

In general non-linear situation Definition 1 should be changed to

Definition 2 System $\mathcal{E}$ is formally integrable if the maps $\pi_{i+1,i} : \mathcal{E}_{i+1} \to \mathcal{E}_i$ are submersions.

Proposition 5 *A system $\mathcal{E}$ is formally integrable iff the modules $\mathcal{F}g_s^*$ are projective and the maps $\pi_{i+1,i}^* : \mathcal{E}_i^* \to \mathcal{E}_{i+1}^*$ are injective.*

Note that whenever prolongations $\mathcal{E}_{k+l}$ exist and k is the maximal order, the fibers of the projections $\pi_{t,s} : \mathcal{E}_t \to \mathcal{E}_s$ carry natural affine structures for $t > s \geq k$.

2.5 Linearizations and evolutionary differentiations

Consider a non-linear differential operator $F \in \operatorname{diff}_k(\pi, \nu)$ and two sections $s, h \in C^\infty(\pi)$ (we assume π to be a vector bundle, though it's not essential). Define

$$\ell_{F,s}(h) = \tfrac{d}{dt} F(s + th)|_{t=0}$$

This operator is linear in h and depends on its k-jets, so we have

$$\ell_{F,s} \in \operatorname{Hom}_{C^\infty(M)}(\mathcal{J}^k(\pi), C^\infty(\nu)) = \operatorname{Diff}_k(\pi, \nu).$$

Moreover value of this operator at $x \in M$ depends on k-jet of s, so that $\ell_{F,s} = j_k(s)^*(\ell_F)$. We will also denote $\ell_{F,x^k} = \ell_{F,s}$ for $x_k = [s]_x^k$. This dependence is however non-linear and we get $\ell_F \in \mathfrak{F} \otimes_{C^\infty(M)} \operatorname{Diff}(\pi, \nu)$.

In such a form this operator generalizes to the case of different orders.

Definition 3 Operator $\ell : \mathrm{diff}(\pi,\nu) \to \mathcal{C}\,\mathrm{Diff}(\pi,\nu) = \mathfrak{F} \otimes_{C^\infty(M)} \mathrm{Diff}(\pi,\nu)$ is called the operator of linearization (universal linearization in [40]).

It is instructive to note that whenever the evolutionary PDE (t being an extra variable)

$$\partial_t u = G(u), \quad G \in \mathrm{diff}(\pi,\pi),\ u \in C^\infty(\pi),$$

with initial condition $u(0) = s$ is solvable, then for each $x_k = [s]_x^k$ we get: $\ell_{F,x_k}(G) = \frac{d}{dt} F\big(u(t)\big)|_{t=0}$ for (any if non-unique) solution $u(t)$.

In canonical coordinates (trivializing ν) linearization of $F = (F_1, \dots, F_r)$ is $\ell_F = \ell(F) = (\ell(F_1), \dots, \ell(F_r))$ with

$$\ell(F_i) = \sum (\partial_{p^j_\sigma} F_i) \cdot \mathcal{D}_\sigma^{[j]},$$

where $\mathcal{D}_\sigma^{[j]}$ denotes the operator $\mathcal{D}_\sigma$ applied to the j-th component of the section from $C^\infty(\pi)$.

Recall that $\mathfrak{F}$ is an algebra of functions on $J^\infty\pi$ with usual multiplication and $\mathrm{diff}(\pi,\nu)$ is a lef $\mathfrak{F}$-module: $\mathfrak{F}_i \cdot \mathrm{diff}_k(\pi,\nu) \subset \mathrm{diff}_{\max\{i,k\}}(\pi,\nu)$. With respect to this structure the operator of linearization satisfies the Leibniz rule:

$$\ell_{H\cdot F} = \ell_H \cdot F + H \cdot \ell_F, \quad H \in \mathrm{diff}(\pi,\mathbf{1}),\ F \in \mathrm{diff}(\pi,\nu). \tag{9}$$

Since ℓ_F is a derivation in F, we can introduce the operator $\mathcal{Э}_G$ by the formula

$$\mathcal{Э}_G^\nu(F) = \ell_F(G), \quad F \in \mathrm{diff}(\pi,\nu),\ G \in \mathrm{diff}(\pi,\pi).$$

Definition 4 The operator $\mathcal{Э}_G^\nu : \mathrm{diff}(\pi,\nu) \to \mathrm{diff}(\pi,\nu)$ is called the evolutionary differentiation corresponding to $G \in \mathrm{diff}(\pi,\pi)$.

In canonical coordinates with $G = (G_1, \dots, G_m)$ the i-th component of the evolutionary differentiation equals

$$\mathcal{Э}_{G;\,i}^\nu = \sum (\mathcal{D}_\sigma G_j) \cdot \partial_{p^j_\sigma}{}^{[i]},$$

where $\partial_{p^j_\sigma}{}^{[i]}$ denotes the operator $\partial_{p^j_\sigma}$ applied to the i-th component of the section from $C^\infty(\nu)$.

As a consequence of (9) evolutionary differentiations satisfy the Leibniz rule:

$$\mathcal{Э}_G^\nu(H \cdot F) = \mathcal{Э}_G^\nu(H) \cdot F + H \cdot \mathcal{Э}_G^\nu(F), \quad H \in \mathrm{diff}(\pi,\mathbf{1}),\ F \in \mathrm{diff}(\pi,\nu). \tag{10}$$

Moreover since linear differential operators commute with $\frac{d}{dt}$, we get:

$$\hat{K} \circ \mathcal{Э}_G^\nu = \mathcal{Э}_G^\xi \circ \hat{K}, \quad \forall K \in \mathrm{Diff}(\nu,\xi). \tag{11}$$

Proposition 6 *[40]. $\mathbb{R}$-linear maps satisfying (10) and (11) are evolutionary differentiations and only they.*

Corollary 7 *The space $\mathfrak{Ev}(\pi,\nu) = \{\mathcal{Э}_G^\nu : G \in \mathrm{diff}(\pi,\pi)\}$ for fixed vector bundles π,ν is a Lie algebra with respect to the commutator.*

Consider the surjective $\mathbb{R}$-linear map

$$\partial^{\pi} : \mathrm{diff}(\pi,\pi) \to \mathfrak{Ev}(\pi,\pi), \qquad G \mapsto \partial^{\pi}_{G}. \tag{12}$$

It is injective because $\partial^{\pi}_{G}(\mathrm{Id}) = G$, and so can be used to introduce Lie algebra structure on $\mathrm{diff}(\pi,\pi)$, with respect to which (12) is an anti-isomorphism of Lie algebras:

$$\partial^{\pi}_{\{F,G\}} = [\partial^{\pi}_{G}, \partial^{\pi}_{F}], \qquad F, G \in \mathrm{diff}(\pi,\pi).$$

Definition 5 The bracket $\{F,G\}$ is called the higher Jacobi bracket.

This bracket generalizes the Lagrange-Jacobi bracket from classical mechanics and contact geometry as well as Poisson bracket from symplectic geometry. It coincides with the commutator for linear differential operators.

We can calculate $\{F,G\} = \partial^{\pi}_{G}(F) - \partial^{\pi}_{F}(G) = \ell_F(G) - \ell_G(F)$. In canonical coordinates the bracket writes:

$$\{F,G\}_i = \sum \big(\mathcal{D}_\sigma(G_j)\cdot \partial_{p^j_\sigma} F_i - \mathcal{D}_\sigma(F_j)\cdot \partial_{p^j_\sigma} G_i\big).$$

2.6 Brackets and multi-brackets of differential operators

Let $\pi = m \cdot \mathbf{1}$ be the trivial bundle of rank m. Then linearization of the operator $F \in \mathrm{diff}(\pi, \mathbf{1})$ can be written in components: $\ell(F) = (\ell_1(F), \dots, \ell_m(F))$.

Multi-bracket of $(m+1)$ differential operators F_i on π is another differential operator on π, given by the formula [46]:

$$\{F_1, \dots, F_{m+1}\} = \frac{1}{m!} \sum_{\alpha\in S_m, \beta\in S_{m+1}} (-1)^{\alpha} (-1)^{\beta}\, \ell_{\alpha(1)}(F_{\beta(1)}) \circ \dots \circ \ell_{\alpha(m)}(F_{\beta(m)}) \left(F_{\beta(m+1)}\right).$$

When $m = 1$ we obtain the higher Jacobi bracket.

For linear vector differential operators $\nabla_i : m \cdot C^\infty_{\mathrm{loc}}(M) \to C^\infty_{\mathrm{loc}}(M)$, represented as rows $(\nabla^1_i, \dots, \nabla^m_i)$ of scalar linear differential operators, the multi-bracket has the form:

$$\{\nabla_1, \dots, \nabla_{m+1}\} = \sum_{k=1}^{m+1} (-1)^{k-1}\, \mathrm{Ndet}[\nabla^j_i]^{1\le j\le m}_{i\ne k} \circ \nabla_k,$$

where Ndet is a version of non-commutative determinant [46].

If we interchange Ndet and ∇_k in the above formula, we obtain the opposite multi-bracket $\{F_1, \dots, F_{m+1}\}^\dagger$ (taking another representative for Ndet).

Theorem 8 *[52]. Let $F_i \in \mathrm{diff}(\pi, \mathbf{1})$ be vector differential operators, $1 \le i \le m+2$, and let $F_{k,i}$ denote component i of F_k and $\{\cdots\}_i$ the i^{th} component of the multi-bracket. Then the multi-bracket and the opposite multi-bracket are related by the following identities (check means absence of the argument) for any $1 \le i \le m$:*

$$\sum_{k=1}^{m+2} (-1)^k \Big[\ell_{\{F_1,\dots,\check{F}_k,\dots,F_{m+2}\}^\dagger_i} F_k - \ell_{F_{k,i}} \{F_1, \dots, \check{F}_k, \dots, F_{m+2}\}\Big] = 0.$$

For $m = 1$ this formula becomes the standard Jacobi identity. In this case $F_i \in \mathrm{diff}(\mathbf{1},\mathbf{1})$ are scalar differential operators, multi-bracket becomes the higher Jacobi bracket $\{F,G\}$ and we get:

$$\sum_{\text{cyclic}} \left(\ell_F\{G,H\} - \ell_{\{G,H\}}F\right) = \sum_{\text{cyclic}} \{F,\{G,H\}\} = 0.$$

Thus the multi-bracket identities could be considered as generalized Jacobi identities (but neither in the sense of Nambu, generalized Poisson, nor as SH-algebras [76, 70]). We called them non-commutative Plücker identities in [52], because their symbolic analogs are precisely the standard Plücker formulas. Symbolic counterpart of the above identities can be interpreted as multi-version of the integrability of characteristics ([29, 52]).

Finally we give a coordinate representation of the introduced multi-bracket. As in the classical contact geometry there is a variety of brackets (see more in [55]). The following is the multi-bracket analog of the Mayer bracket:

$$[F_1,\dots,F_{m+1}] = \frac{1}{m!} \sum_{\substack{\sigma\in S_{m+1}\\ \nu\in S_m}} \frac{\mathrm{sgn}(\sigma)}{\mathrm{sgn}(\nu)} \sum_{\substack{1\le i\le m\\ |\tau_i|=k_{\sigma(i)}}} \prod_{j=1}^{m} \frac{\partial F_{\sigma(j)}}{\partial p^{\nu(j)}_{\tau_j}}\, \mathcal{D}_{\tau_1+\dots+\tau_m} F_{\sigma(m+1)},$$

where $F_i \in \mathrm{diff}_{k_i}(m\cdot\mathbf{1},\mathbf{1})$. For $m = 1$ this gives Mayer brackets instead of Jacobi brackets [43]. We have ([52]):

Proposition 9 *Restrictions of the two multi-brackets to the system $\mathcal{E} = \{F_1 = \dots = F_{m+1} = 0\}$ coincide:*

$$\{F_1,\dots,F_{m+1}\} \equiv [F_1,\dots,F_{m+1}] \bmod \mathcal{J}_{k_1+\dots+k_{m+1}-1}(F_1,\dots,F_{m+1}).$$

3 Formal theory of PDEs

3.1 Symbolic systems

Consider vector spaces T of dimension n and N of dimension m (usually over the field $\mathbb{R}$, but also possible over $\mathbb{C}$). The symmetric power $ST^* = \oplus_{i\ge 0} S^iT^*$ can be identified with the space of polynomials on T.

Spencer δ-complex is the graded de Rham complex of N-valued differential forms on T with polynomial coefficients:

$$0 \to S^kT^*\otimes N \xrightarrow{\delta} S^{k-1}T^*\otimes N\otimes T^* \xrightarrow{\delta} \cdots \xrightarrow{\delta} S^{k-n}T^*\otimes N\otimes \Lambda^nT^* \to 0,$$

where $S^iT^* = 0$ for $i < 0$. By Poincaré lemma δ-complex is exact.

For a linear subspace $h \subset S^kT^*\otimes N$ its *first prolongation* is

$$h^{(1)} = \{p \in S^{k+1}T^*\otimes N \mid \delta p \in h\otimes T^*\}$$

Higher prolongations are defined inductively and satisfy $(h^{(k)})^{(l)} = h^{(k+l)}$.

Definition 6 *Symbolic system* is a sequence of subspaces $g_k \subset S^kT^*\otimes N$ such that $g_{k+1} \subset g_k^{(1)}$, $k \ge 0$.

If $\mathcal{E}$ is a system of PDEs of maximal order k and $x_k \in \mathcal{E}_k$, then the symbols of $\mathcal{E}$, namely $\{g_i(x_k)\}$ form a symbolic system.

We usually assume $g_0 = N$ (if $g_0 \subsetneqq N$ one can shrink N). With every such a system we associate its Spencer δ-complex of order k:

$$0 \to g_k \xrightarrow{\delta} g_{k-1} \otimes T^* \xrightarrow{\delta} g_{k-2} \otimes \Lambda^2 T^* \to \cdots \xrightarrow{\delta} g_{k-n} \otimes \Lambda^n T^* \to 0.$$

The cohomology group at the term $g_i \otimes \Lambda^j T^*$ is denoted by $H^{i,j}(g)$ and is called Spencer δ-cohomology.

When g is the symbolic system corresponding to a system of PDEs we denote the cohomology by $H^{i,j}(\mathcal{E}; x_k)$ and often omit reference to the point.

Another way to deal with the system $g \overset{i}{\hookrightarrow} ST^* \otimes N$ is to consider its dual $g^* = \oplus g_k^*$, which is an epimorphic image of $ST \otimes N^*$ via the map i^*. The last space is naturally an ST-module and we can try to carry the module structure to g^* by the formula $w \cdot i^*(v) = i^*(w \cdot v)$, $w \in ST$, $v \in ST \otimes N^*$. Correctness of this operation has the following obvious meaning:

Proposition 10 *System $g \subset ST^* \otimes N$ is symbolic iff g^* is an ST-module.*

Orders of the system is the following set:

$$\mathrm{ord}(g) = \{k \in \mathbb{Z}_+ \mid g_k \neq g_{k-1}^{(1)}\}.$$

Multiplicity of an order k is $m(k) = \dim g_{k-1}^{(1)}/g_k$ and this equals to the dimension of the Spencer δ-cohomology group $H^{k-1,1}(g)$.

Hilbert basis theorem implies finiteness of the set of orders.

Definition 7 Call *formal codimension* of a symbolic system g the number of elements in $\mathrm{ord}(g)$ counted with multiplicities. In other words

$$\mathrm{codim}(g) = \sum_{k=1}^{\infty} \dim H^{k-1,1}(g).$$

3.2 Spencer δ-cohomology

Let us show how to calculate the Spencer δ-cohomology in some important cases. Denote $m = \dim N$, $r = \mathrm{codim}(g)$ and $U = \mathbb{R}^r$. Then minimal resolution of the symbolic module starts as follows:

$$\cdots \to ST \otimes U^* \longrightarrow ST \otimes N^* \to g^* \to 0.$$

Definition 8 Call a symbolic system g *generalized complete intersection* if the symbolic module satisfies: $\mathrm{depth}\,\mathrm{ann}(g^*) \geq r - m + 1$.

This condition will be interpreted for systems of PDEs in §3.6. It is a condition of general position for module g^* in the range $m < r < m + n$.

Any generalized complete intersection g is a Cohen-Macaulay system. By standard theorems from commutative algebra we have in fact equality for depth in the definition above.

Theorem 11 *[52]. If g is a generalized complete intersection, then the only non-zero Spencer δ-cohomology are given by the formula:*

$$H^{*,j}(g) = \begin{cases} N & \text{for } j = 0, \\ U & \text{for } j = 1, \\ S^{j-2}N^* \otimes \Lambda^{m+j-1}U & \text{for } 2 \leq j \leq r+1-m \leq n. \end{cases}$$

In this formula we suppressed bi-grading. If g corresponds to a system $\mathcal{E}$ of different orders PDEs, then $H^{*,j}$ is a sum of different cohomology spaces. They can be specified as follows (if there're several equal $H^{i,j}$ in the sum below, we count only one and the rest contributes to the growth of dimension):

$$H^{*,0}(g) = H^{0,0}(g), \qquad H^{*,1}(g) = \bigoplus_{i \in \mathrm{ord}(g)} H^{i-1,1}(g),$$
$$H^{*,2}(g) = \bigoplus_{i_1,\dots,i_{m+1} \in \mathrm{ord}(g)} H^{i_1+\dots+i_{m+1}-2,2}(g),$$
$$H^{*,3}(g) = \bigoplus_{i_1,\dots,i_{m+2} \in \mathrm{ord}(g)} H^{i_1+\dots+i_{m+2}-3,3}(g) \quad \text{etc}\dots$$

One of the most important techniques in calculating Spencer δ-cohomology of a symbolic system g comes from commutative algebra, because they $\mathbb{R}$-dualize to Koszul homology of the symbolic module g^* ([88]). In particular, homology calculus can be equivalently represented by calculating free resolvents of g^*, see [26].

However Spencer δ-cohomology are related to certain constructions specific to PDEs, which we are going to describe.

Having a symbolic system $g = \{g_l \subset S^l T^* \otimes N\}$ and a subspace $V^* \subset T^*$ we define another system $\tilde{g} = \{g_l \cap S^l V^* \otimes N\} \subset SV^* \otimes N$. This is a symbolic system, called the *V^*-reduction*.

It is important that such $\tilde{g}$ are precisely the symbolic systems corresponding to symmetry reductions, with respect to Lie group actions [2].

Theorem 12 *[44]. Let g be a Cohen-Macaulay symbolic system and a subspace $V^* \subset T^*$ be transversal to the characteristic variety of g:*

$$\mathrm{codim}(\mathrm{Char}^{\mathbb{C}}(g) \cap \mathbb{P}^{\mathbb{C}} V^*) = \mathrm{codim}\,\mathrm{Char}^{\mathbb{C}}(g).$$

Then Spencer δ-cohomology of the system g and its V^*-reduction $\tilde{g}$ are isomorphic:

$$H^{i,j}(g) \simeq H^{i,j}(\tilde{g}).$$

Another important transformation is related to solving Cauchy problem for general PDEs. Let $W \subset T$. The following exact sequence allows to project along the subspace $V^* = \mathrm{ann}(W)$:

$$0 \to V^* \hookrightarrow T^* \to W^* \to 0$$

Applying this projection to the symbolic system g we get a new symbolic system $\bar{g}_k \subset S^k W^* \otimes N$, called W-restriction.

In order to describe the result we need to introduce some concepts.

The first is involutivity. With every symbolic system $g \subset ST^* \otimes N$ and any $k \geq 0$ we can relate the symbolic system $g^{|k\rangle}$, which is generated by all differential corollaries of the system deduced from the order k:

$$g_i^{|k\rangle} = \begin{cases} S^i T^* \otimes N, & \text{for } i < k; \\ g_k^{(i-k)}, & \text{for } i \geq k. \end{cases}$$

Note that g is a system of pure order k if and only if $g = g^{|k\rangle}$. In this case classical Cartan definition of involutivity can be equivalently expressed via vanishing of Spencer δ-cohomology (see Serre's letter in [28]):

$$H^{i,j}(g) = 0 \quad \forall i \neq k-1.$$

For a system of different orders we have:

Definition 9 A system g is involutive if all systems $g^{|k\rangle}$ are involutive.

The number of conditions in this definition is not infinite, since only $k \in \operatorname{ord}(g)$ are essential. This general involutivity can still be expressed via vanishing of δ-cohomology for systems $g^{|k\rangle}$, but not for the system g ([51]).

Let us denote

$$\Upsilon^{i,j} = \bigoplus_{r>0} S^r V^* \otimes \delta(S^{i+1-r} W^* \otimes \Lambda^{j-1} W^*) \otimes N, \quad \Theta^{i,j} = \bigoplus_{q>0} \Upsilon^{i,q} \otimes \Lambda^{j-q} V^*,$$

where δ is the Spencer operator. Let also $\Pi^{i,j} = \delta(S^{i+1} V^* \otimes N \otimes \Lambda^{j-1} V^*)$.

Call a subspace $V^* \subset T^*$ strongly non-characteristic for a symbolic system g if $g_k \cap V^* \cdot S^{k-1} T^* \otimes N = 0$ for $k = r_{\min}(g)$ the minimal order of the system.

Theorem 13 *[51]. Let V^* be a strongly non-characteristic subspace for a symbolic system g. If g is involutive, then its W-restriction $\bar{g}$ is also involutive.*

Moreover the Spencer δ-cohomology of g and $\bar{g}$ are related by the formula:

$$H^{i,j}(g) \simeq \bigoplus_{q>0} H^{i,q}(\bar{g}) \otimes \Lambda^{j-q} V^* \oplus \delta^{i+1}_{r_{min}(g)} \cdot [\Theta^{i,j} \oplus \Pi^{i,j}] \oplus \delta^i_0 \delta^j_0 \cdot H^{0,0}(\bar{g}),$$

where δ^t_s is the Kronecker symbol.

If $\bar{g}$ is an involutive system of pure order $k = r_{min}(\bar{g}) = r_{max}(\bar{g})$, then g is also an involutive system of pure order k and the above formula holds.

The first two parts of this theorem generalize previous results of pure first order by Quillen and Guillemin, see [31].

3.3 Geometric structures

These are given by specification of a Lie group G in light of Klein's Erlangen program [39], though prolongations usually make this into infinite-dimensional Lie pseudo-group, see [28, 86, 94] and also §4.5. Not going much into details, we consider calculation of Spencer δ-cohomology and restrict, for simplicity, to the first order structures.

They correspond to G-structures, discussed in [89]. More general cases are studied in [30, 64]. Thus $g = g^{|1\rangle}$ is generated in order 1 with subspace $g_1 = \mathfrak{g} \subset \mathrm{gl}(n)$ being a matrix Lie algebra, corresponding to G.

Respective system of PDEs describes equivalence of a geometric structure, governed by a Lie group G, to the standard flat model. PDEs describing automorphism groups can be investigated similarly.

As we shall see in the next section, the group $H^{*,2}(g)$ plays an important role in investigation of formal integrability. For geometric structures this is the space of curvatures/torsions. We shall illustrate this with three examples:

(a) Almost complex geometry: $\mathfrak{g} = \mathrm{gl}(\frac{n}{2}, \mathbb{C})$. It is given by a tensor $J \in C^\infty(T^*M \otimes TM)$, $J^2 = -\mathbf{1}$;

(b) Riemannian geometry: $\mathfrak{g} = \mathrm{so}(n)$. It is given by a tensor $q \in C^\infty(S^2T^*M)$, $q > 0$;

(c) Almost symplectic geometry: $\mathfrak{g} = \mathrm{sp}(\frac{n}{2})$. It is given by a tensor $\omega \in C^\infty(\Lambda^2 T^*M)$, $\omega^n \neq 0$.

In all three cases $T = T_xM = N$ and there is a linear structure J, q or ω respectively on T.

In the first case (T, J) is a complex space and we can identify $g_1 = T^* \otimes_{\mathbb{C}} T$. The prolongations are $g_i = S^i_{\mathbb{C}} T^* \otimes_{\mathbb{C}} T$ (all tensor products over $\mathbb{C}$). The only non-zero Spencer δ-cohomology are:

$$H^{0,k}(g) = \Lambda^k_{\bar{\mathbb{C}}} T^* \otimes_{\bar{\mathbb{C}}} T,$$

which is the space of all skew-symmetric k-linear $\mathbb{C}$-antilinear T-valued forms on T. The system is involutive.

In the second system identification $T \overset{q}{\simeq} T^*$ yields $g_1 = \Lambda^2 T^*$. Since $g_2 = T^* \otimes \Lambda^2 T^* \cap S^2 T^* \otimes T^* = 0$, prolongations vanish $g_{1+i} = 0$ and the system is of finite type. The only non-zero Spencer δ-cohomology are:

$$H^{0,k}(g) = \Lambda^k T^* \otimes T, \quad H^{1,k} = \mathrm{Ker}(\delta : \Lambda^2 T^* \otimes \Lambda^k T^* \to T^* \otimes \Lambda^{k+1} T^*).$$

Thus g is not involutive. We can rewrite the cohomology in bi-grade $(1, 2)$ as $H^{1,2}(g) = \mathrm{Ker}(S^2 \Lambda^2 T^* \to \Lambda^4 T^*)$. Note that $H^{0,2}$ is the space of torsions and $H^{1,2}$ the space of curvature tensors.

In the last case we identify $T \overset{\omega}{\simeq} T^*$ and then get $g_1 = S^2 T^*$. Therefore prolongations $g_i = S^{i+1} T^*$ and the system is of infinite type. The only non-zero Spencer δ-cohomology are:

$$H^{0,k}(g) = \Lambda^{k+1} T^*.$$

The system is involutive.

3.4 Cartan connection and Weyl tensor

We define *regular system of PDEs* $\mathcal{E}$ of maximal order k as a submanifold $\mathcal{E}_k \subset J^k\pi$ co-filtered by $\mathcal{E}_l$, with $\mathcal{E}_i^{(1)} \supset \mathcal{E}_{i+1}$ and $\pi_i : \mathcal{E}_{i+1} \to \mathcal{E}_i$ being a bundle map, such that the symbolic system and the Spencer δ-cohomology form graded bundles over it. We define orders $\mathrm{ord}(\mathcal{E})$ of the system and its formal codimension $\mathrm{codim}(\mathcal{E})$ as these quantities for the symbolic system.

Cartan distribution on $\mathcal{E}_k$ is $\mathcal{C}_{\mathcal{E}_k} = \mathcal{C}_k \cap T\mathcal{E}_k$. *Cartan connection* on $\mathcal{E}_k$ is a horizontal subdistribution in it, i.e. a smooth family $H(x_k) \subset \mathcal{C}_{\mathcal{E}_k}(x_k)$, $x_k \in \mathcal{E}_k$, such that $d\pi_k : H(x_k) \to T_xM$ is an isomorphism. A Cartan connection yields the splitting $\mathcal{C}_{\mathcal{E}_k}(x_k) \simeq H(x_k) \oplus g_k(x_k)$ of the Cartan distribution into horizontal and vertical components.

Given a distribution Π on a manifold its *first derived differential system* $\partial\Pi$ is generated by the commutators of its sections. In the regular case it is a distribution and one gets the effective normal bundle $\nu = \partial\Pi/\Pi$. The *curvature* of Π is the vector-valued 2-form $\Xi_\Pi \in \Lambda^2\Pi^* \otimes \nu$ given by the formula:

$$\Xi_\Pi(\xi,\eta) = [\xi,\eta] \mod \Pi, \quad \xi,\eta \in C^\infty(\Pi)$$

(it is straightforward to check that Ξ_Π is a tensor).

The *metasymplectic structure* Ω_k on $J^k(\pi)$ is the curvature of the Cartan distribution [64, 48]. At a point x_k it is a 2-form on $\mathcal{C}_k(x_k)$ with values in the vector space $F_{k-1}(x_{k-1}) = T_{x_{k-1}}\left[\pi^{-1}_{k-1,k-2}(x_{k-2})\right] \simeq S^{k-1}T^*_xM \otimes N_x$.

To describe it fix a point $x_{k+1} \in J^{k+1}(\pi)$ over x_k and decompose $\mathcal{C}_k(x_k) = L(x_{k+1}) \oplus F_k(x_k)$. Then $\Omega_k(\xi,\eta) = 0$ if both ξ,η belong simultaneously either to $L(x_{k+1})$ or to $F_k(x_k)$. But if $\xi \in L(x_{k+1})$ corresponds to $X = d\pi_k(\xi) \in T_xM$ and $\eta \in F_k(x_k)$ corresponds to $\theta \in S^kT^*_xM \otimes N_x$, then the value of $\Omega_k(\xi,\eta)$ equals

$$\Omega_k(X,\theta) = \delta_X\theta \in S^{k-1}T^*_xM \otimes N_x,$$

where $\delta_X = i_X \circ \delta$ is the differentiation along X. The introduced structure does not depend on the point x_{k+1} determining the splitting because the subspace $L(x_{k+1})$ is Ω_k-isotropic.

Restriction of the metasymplectic structure $\Omega_k \in F_{k-1} \otimes \Lambda^2\mathcal{C}^*_k$ to the equation is the tensor $\Omega_{\mathcal{E}_k} \in g_{k-1} \otimes \Lambda^2\mathcal{C}^*_{\mathcal{E}_k}$. Given a Cartan connection H we define its curvature at x_k to be $\Omega_{\mathcal{E}_k}|_{H(x_k)} \in g_{k-1} \otimes \Lambda^2T^*_xM$. Considered as an element of the Spencer complex it is δ-closed and change of the Cartan connection effects in a shift by a δ-exact element.

The *Weyl tensor* $W_k(\mathcal{E};x_k)$ of the PDEs system $\mathcal{E}$ is the δ-cohomology class $[\Omega_{\mathcal{E}_k}|_{H(x_k)}] \in H^{k-1,2}(\mathcal{E};x_k)$ ([64]). For G-structures it coincides with the classical structural function [89]. For more general geometric structures it equals torsion/curvature tensor [30].

Prolongation $\mathcal{E}_{k+1} = \mathcal{E}^{(1)}_k$ is called regular if $\pi_{k+1,k} : \mathcal{E}_{k+1} \to \mathcal{E}_k$ is a bundle map. For regular systems a necessary and sufficient condition for regularity of the first prolongation is vanishing of the Weyl tensor: $W_k(\mathcal{E}) = 0$.

This gives the following criterion of formal integrability:

Theorem 14 *Let $\mathcal{E} = \{\mathcal{E}_l\}_{l=0}^k$ be a regular system of maximal order k. Then the system is formally integrable iff $W_i(\mathcal{E}) = 0$ for all $i \geq k$.*

Note that the number of conditions is indeed finite due to Poincaré δ-lemma: starting from some number i_0 all groups $H^{i,2}(\mathcal{E}) = 0$ for $i > i_0$ (see a bound for i_0 in [90]). This in fact was an original sufficient (cf. to necessary and sufficient in the above statement) criterion of formal integrability in [23, 88]: If all second cohomology groups $H^{i,2}$ vanish, $i \geq k$, then the regular system is formally integrable.

Tensor $W_k(\mathcal{E})$ plays a central role in equivalence problems [48]. Calculating Weyl tensor is a complicated issue, see e.g. [43, 44, 45], where it was calculated for complete intersection systems of PDEs. Let us perform calculation for the examples from §3.3.

(1) In this case $W_1 = N_J$ is the Nijenhuis tensor of the almost complex structure J. Its vanishing gives integrability condition of almost complex structure. The space $H^{0,2}(g) = \Lambda^2_{\bar{\mathbb{C}}} T^* \otimes_{\bar{\mathbb{C}}} T$ is the space of Nijenhuis tensors.

(2) Here Cartan connection is a linear connection ∇, preserving q. Tensor $W_1 = T_\nabla \in H^{0,2}$ is the torsion and its vanishing leads us to Levi-Civita connection ∇_q. The next Weyl tensor is the Riemannian curvature $R_q \in H^{1,2}$ and its vanishing yields flatness of the metric q.

(3) Curvature is $W_1 = d\omega \in H^{0,2}$. Thus integrability $W_1 = 0$ gives us symplectic structure ω.

Remark 4 Looking at these examples we observe that searching for involutivity is sometimes superfluous: All three geometries are equally important and from the point of view of getting solutions one just studies formal integrability, which usually occurs at smaller number of prolongations than involutivity.

Let us finish by mentioning without calculations that $W_k(\mathcal{E})$ equals the conformal Weyl tensor for the conformal Lie algebra $\mathfrak{g} = \mathrm{co}(n)$ and the Weyl projective tensor for the projective Lie algebra $\mathfrak{g} = \mathrm{sl}(n+1)$. Whence the name.

3.5 Compatibility and solvability

Investigation of overdetermined systems of PDEs begins with checking compatibility conditions. Riquier-Janet theory makes finding compatibility conditions algorithmic.

With Riquier approach [85] one expresses certain higher order derivatives via others (i.e. bring equations to the orthonomic form), differentiate PDEs and substitute the expressed quantities. If new equations arise, the system is called active, otherwise passive. In modern language we talk of formal integrability. Riquier's test on passivity allows to disclose the compatibility conditions.

Janet monomials [37] (and also Thomas's [97]) allow to check compatibility via cross differentiations of respective equations, which is determined by their differential monomials in certain ordering. To a large extent this can be seen as an origin of computer differential algebra (Gröbner bases etc).

Being algorithmic, these approaches are heavily calculational and so good luck in generators in $\mathcal{E}$ and coordinates in jet spaces plays a major role. On the contrary Cartan's theory [10] has a geometric base (Vessiot's dual approach [100] is of the same flavor) and so coupled with Spencer's homological technique [87] allows to calculate compatibility conditions in a visibly minimal number of steps.

Remark 5 Original Cartan's approach aims though to involutivity, not just to formal integrability, cf. Remark 4. In this respect Riquier-Janet theory is more economic.

Using the machinery of the previous section we can describe one step prolongation as follows. Assume $\mathcal{E}$ is a regular system of PDEs of maximal order k, which includes compatibility to order k. Then $W_k(\mathcal{E}) \in H^{k-1,2}(\mathcal{E})$ is precisely the obstruction to prolongation to $(k+1)$-st jets.

The number of compatibility conditions is $\dim H^{k-1,2}(\mathcal{E})$ (this quantity is constant along $\mathcal{E}$ due to regularity) and they are just components of the Weyl tensor $W_k(\mathcal{E})$. These latter are certain differential equations of $\mathrm{ord} \le k$.

If $W_k(\mathcal{E}) = 0$, the system can be prolonged to level $(k+1)$ and we get a system $\mathcal{E}_{k+1}$, with projections $\pi_{k+1,k} : \mathcal{E}_{k+1} \to \mathcal{E}_k$ being a vector bundle, so that we get new regular system and can continue prolongations. By Hilbert's theorem $H^{i,2}(\mathcal{E}) = 0$ starting from some number i_0. Then there's no more obstructions to compatibility and the system is formally integrable.

If the Weyl tensor is non-zero, we disclose new equations in the system $\mathcal{E}$, which are differential corollaries of $\operatorname{ord} \leq k$, and so we change the system by adding them. The new system is

$$\tilde{\mathcal{E}} = \mathcal{E} \cap \pi_{k+1,k}(\mathcal{E}_k^{(1)}) = \{x_k \in \mathcal{E}_k : W_k(\mathcal{E}; x_k) = 0\}.$$

We restart investigation of formal integrability with this new system of equations. This approach is called prolongation-projection method.

The following statement, known as Cartan-Kuranishi theorem, states that we do not continue forever:

Theorem 15 *After a finite number of prolongations-projections system $\mathcal{E}$ will be transformed into a formally integrable system $\bar{\mathcal{E}} \subset \mathcal{E}$.*

This statement was formulated by Cartan in [10] without precise conditions. It was proved by Kuranishi [56] under suitable regularity assumptions (see also [73]), but essentially the proof was published long before by the Russian school [84, 21].

With regularity assumptions we remove points, where the ranks of symbol bundles/δ-cohomology drop, together with their projections and prolongations. One hopes that most points will survive, so that the above theorem holds at a generic point.

While in general this is not known, it holds in some good situations. In algebraic case the result is due to Pommaret [82] and in analytic case due to Malgrange [72].

Note that we started with regular systems $\mathcal{E}$ of maximal order k, though one should start from $\mathcal{E}_1$. Arriving to $\mathcal{E}_l$ one can either add compatibility conditions or new equations of the system of order $l+1$. In the latter case the projection $\pi_{l+1,l} : \mathcal{E}_{l+1} \to \mathcal{E}_l$ is a fiber bundle (in regular case). The prolongation-projection method can be generalized to this situation and Theorem 15 holds in the same range of assumptions.

As a result of the method we get a minimal formally integrable sub-system $\bar{\mathcal{E}} \subset \mathcal{E}$. If it is non-empty the system $\mathcal{E}$ is called (formally) solvable. Indeed all (formal) solutions of $\mathcal{E}$ coincide with these of $\bar{\mathcal{E}}$. This alternative E.Cartan [10] characterized as follows: *"after a finite number of prolongations the system becomes involutive or contradictory"*.

3.6 Formal integrability via multi-brackets and Massey product

Though Weyl tensors $W_k(\mathcal{E})$ are precisely compatibility conditions, it is important to have a good calculational formula for the latter, at least for some classes of PDEs. The following is a wide class of systems, important in applications.

Let $\mathcal{E} \subset J^k(\pi)$ be a regular system of maximal order k, consisting of $r = \operatorname{codim}(\mathcal{E})$ differential equations on $m = \operatorname{rank}(\pi)$ unknown functions.

Definition 10 System $\mathcal{E}$ is of *generalized complete intersection type* if

(a) $m \leq r < n + m$;

(b) The characteristic variety has $\dim_{\mathbb{C}} \operatorname{Char}^{\mathbb{C}}_{x_k}(\mathcal{E}) = n+m-r-2$ at each point $x_k \in \mathcal{E}$ (we assume $\dim \emptyset = -1$);

(c) The characteristic sheaf $\mathcal{K}$ over $\mathrm{Char}^{\mathbb{C}}_{x_k}(\mathcal{E}) \subset \mathbb{P}^{\mathbb{C}}T^*$ has fibers of dimension 1 everywhere.

If $\mathcal{E}$ is a generalized complete intersection in this sense, then its symbolic system g is a generalized complete intersection in the sense of definition 8.

Note that vanishing of the multi-brackets due to the system is a necessary condition for formal integrability, because they belong to differential ideal of the system.

Theorem 16 *[46, 52]. Consider a system of PDEs*

$$\mathcal{E} = \Bigl\{ F_i\Bigl(x^1, \dots, x^n, u^1, \dots, u^m, \frac{\partial^{|\sigma|} u^j}{\partial x^\sigma}\Bigr) = 0 \,:\, 1 \le i \le r \Bigr\}, \quad \mathrm{ord}(F_i) = k_i.$$

If $\mathcal{E}$ is a system of generalized complete intersection type, then it is formally integrable if and only if the multi-brackets vanish due to the system:

$$\{F_{i_1}, \dots, F_{i_{m+1}}\} \bmod \mathcal{J}_{k_{i_1}+\dots+k_{i_{m+1}}-1}(F_1, \dots, F_r) = 0.$$

In particular, we get the following compatibility criterion for scalar PDEs:

Corollary 17 *Let $\mathcal{E} = \{F_1[u] = 0, \dots, F_r[u] = 0\}$ be a scalar system of complete intersection type, i.e. $r \le n$ and $\mathrm{codim}_{\mathbb{C}} \mathrm{Char}^{\mathbb{C}}(\mathcal{E}) = r$. Then formal integrability expresses via Mayer-Jacobi brackets as follows:*

$$\{F_i, F_j\} = 0 \bmod \mathcal{J}_{k_i+k_j-1}(F_1, \dots, F_r), \quad \forall\, 1 \le i < j \le r.$$

This criterion is effective in the study of not only compatibility, but also solvability of systems of PDEs. Examples of applications are [22, 52]. Moreover, since differential syzygy is provided explicitly, it is more effective than the method of differential Gröbner basis or its modifications [41].

Let now describe a sketch of the general idea how to investigate systems of PDEs $\mathcal{E} = \{F_1[u^1, \dots, u^m] = 0, \dots, F_r[u^1, \dots, u^m] = 0\}$ for compatibility.

Take a pair of equations F_i and F_j, $i < j$. Even though the system $\{F_i = 0, F_j = 0\}$ is underdetermined (for $m > 1$) it can possess compatibility conditions $\Theta_{ij} = 0$ of order t_{ij} (this actually means that after a change of coordinates this pair of PDEs will involve only one dependent function; but rigorously can be expressed only via non-vanishing second Spencer δ-cohomology). We denote $\Theta^{\mathcal{E}}_{ij} = \Theta_{ij} \bmod \mathcal{J}_{t_{ij}}(F_1, \dots, F_r)$. Thus we get compatibility conditions $\Theta^{\mathcal{E}}_{ij} = 0$ of orders $\tau_{ij} \le t_{ij}$.

Then we look to triples F_i, F_j, F_h with $i < j < h$, get in a similar way compatibility conditions $\Theta^{\mathcal{E}}_{ijh} = 0$ of orders τ_{ijh} and so forth. In general we get "generalized s-brackets" $\Theta^{\mathcal{E}}_{i_1 \dots i_s}$ of orders $\tau_{i_1 \dots i_s}$ for all $2 \le s \le r$ (see [43] §3.2 for an example of 3-bracket in the case $n = k = 2$, $m = 1$, $r = 3$).

The formula of the operator $\Theta^{\mathcal{E}}_{i_1 \dots i_s}$ and the number $\tau_{i_1 \dots i_s}$ ($i_1 < \dots < i_s$) strongly depends on the type of the system and varies with the type of characteristic variety/symbolic module. For each type of normal form or singularity one gets own formulas. In the range $m \le r < m + n$ the generic condition is that all generalized s-brackets for $s \le m$ are void and the first obstruction to formal integrability are $(m+1)$ multi-brackets. Theorem 16 states that this will be the only set of compatibility conditions.

Remark 6 Important case of an overdetermined system with $r = m$ constitute Einstein-Hilbert field equations [18]. They possess compatibility conditions, which hold identically, implying formal integrability of the system.

Note that the idea of calculating successively 2-product for a pair, then 3-product for a triple (in the case it vanishes for all sub-pairs) etc is very similar to Massey products in topology and algebra. Resembling situation is observed in deformation theory of module structures[4].

We note however that in the context of PDEs the situation is governed by order. We examine the set $\{\tau_{ij},\tau_{ijh},\dots,\tau_{1\dots r}\}$ (some numbers can be omitted if respective $\Theta^{\mathcal{E}}_{i_1\dots i_s}$ are void) and take the minimal order. These compatibility conditions are investigated first. Being non-zero, they are added to $\mathcal{E}$ and one considers a new system $\tilde{\mathcal{E}}$ (which can be simpler with lower order compatibility conditions, cf. §3.5). If these compatibility conditions are satisfied, we take the next ones and so on. The procedure is finite in the same sense as in Cartan-Kuranishi prolongation-projection theorem.

3.7 Integral Grassmanians revisited

A system of PDEs $\mathcal{E}\subset J^1(E,m)$ is said to be *determined* if $\operatorname{codim}\mathcal{E}=m$ and $\operatorname{codim}_{\mathbb{C}}\operatorname{Char}^{\mathbb{C}}(\mathcal{E})=1$. We usually represent such systems as the kernel of a (non-linear) operator $F:C^\infty(\pi)\to C^\infty(\nu)$ with $\operatorname{rank}\pi=\operatorname{rank}\nu=m$.

In this section we restrict to the case $k=1$ of first order equations. Let w_i^1 be Stiefel-Whitney classes of the tautological vector bundle over the Grassmanian $I(x_1)$, as before.

Theorem 18 *[65]. Let $\mathcal{E}\subset J^1(E,m)$ be a determined system such that the characteristic variety $\operatorname{Char}^{\mathbb{C}}_{x_1}(\mathcal{E})$ does not belong to a hyperplane for any $x_1\in\mathcal{E}$. Then the embedding $I\mathcal{E}(x_1)\hookrightarrow I(x_1)$ induces an isomorphism of cohomology with $\mathbb{Z}_2$-coefficients up to dimension n in all cases except the following:*

(a) $m=2, n\geq 3$. *Then $H^*(I\mathcal{E}(x_1),\mathbb{Z}_2)$ is isomorphic to the algebra $\mathbb{Z}_2[w_1^1,\dots,w_n^1,U_{n-1},\operatorname{Sq}U_{n-1}]$ up to dimension n, where U_{n-1} has dimension $n-1$ and Sq is the Steenrod square.*

(b) $m=3, n=2$. *Then $H^*(I\mathcal{E}(x_1),\mathbb{Z}_2)$ is isomorphic to the algebra $\mathbb{Z}_2[w_1^1,w_2^1,\rho_1,\dots,\rho_r]$ up to dimension 2, where dimensions of ρ_i equal 2 and r is a number of components of $\operatorname{Char}^{\mathbb{C}}(\mathcal{E},x_1)$ with the fibers of the kernel sheaf $\mathcal{K}$ of dimension 1.*

(c) $m=n=2$. *Then $I\mathcal{E}(x_1)$ is diffeomorphic to the torus $S^1\times S^1$ in hyperbolic case or to the complex projective line $\mathbb{CP}^1$ in elliptic case.*

This yields calculation of cohomology of integral Grassmanians for determined systems. Underdetermined systems can be treated similarly.

Finding cohomology of $I\mathcal{E}(x_k)$ in general overdetermined case seems to be a hopeless problem. However in many cases they stabilize after a sufficient number of prolongations. This constitutes a topological version of the Cartan-Kuranishi theorem:

Theorem 19 *[65]. Let $\mathcal{E}$ be a system of differential equations of pure first order $\mathcal{E}_1\subset J^1(E,m)$. Suppose that it is formally integrable and characteristically regular and such that the characteristic variety $\operatorname{Char}^{\mathbb{C}}_{x_1}(\mathcal{E})$ does not belong to a hyperplane for any $x_1\in\mathcal{E}_1$. Assume also that $\dim_{\mathbb{C}}\operatorname{Char}^{\mathbb{C}}(\mathcal{E})>0$. Then the embeddings $I\mathcal{E}_l(x_l)\hookrightarrow I(x_l)$ induce an isomorphism in cohomology with $\mathbb{Z}_2$-coefficients up to dimension n for sufficiently large values of l.*

[4]We thank A. Laudal for a fruitful discussion on this topic.

Let $I^k(E,m) = \bigcup I(x_k)$ be the total space of all integral Grassmanians. Then any integral n-dimensional manifold $L \subset J^k(E,m)$ defines a *tangential map* $t_L : L \to I^k(E,m)$, where $t_L : L \ni x_k \longmapsto T_{x_k}L \in I(x_k)$. Each cohomology class $\kappa \in H^l(I^k(E,m),\mathbb{Z}_2)$ gives rise to a *characteristic class* $\kappa(L) = t_L^*(\kappa)$ on integral manifolds L. By Theorem 2 $H^*(I^k(E,m),\mathbb{Z}_2)$, considered as an algebra over $H^*(J^k(E,m),\mathbb{Z}_2)$, is generated by the Stiefel-Whitney classes $w_1^k,\dots,w_n^k$ of the tautological bundle over $I^k(E,m)$ up to dimension n. Moreover since the bundle $J^l(E,m) \to J^{l-1}(E,m)$ is affine for $l > 1$ and for $l = 1$ is the standard Grassmanian bundle, we get that $H^*(J^k(E,m),\mathbb{Z}_2) = H^*(J^1(E,m),\mathbb{Z}_2)$, considered as an algebra over $H^*(E,\mathbb{Z}_2)$, is generated by the Stiefel-Whitney classes $w_1,\dots,w_n$ of the tautological bundle over $J^1(E,m)$, provided that $\pi_1(E) = 0$.

Let $\mathcal{E}_k \subset J^k(E,m)$ be a formally integrable system of PDEs of maximal order k and $I\mathcal{E}_{k+l} = \bigcup I\mathcal{E}_{k+l}(x_{k+l}) \subset I^{k+l}(E,m)$ be a total space of all integral Grassmanians associated with l-th prolongation $\mathcal{E}_{k+l} = \mathcal{E}_k^{(l)}$. If the differential equation satisfies the conditions of Theorem 19, then for a sufficiently large l the cohomology $H^*(I\mathcal{E}_{k+l},\mathbb{Z}_2)$, considered as an algebra over $H^*(\mathcal{E}_{k+l},\mathbb{Z}_2)$, is generated by the Stiefel-Whitney classes $w_1^k,\dots,w_n^k$ of the tautological vector bundle up to dimension n. On the other hand all bundles $\pi_{k+l,k+l-1} : \mathcal{E}_{k+l} \to \mathcal{E}_{k+l-1}$ are affine for $l > 0$ and hence $H^*(\mathcal{E}_{k+l},\mathbb{Z}_2) = H^*(\mathcal{E}_k,\mathbb{Z}_2)$.

By a *Cauchy data* we mean an $(n-1)$-dimensional integral manifold $\Gamma \subset \mathcal{E}_{k+l}$ together with a section $\gamma : \Gamma \to I\mathcal{E}_{k+l}$ such that $T_{x_{k+l}}\Gamma \subset \gamma(x_{k+l})$ for all $x_{k+l} \in \Gamma$. A solution of the Cauchy problem is an integral n-dimensional submanifold $L \subset \mathcal{E}_{k+l}$ with boundary such that $\Gamma = \partial L$. Each cohomology class $\theta \in H^{n-1}(I\mathcal{E}_{k+l},\mathbb{Z}_2)$ defines a characteristic number $\theta(\Gamma) = \langle \gamma^*\theta, \Gamma\rangle$.

Theorem 20 *If the Cauchy problem* (Γ,γ) *has a solution, then the characteristic numbers* $\theta(\Gamma)$ *vanish for all* $\theta \in H^{n-1}(I\mathcal{E}_{k+l},\mathbb{Z}_2)$.

4 Local and global aspects

4.1 Existence theorems

System of PDEs $\mathcal{E}$ is called locally/globally integrable, if for its infinite prolongation $\mathcal{E}_\infty$ and any admissible jet $x_\infty \in \mathcal{E}_\infty$ there exists a local/global smooth solution $s \in C^\infty(\pi)$ with $[s]_x^\infty = x_\infty$ (this clearly can be generalized to more general spaces $J^k(E;m)$ of jets).

If the system $\mathcal{E}$ is of finite type and formally integrable, then it is locally integrable. Indeed, the Cartan distribution $\mathcal{C}_{\mathcal{E}_k}$ for k so large that $g_k = 0$ has rank n and is integrable by the Frobenius theorem. Its local integral leaves are solutions of the system $\mathcal{E}$.

For infinite type systems, $\dim g_k \not\to 0$, formal integrability does not imply local integrability in general, some additional conditions should be assumed. One sufficient condition for existence of local solutions is analyticity as Cartan-Kähler theorem claims [38]:

Theorem 21 *Let a system* $\mathcal{E}$ *be analytic, regular and formally integrable. Then it is locally integrable, i.e. any admissible jet* $x_\infty \in \mathcal{E}_\infty$ *is the jet of a local analytic solution.*

This theorem is a generalization of Cauchy-Kovalevskaya theorem [80]. Other generalizations are known, see [19]. In particular we should mention Ovsyannikov's theorem, according to which for a system $\mathcal{E}$, written in the orthonomic form

$$\frac{\partial^{k_i} u^i}{\partial t^{k_i}} = F_i(t, x, u, \mathcal{D}_\sigma u), \quad i = 1, \dots, m$$

(F_i does not contain derivatives of order higher than k_i and it is free of $\partial_t^{k_i} u^i$) it is enough to require analyticity only in $x = (x^1, \dots, x^{n-1})$ and $u, \mathcal{D}_\sigma u$, and also continuity in t in order to get solution to the Cauchy initial value problem

$$\frac{\partial^l u^i}{\partial t^l}(0, x) = U_l^i(x), \quad l = 0, \dots, k_i - 1,\ i = 1, \dots, m. \tag{13}$$

Notice that the solution to a formally integrable system $\mathcal{E}$ in general form of Theorem 21 is given by a sequence of solutions of Cauchy problems, see [38, 7]. The initial value problem is specified similar to (13), with prescribed collection of s_p functions of p arguments, ..., s_0 constants, where s_i are Cartan characters [10], see also §4.3. Thus one can, in principle, slightly relax analyticity conditions of Theorem 21 via Ovsyannikov's approach.

Also note that according to Holmgren's theorem ([80]) a local solution to the Cauchy problem for a formally integrable analytic system $\mathcal{E}$ is unique even if the (non-characteristic) initial data is only smooth. In general smooth case formal integrability implies local integrability only in certain cases.

One of such cases is when the system $\mathcal{E}$ is purely hyperbolic, i.e. when complex characteristics complexify the real ones: $\mathrm{Char}^{\mathbb{C}}(\mathcal{E}) = (\mathrm{Char}(\mathcal{E}))^{\mathbb{C}}$. In other words any real plane of complimentary dimension in $\mathbb{P}^{\mathbb{C}}T^*$ intersects $\mathrm{Char}(\mathcal{E})$ in $\deg \mathrm{Char}^{\mathbb{C}}(\mathcal{E})$ real different points.

More general case is represented by involutive hyperbolic systems, which are given by the condition that on each step of Cartan-Kähler method the arising determined systems are hyperbolic (these systems are non-unique, see [101] for the precise definition) on the symbolic level.

Theorem 22 *[101]. If an involutive hyperbolic system $\mathcal{E}$ is formally integrable and the Cauchy data is non-characteristic, there exists a local solution. In particular, if not all covectors are characteristic for $\mathcal{E}$, then there is a local solution through almost any admissible jet $x_\infty \in \mathcal{E}_\infty$.*

In particular one can solve locally the Cauchy problem for the second order hyperbolic quasi-linear systems of the type arising in general relativity [12].

Another important example constitute purely elliptic systems, i.e. such $\mathcal{E}$ that the real characteristic variety is empty: $\mathrm{Char}(\mathcal{E}) = \emptyset$.

A system $\mathcal{E} = \{F_i(x, u, \mathcal{D}_\sigma u) = f_i(x)\}$ (independent and dependent variables x, u are multi-dimensional) is said to be of analytic type if locally near any point $x_k \in \mathcal{E}_k$ the (non-linear in general) differential operator $F = (F_i)_{i=1}^r$ is analytic in a certain chart (but the charts may overlap smoothly and $f = (f_i)_{i=1}^r$ is just smooth).

Theorem 23 *[87, 71]. Elliptic formally integrable system $\mathcal{E}$ of analytic type is locally integrable.*

Spencer conjectured [88] that any formally integrable elliptic system is locally integrable[5], but this was not proved in the full generality.

[5] In fact, we cannot prescribe the value of jet of the solution, and so it's better to talk here of local solvability from the next section.

A certain progress was due to the works of MacKichan [74, 75] and Sweeney [91]. They studied solvability of the Neumann problem and related this to the δ-estimate on a linear operator Δ: For any large k in

$$0 \to g_{k+1} \xrightarrow{\delta} g_k \otimes T^* \xrightarrow{\delta} g_{k-1} \otimes \Lambda^2 T^* \text{ and any } \xi \in g_k \otimes T^* \cap \operatorname{Ker} \delta^*$$

(operator δ^* is conjugated to δ with respect to some Hermitian metrics on T^*, π, ν), we have $\|\delta\xi\| \geq \frac{k}{\sqrt{2}}\|\xi\|$. Their results imply (see [88], also [95]):

Theorem 24 *Let $\mathcal{E} = \operatorname{Ker} \Delta$ be a formally integrable system, with not all covectors being characteristic. Suppose that the operator Δ satisfies the δ-estimate. Then the system $\mathcal{E}$ is locally integrable.*

4.2 Local solvability

$\mathcal{E}$ is called solvable if we can guarantee existence of a local/global solution. Obviously one should first carry prolongation-projection method to get a maximal formally integrable system $\bar{\mathcal{E}} \subset \mathcal{E}$ (this is usually called bringing $\mathcal{E}$ to an involutive form, compare though with Remark 4), so we can assume that already $\mathcal{E}$ fulfills compatibility conditions.

In light of Cartan-Kähler theorem one would like to specify an admissible jet of solution. It is possible for hyperbolic systems and their generalizations (see Theorem 22), but not for all systems. However even the problem of finding some solution can be non-solvable.

Consider at first smooth linear differential operators. Let us restrict to $\mathbb{C}$-scalar PDEs, i.e. differential operators $\Delta = \Delta_1 + i\Delta_2 : C^\infty(M;\mathbb{C}) \to C^\infty(M;\mathbb{C})$ of order k (Δ_i are real, so one can think of determined real system $\mathcal{E}$ with $\operatorname{rank} \pi = 2$, but it is of special type: $\operatorname{codim} \operatorname{Char}(\mathcal{E}) = 2$).

The first example of operator of this type such that the corresponding PDE $\Delta(u) = f$ not locally solvable for some smooth $f \in C^\infty(M;\mathbb{C})$, was constructed by H. Levi [60][6].

This example was later generalized by Hörmander, Grushin and others. In fact, Hörmander found a necessary condition of solvability for principal type operators. Using the bracket approach of §3.6 we can formulate it as follows. Let $\mathcal{E}$ be the above PDE written as the real system $\{\Delta_1(u) = f_1, \Delta_2(u) = f_2\}$ ($\Delta_1 = \operatorname{Re} \Delta$, $\Delta_2 = \operatorname{Im} \Delta$ and similar for f) and let σ_s denote the symbol of order s. Then local solvability of a system $\mathcal{E}$ implies

$$\sigma_{2k-1}\Big(\{\Delta_1(u), \Delta_2(u)\} \bmod \mathcal{E}\Big) = 0. \qquad (14)$$

Hörmander formulated his condition differently [35]. Namely denote $H = \sigma_k(\Delta_1)$, $F = \sigma_k(\Delta_2)$. Then the Poisson bracket $\{H, F\}$ vanishes on $\Sigma = \operatorname{Char}_{\text{aff}}(\mathcal{E}) = \{H = 0, F = 0\} \subset T^*M$. In other words, whenever Σ is a submanifold, it is involutive[7].

The condition that the operator Δ has a principal type means that the Hamiltonian vector field X_H on $\{H = 0\}$ is not tangent to the fibers of the projection $T^*M \to M$ (it is

[6] His Δ was a very nice analytic operator of order 1, namely Cauchy-Riemann operator on the boundary of the pseudo-convex set $\{|z_1|^2 + 2\operatorname{Im} z_2 < 0\} \subset \mathbb{C}^2$.

[7] Note that for real problems, when $F = 0$ and $\operatorname{Im}(u) = 0$ this condition is void. Indeed linear determined PDEs of principal type with real (nonconstant) coefficients in the principal part are locally solvable [34] (this also follows from Theorem 25).

possible to multiply Δ by a function $a \in C^\infty(M;\mathbb{C})$, so that a particular choice of H and F is not essential, but the condition involves linear combinations of dH and dF).

Trajectory of the vector field X_H are called bi-characteristics and considered on the invariant manifold $\{H=0\}$ they are called null bi-characteristics.

Note that even when Σ is not a submanifold, condition (14) can be reformulated as follows

Along null bi-characteristics X_H function F vanishes to at least 2^{nd} *order.*

This condition was refined by Nirenberg-Treves to the following condition:

Along null bi-characteristics X_H function F *does not change its sign.* $(\mathcal{P})$

If Σ is a submanifold in T^*M and dH, dF are independent on its normal bundle, this condition is equivalent to (14). In general it strictly includes (implies) Hörmander condition. Indeed if order of zero for F along null bi-characteristics is finite, it should be even.

It turns out that this new condition is not only necessary, but also sufficient for solvability ([77] with the condition of order $k=1$ or base dimension $n=2$ or that the principal part is analytic; [4] in general):

Theorem 25 *If Δ is of principal type and satisfies condition $(\mathcal{P})$, then for any smooth f linear PDE $\Delta(u)=f$ is locally solvable.*

This theorem was generalized to pseudo-differential operators (see [59] for important partial cases and review; [15] in general), with condition $(\mathcal{P})$ being changed to a similar condition (Ψ). In such a form it is sometimes possible to give a sufficient condition for global solvability (see loc.cit).

In the second paper [77] Nirenberg and Treves gave a vector version (determined system of special type with $\operatorname{rank}_{\mathbb{R}}(\pi)=2m$) of the above theorem:

Theorem 26 *If $\Delta=\Delta_1+i\Delta_2\in \operatorname{Diff}_k(m\cdot\mathbf{1}_{\mathbb{C}}, m\cdot\mathbf{1}_{\mathbb{C}})$ is a complex smooth operator of principal type and the homogeneous Hamiltonian $H+iF=\det[\sigma_\Delta]$ of order mk satisfies condition $(\mathcal{P})$, then the system of linear PDEs $\Delta(u)=f$ is locally solvable for any smooth vector-valued function $f\in C^\infty(M;\mathbb{C}^m)$.*

Note however that the principal type condition of [77] is formulated so that multiple characteristics are excluded (this is equivalent to the claim that $\mathcal{K}$ is a 1-dimensional bundle over $\operatorname{Char}^{\mathbb{C}}(\mathcal{E})$), though in some cases this condition can be relaxed.

For general systems the solvability question is still open and one can be tempted to approach it via successive sequence of determined systems, like in Cartan-Kähler theorem (see Guillemin's normal forms in [31, 7]).

Remark 7 The solutions obtained via the above methods are usually distributions, though in some cases they can be proved to be smooth by using elliptic regularity or Sobolev's embedding theorem [35, 80]. The methods can be generalized to weakly non-linear situations, but for strongly non-linear PDEs effects of multi-valued solutions require new insight [58].

Finally let us consider an important case of evolutionary PDEs $u_t=L[u]$, where L is a non-linear differential operator involving only $\mathcal{D}_x$ differentiations, in the splitting of base coordinates $\mathbb{R}\times U=\{(t,x)\}$. The Cauchy problem for such systems is often posed on the characteristic submanifold $\Sigma^{n-1}=\{t=0\}$, which contradicts the approach of Cartan-Kähler theorem.

Nevertheless in many cases it is possible to show that the solution exists. For instance, consider the system $\partial_t u = Au + F(t, x, u)$ with A being a determined linear differential operator on the space $\mathcal{W}$ of smooth vector-functions of x and F can be non-linear (usually of lower order). If the homogeneous linear system $\partial_t u = Au$ is solvable and e^{At} is a semi-group (on a certain Banach completion of $\mathcal{W}$), then provided that F is Lipschitz on $\mathcal{W}$, we can guarantee existence of a local solution to the initial value problem $u(0, x) = u_0(x)$ (in fact weak solutions; strong solutions are guaranteed if $\mathcal{W}$ can be chosen a reflexive Banach space [93]).

This scheme works well for differential operators A with constant coefficients. Moreover, global solvability can be achieved. Consider, for instance a non-autonomous reaction-diffusion equation

$$\partial_t u = a\Delta u - f(t, u) + g(t, x),$$

where $x \in U \Subset \mathbb{R}^{n-1}$, $a \in \mathrm{Gl}_+(\mathbb{R}^m)$ is a positive constant matrix, Δ the Laplace operator and the functions f, g belong to certain Hölder spaces. The boundary behavior is governed by Dirichlet or Neumann or periodic conditions. Then provided that function f has a limited growth behavior at infinity (see [11] for details) the initial problem $u(0, x) = u_0(x)$ for this system is globally solvable.

Similar schemes (with characteristic Cauchy problems) work also for PDEs involving higher derivatives in t, for example damped hyperbolic equation [11]. This allows to consider evolutionary PDEs as dynamical systems. In fact, bracket approach for compatibility and generalized Lagrange-Charpit method of §4.4 allows to establish and investigate finite-dimensional sub-dynamics, see [42, 67].

4.3 Dimension of the solutions space

In his study of systems $\mathcal{E}$ of PDEs [10] (interpreted as exterior differential systems) Cartan constructed a sequence of numbers s_i, which are basic for his involutivity test. These numbers depend on the flag of subspaces one chooses for investigation of the system and so have no invariant meaning.

The classical formulation is that a general solution depends on s_p functions of p variables, s_{p-1} functions of $(p-1)$ variables, $\ldots$, s_1 functions of 1 variable and s_0 constants (we adopt here the notations from [7]; in Cartan's notations [10] we should rather write s_p, $s_p + s_{p-1}$, $s_p + s_{p-1} + s_{p-2}$ etc). However as Cartan notices just after the formulation [10], this statement has only a calculational meaning.

Nevertheless two numbers are absolute invariants and play an important role. These are Cartan genre, i.e. the maximal number p such that $s_p \neq 0$, but $s_{p+1} = 0$, and Cartan integer $\sigma = s_p$. As a result of Cartan's test a general solution depends on σ functions of p variables (and some number of functions of lower number of variables, but this number can vary depending on a way we parametrize the solutions).

Here in analytical category a general solution is a local analytic solution obtained as a result of application of Cartan-Kähler theorem and thus being parametrized by the Cauchy data. In smooth category one needs a condition to ensure existence of solutions with any admissible jet, see §4.1-4.2.

In general we can calculate these numbers in formal category. We call p *functional dimension* and σ *functional rank* of the solutions space $\mathrm{Sol}(\mathcal{E})$ [47]. These numbers can be computed via the characteristic variety. If the characteristic sheaf $\mathcal{K}$ over $\mathrm{Char}^{\mathbb{C}}(\mathcal{E})$ has fibers of dimension k, then

$$p = \dim \mathrm{Char}^{\mathbb{C}}(\mathcal{E}) + 1,\ \sigma = k \cdot \deg \mathrm{Char}^{\mathbb{C}}(\mathcal{E}).$$

The first formula is a part of Hilbert-Serre theorem ([33]), while the second is more complicated. Actually Cartan integer σ was calculated in [7] in general situation and the formula is as follows.

Let $\mathrm{Char}^{\mathbb{C}}(g) = \cup_\epsilon \Sigma_\epsilon$ be the decomposition of the characteristic variety into irreducible components and $d_\epsilon = \dim \mathcal{K}_x$ for a generic point $x \in \Sigma_\epsilon$. Then

$$\sigma = \sum d_\epsilon \cdot \deg \Sigma_\epsilon.$$

The clue to this formula is commutative algebra. Namely *Hilbert polynomial* ([33]) of the symbolic module g^* equals

$$P_{\mathcal{E}}(z) = \sigma z^p + \dots$$

A powerful method to calculate the Hilbert polynomial is resolution of a module. In our case a resolution of the symbolic module g^* exists and it can be expressed via the Spencer δ-cohomology. Indeed, the Spencer cohomology of the symbolic system g is $\mathbb{R}$-dual to the Koszul homology of the module g^* and for algebraic situation this resolution was found in [26].

This yields the following formulae [47]. Let

$$\tbinom{z+k}{k} = \frac{1}{k!}(z+1)\cdot(z+2)\cdots(z+k).$$

Denote $S_j(k_1,\dots,k_n) = \sum\limits_{i_1<\dots<i_j} k_{i_1}\cdots k_{i_j}$ the j-th symmetric polynomial and let also

$$s_i^n = \frac{(n-i)!}{n!}\, S_i(1,\dots,n)$$

Thus

$$s_0^n = 1, \quad s_1^n = \frac{n+1}{2}, \quad s_2^n = \frac{(n+1)(3n+2)}{4\cdot 3!}, \quad s_3^n = \frac{n(n+1)^2}{2\cdot 4!},$$

$$s_4^n = \frac{(n+1)(15n^3+15n^2-10n-8)}{48\cdot 5!} \qquad \text{etc.}$$

If we decompose

$$\tbinom{z+n}{n} = \sum_{i=0}^{n} s_i^n \frac{z^{n-i}}{(n-i)!},$$

then we get the expression for the Hilbert polynomial

$$P_{\mathcal{E}}(z) = \sum_{i,j,q} (-1)^i h^{q,i} s_j^n \frac{(z-q-i)^{n-j}}{(n-j)!} = \sum_{k=0}^{n} b_k \frac{z^{n-k}}{(n-k)!},$$

where

$$b_k = \sum_{j=0}^{k} \sum_{q,i} (-1)^{i+j+k} h^{q,i} s_j^n \frac{(q+i)^{k-j}}{(k-j)!}.$$

Let us compute these dimensional characteristics p, σ for two important classes of PDEs.

If $\mathcal{E}$ is an involutive systems, then $H^{i,j}(\mathcal{E}) = 0$ for $i \notin \operatorname{ord}(\mathcal{E}) - 1$, $(i,j) \neq (0,0)$, and the above formula becomes more comprehensible.

Let us restrict for simplicity to the case of systems of PDEs $\mathcal{E}$ of pure first order. Then

$$P_{\mathcal{E}}(z) = h^{0,0}\binom{z+n}{n} - h^{0,1}\binom{z+n+1}{n+1} + h^{0,2}\binom{z+n+2}{n+2} - \dots$$
$$= b_1 \frac{z^{n-1}}{(n-1)!} + b_2 \frac{z^{n-2}}{(n-2)!} + \dots + b_0.$$

Vanishing of the first coefficient $b_0 = 0$ is equivalent to vanishing of Euler characteristic for the Spencer δ-complex, $\chi = \sum_i (-1)^i h^{0,i} = 0$, and this is equivalent to the claim that not all the covectors from ${}^{\mathbb{C}}T^* \setminus 0$ are characteristic for the system g.

The other numbers b_i are given by the above general formulas, though now they essentially simplify. For instance

$$b_1 = \tfrac{n+1}{2} b_0 - \sum (-1)^i h^{0,i} i = \sum (-1)^{i+1} i \cdot h^{0,i}.$$

If $\operatorname{codim} \operatorname{Char}^{\mathbb{C}}(\mathcal{E}) = n - p > 1$, then $b_1 = 0$ and in fact then $b_i = 0$ for $i < n - p$, but $b_{n-p} = \sigma$.

Theorem 27 *[47]. If* $\operatorname{codim} \operatorname{Char}^{\mathbb{C}}(\mathcal{E}) = n - p$, *then the functional rank of the system equals*

$$\sigma = \sum_i (-1)^i h^{0,i} \frac{(-i)^{n-p}}{(n-p)!}.$$

One can extend the above formula for general involutive system and thus compute the functional dimension and functional rank of the solutions space (some interesting calculations can be found in classical works [37, 10]).

Consider also an important partial case of Cohen-Macaulay systems:

Theorem 28 *[52]. Let $\mathcal{E}$ be a formally integrable system of generalized complete intersection type with orders $k_1, \dots, k_r$. Then the space $\mathcal{S}ol_{\mathcal{E}}$ has formal functional dimension and rank equal respectively*

$$p = m + n - r - 1, \qquad \sigma = S_{r-m+1}(k_1, \dots, k_r).$$

4.4 Integrability methods

Most classical methods for integration of PDEs are related to symmetries ([61, 24, 20]).

A *symmetry* of a system $\mathcal{E}$ is a Lie transformation of $J^k\pi$, resp. $J^k(E,m)$, that preserves $\mathcal{E}$ (k is the maximal order of $\mathcal{E}$). Internal symmetry is a structural diffeomorphism of $\mathcal{E}$, i.e. a diffeomorphism of $\mathcal{E}_k$ (not necessary inducing diffeomorphisms of $\mathcal{E}_l$ for $l \le k$) that preserves the Cartan distributions $\mathcal{C}_{\mathcal{E}_k}$. In many important cases, the systems $\mathcal{E}$ are rigid [40], in which case internal and external symmetries coincide.

In practice the group (in fact, pseudo-group, see the next section) of symmetries $\mathrm{Sym}(\mathcal{E})$ is difficult to calculate and it is much easier to work with the corresponding Lie algebra of infinitesimal symmetries ([62, 63]). These are Lie vector fields X_φ on the space[8] $J^k\pi$, which are tangent to $\mathcal{E}$.

The generating function φ has order 0 or 1 in the classical case (point or contact transformations). Equation for φ to be a symmetry of a system $\mathcal{E} = \{F_\alpha = 0\}$ can be written in the form (for some differential operators Q_α):

$$X_\varphi(F_\alpha)|_{\mathcal{E}} = 0 \quad \Leftrightarrow \quad \ell_{F_\alpha}\varphi = \sum Q_\alpha F_\alpha.$$

Notice that when the system is scalar, i.e. $\pi = \mathbf{1}$, and $\deg F_\alpha = k_\alpha$, $\deg\varphi = \varkappa$, then the defining equations can be written in the form

$$\{F_\alpha, \varphi\} = 0 \mod \mathcal{J}_{k_\alpha+\varkappa-1}(\mathcal{E}). \tag{15}$$

When $\varphi \in \mathfrak{F}_i$, $i > 1$, the field X_φ does not define a flow on any finite jet-space, but rather on $J^\infty(\pi)$. If this flow leaves $\mathcal{E}_\infty$ invariant, then φ (or X_φ) is called *higher symmetry* ([40]). Denoting by $\ell_F^{\mathcal{E}}$ the restriction $\ell_F|_{\mathcal{E}_\infty}$ we obtain the defining equations of higher symmetries:

$$\varphi \in \mathrm{sym}(\mathcal{E}) \quad \Leftrightarrow \quad \ell^{\mathcal{E}}_{F_\alpha}(\varphi) = 0.$$

Here $\mathrm{sym}(\mathcal{E}) = D_{\mathcal{C}}(\mathcal{E}_\infty)/\mathcal{C}D(\mathcal{E}_\infty)$ is the quotient of the Lie algebra $D_{\mathcal{C}}$ of all symmetries of the Cartan distribution $\mathcal{C}_{\mathcal{E}}$ on $\mathcal{E}_\infty$ by the space $\mathcal{C}D$ of trivial symmetries, tangent to the distribution $\mathcal{C}_{\mathcal{E}}$.

Conservation laws ω_ψ with generating function ψ are obtained from the dual equation

$$(\ell^{\mathcal{E}}_{F_\alpha})^*(\psi) = 0,$$

where Δ^* is the formally dual to an operator Δ.

Remark 8 Both symmetries and conservation laws enter variational bi-complex or equivalently $\mathcal{C}$-spectral sequence for the system $\mathcal{E}$, see [99, 92, 53, 1] and references therein.

Notice that classical (point and contact) symmetries as well as classical conservation laws are widely used to find classes of exact solutions and partially integrate the system, see [13, 78]. In fact, almost all known exact methods are based on the idea of symmetry or intermediate integral [25, 20].

Due to Corollary 17 this also holds for higher symmetries/conservation laws. Indeed if $\mathcal{G} = \langle\varphi_1, \dots, \varphi_s\rangle \subset \mathrm{sym}(\mathcal{E})$ is a Lie subalgebra of symmetries of a compatible system $\mathcal{E}$, then the joint system $\tilde{\mathcal{E}} = \mathcal{E} \cap \{\varphi_1 = 0, \dots, \varphi_s = 0\}$, provided that regularity assumptions are satisfied (this includes certain non-degeneracy condition), is compatible too.

Classical Lagrange-Charpit method [24, 27] for first order PDEs consists in a special type overdetermination of the given system $\mathcal{E}$, so that the new system is again compatible[9].

[8] We take the affine chart to have formulas (2) representing X_φ.

[9] This stays in contrast with the method of differential ansatz, where the additional equations are imposed with only condition that the joint system is solvable.

Generalized Lagrange-Charpit method [45] works for any system of PDEs and it also consists in overdetermination to a compatible system.

For systems of scalar PDEs it is often more convenient to impose additional equations $F_{r+1},\dots,F_{r+s}$ to the system $\mathcal{E}=\{F_1=0,\dots,F_r=0\}$, so that the joint system $\tilde{\mathcal{E}}=\{F_1=0,\dots,F_{r+s}=0\}$ is of complete intersection type. Then if $\mathcal{E}$ is compatible, the compatibility of the sub-system $\tilde{\mathcal{E}}\subset\mathcal{E}$ can be expressed as follows (see Corollary 17):

$$\{F_i,F_j\}=0 \bmod \mathcal{J}_{k_i+k_j-1}(\tilde{\mathcal{E}}) \quad \text{for} \quad 1\le i\le r+s,\ r<j\le r+s.$$

Note that (15) is a particular case of these equations. For a system of vector PDEs ($\operatorname{rank}\pi>1$) the corresponding situation, when the compatibility condition writes effectively, should be the generalized complete intersection (see Theorem 16), and then the conditions of generalized Lagrange-Charpit method can be written via multi-brackets.

Let us remark that intermediate integrals are partial cases of this approach (we called additional PDEs $F_{r+1}=0,\dots,F_{r+s}=0$ *auxiliary integrals* in [44]). More generally, most integrability schemes (Lax pairs, Sato theory, commuting hierarchies etc) are closely related to compatibility criteria.

For instance, Backlund transformations [83, 36] can be treated as follows. Let $\mathcal{E}_1=\{F_1=0,\dots,F_r=0\}\subset J^\infty(\pi_1)$ be a compatible system. Extend $\pi_1\hookrightarrow\pi=\pi_1\oplus\pi_2$ and let us impose new PDEs $\{F_{r+1}=0,\dots,F_{r+s}=0\}$, which are not auxiliary integrals in the sense that the joint system $\tilde{\mathcal{E}}=\{F_1=0,\dots,F_{r+s}=0\}$ is not compatible. If the compatibility conditions modulo the system $\mathcal{E}_1$ are reduced to a compatible system $\mathcal{E}_2\subset J^\infty(\pi_2)$, then any solution of $\mathcal{E}_1$ gives (families of) solutions of $\mathcal{E}_2$.

For the sin-Gordon equation $u_{xy}=\sin u$ the additional equations are $v_x=\sin w$, $w_y=\sin v$, $u=v+w$ and we get $\mathcal{E}_2=\mathcal{E}_1$ for $w=u-v$; here $\pi_1=\mathbf{1}$ (fiber coordinate u) and $\pi_2=\mathbf{1}$ (fiber coordinate w).

Finally consider the classical Darboux method of integrability [14, 25, 3]. It is applied to hyperbolic second-order PDEs $F=0$ on the plane (if quasi-linear, then local point transformation brings it to the form $u_{xy}=f(x,y,u,u_x,u_y)$; in general denote the characteristic fields by X,Y), which by a sequence of Laplace transformations reduce to the trivial PDE $u_{xy}=0$.

In this case the equation possesses a closed form general solution depending on two arbitrary functions of 1 variable. They are obtained via a pair of intermediate integrals $I_1=0$, $I_2=0$, such that the system $\{F=0,I_1=0\}$ is compatible and has one common characteristic X, while the system $\{F=0,I_2=0\}$ is compatible and has one common characteristic Y. All three equations are compatible as well (and this system is already free of characteristics, i.e. of finite type).

For Liouville equation $u_{xy}=e^u$ the pair of second order intermediate integrals is $I_1=u_{xx}-\frac{1}{2}u_x^2=f(x)$ and $I_2=u_{yy}-\frac{1}{2}u_y^2=g(y)$, i.e. we have $\mathcal{D}_y(I_1)=0$ and $\mathcal{D}_x(I_2)=0$ on $\mathcal{E}$.

Thus Darboux method can be treated as a particular case of generalized Lagrange-Charpit method, but in this case we relax the condition of complete intersection (for overdetermined system in dimension two this yields $\operatorname{Char}^{\mathbb{C}}(\mathcal{E})=\emptyset$) to possibility of common characteristics (in this case criterion of Theorem 16 fails and compatibility conditions become of lower orders and simpler).

4.5 Pseudogroups and differential invariants

Let a group G act on the manifold E by diffeomorphisms. Its action lifts naturally to the jet-space $J^k(E,m)$. An important modification of this situation is when G acts by contact transformations on $J^1(E,1)$.

A general G-representation via Lie transformations is a prolongation of one of these by the Lie-Backlund theorem, see §1.4. We also investigate group actions on differential equations $\mathcal{E}$. We again require that the group acts by symmetries of $\mathcal{C}_{\mathcal{E}}$, but now they need not to be external, and if the system is not rigid (§4.4), they may not to be prolongation of point of contact symmetries.

It is often assumed that G is a Lie group, because then one can exploit the formulas of §1.5 to lift transformations to the higher jets, without usage of the inverse function theorem.

A function I is called a *differential invariant* of order k with respect to the action of G, if it is constant on the orbits $G^k \cdot x_k \subset J^k(E,m)$ of the lifted action. For connected Lie groups this writes simpler: $\hat{X}(I)=0$, $X \in \mathfrak{g}$, where $\mathfrak{g} = \mathrm{Lie}(G)$ is the corresponding Lie algebra.

Denote by $\mathcal{I}_k$ the algebra of differential invariants of order $\leq k$. Then $\mathcal{I} = \cup \mathcal{I}_k$ is a filtered algebra, with the associated graded algebra $\mathfrak{O} = \oplus \mathfrak{O}^k$ called the algebra of covariants ([49]). The latter plays an important role in setting a Spencer-type calculus for pseudo-groups ([50]).

Similar to invariant functions there are defined invariant (multi-) vector fields, invariant differential forms, various invariant tensors, differential operators on jet-spaces etc.

Invariant differentiations play a special role in producing other differential invariants. Levi-Civita connection is one of the most known examples. Tresse derivatives are the very general class of such operations and they are defined as follows.

Suppose we have $n = \dim E - m$ differential invariants $f_1, \dots, f_n$ on $\mathcal{E}_k \subset J^k(E,m)$. Provided $\pi_{k+1,k}(\mathcal{E}_{k+1}) = \mathcal{E}_k$ we define the differential operator

$$\hat{\partial}_i : C^\infty(\mathcal{E}_k) \to C^\infty(\mathcal{E}'_{k+1}),$$

where $\mathcal{E}'_{k+1}$ is the open set of points $x_{k+1} \in \mathcal{E}_{k+1}$ with

$$df_1 \wedge \ldots \wedge df_n|_{L(x_{k+1})} \neq 0. \tag{16}$$

We require that $\{f_i\}_{i=1}^n$ are such that $\mathcal{E}'_{k+1}$ is dense in $\mathcal{E}_{k+1}$. For the trivial equation $\mathcal{E}_i = J^i(E,m)$ this is always the case. But if the equation $\mathcal{E}$ is proper, this is a requirement of "general position" for it. Given condition (16) we write:

$$df|_{L(x_{k+1})} = \sum_{i=1}^n \hat{\partial}_i(f)(x_{k+1})\, df_i|_{L(x_{k+1})},$$

which defines the function $\hat{\partial}_i(f)$ uniquely at all the points $x_{k+1} \in \mathcal{E}'_{k+1}$. This yields an invariant differentiation $\hat{\partial}_i = \hat{\partial}/\hat{\partial} f_i : \mathcal{I}_k \to \mathcal{I}_{k+1}$. The expressions $\hat{\partial}_i(f) = \hat{\partial} f/\hat{\partial} f_i$ are called Tresse derivatives of f with respect to f_i ([50]).

For affine charts $J^k(\pi) \subset J^k(E,m)$ this definition coincides with the classical one ([98, 96, 78]). Consider some examples of calculations of scalar differential invariants[10].

[10]Some of these facts are contained in classical textbooks. We obtained the formulas thanks to the wonderful Mapple-11 package DiffGeom by I.Anderson.

(1) *Diffeomorphisms of the projective line.*

1a. Left SL_2*-action.* For a diffeomorphism $f : \mathbb{RP}^1 \to \mathbb{RP}^1$ and $g \in \mathrm{SL}_2(\mathbb{R})$ define the left action by $g(f) = g \circ f$. The corresponding Lie algebra $\mathfrak{g} = \mathrm{sl}(2)$ is generated by the vector fields $\langle \partial_u, u\partial_u, u^2\partial_u \rangle$ on $J^0(\mathbb{R})$. The algebra $\mathcal{I}$ of differential invariants is generated by x, the Schwartz derivative

$$j_3 = \frac{2p_1p_3 - 3p_2^2}{2p_1^2} \text{ and all total derivatives } \mathcal{D}_x^k(j_3),\ k > 0.$$

1b. Right SL_2*-action.* The right action of $G = \mathrm{SL}_2(\mathbb{R})$ on $\mathbb{RP}^1$ is defined by the formula: $g(f) = f \circ g^{-1}$. The corresponding Lie algebra $\mathfrak{g} = \mathrm{sl}(2)$ is generated by the vector fields $\langle \partial_x, x\partial_x, x^2\partial_x \rangle$ on $J^0(\mathbb{R})$. The algebra $\mathcal{I}$ of differential invariants is generated by u, the inverse Schwartz derivative

$$J_3 = \frac{2p_1p_3 - 3p_2^2}{2p_1^4} \text{ and the Tresse derivatives } \frac{\hat{\partial}^k}{\hat{\partial} u^k}(J_3),\ k > 0.$$

(2) *Curves in the classical plane geometries.*

2a. Metric plane. The Lie algebra of plane motions $\mathfrak{m}_2$ is generated by the vector fields $\langle \partial_x, \partial_u, x\partial_u - u\partial_x \rangle$ on the plane $\mathbb{R}^2 = J^0(\mathbb{R})$. There is an $\mathfrak{m}_2$-invariant differentiation (metric arc)

$$\nabla = \frac{1}{\sqrt{p_1^2 + 1}} \frac{d}{dx},$$

and the algebra of $\mathfrak{m}_2$-differential invariants is generated by the curvature

$$\kappa_2 = \frac{p_2}{(p_1^2 + 1)^{3/2}} \text{ and the derivatives } \nabla^r \kappa_2,\ r > 0.$$

2b. Conformal plane. The Lie algebra of plane conformal transformations $\mathfrak{co}_2$ is generated by the vector fields $\langle \partial_x, \partial_u, x\partial_u - u\partial_x, x\partial_x + u\partial_u \rangle$ on the plane $\mathbb{R}^2 = J^0(\mathbb{R})$. There is a $\mathfrak{co}_2$-invariant differentiation (conformal arc)

$$\nabla = \frac{p_1^2 + 1}{p_2} \frac{d}{dx},$$

and the algebra of $\mathfrak{co}_2$-differential invariants is generated by the conformal curvature

$$\kappa_3 = \frac{p_1^2 p_3 + p_3 - 3p_1p_2^2}{p_2^2} \text{ and the derivatives } \nabla^r \kappa_3,\ r > 0.$$

2c. Symplectic plane. The Lie algebra of plane symplectic transformations $\mathfrak{sp}_2$ is generated by the vector fields $\langle \partial_x, \partial_u, x\partial_u, u\partial_x, x\partial_x - u\partial_u \rangle$ on the plane $\mathbb{R}^2 = J^0(\mathbb{R})$. There is an $\mathfrak{sp}_2$-invariant differentiation (symplectic arc)

$$\nabla = \frac{1}{\sqrt[3]{p_2}} \frac{d}{dx},$$

and the algebra of $\mathfrak{sp}_2$-differential invariants is generated by the symplectic curvature

$$\kappa_4 = \frac{3p_2p_4 - 5p_3^2}{3p_2^{8/3}} \text{ and the derivatives } \nabla^r \kappa_4,\ r > 0.$$

2d. Affine plane. The Lie algebra of plane affine transformations $\mathfrak{a}_2$ is generated by the vector fields $\langle \partial_x, \partial_u, x\partial_u, u\partial_x, x\partial_x, u\partial_u \rangle$ on the plane $\mathbb{R}^2 = J^0(\mathbb{R})$. There is an $\mathfrak{a}_2$-invariant differentiation (affine arc)

$$\nabla = \frac{p_2}{\sqrt{3p_2p_4 - 5p_3^2}} \frac{d}{dx},$$

and the algebra of $\mathfrak{a}_2$-differential invariants is generated by the affine curvature

$$\kappa_5 = \frac{9p_2^2p_5 + 40p_3^3 - 45p_2p_3p_4}{9(3p_2p_4 - 5p_3^2)^{3/2}} \quad \text{and the derivatives } \nabla^r \kappa_5,\ r > 0.$$

2e. Projective Plane. The Lie algebra of plane projective transformations $\mathfrak{sl}_3$ is generated by the vector fields $\langle \partial_x, \partial_u, x\partial_u, u\partial_x, x\partial_x, u\partial_u, x^2\partial_x + xu\partial_u, xu\partial_x + u^2\partial_u \rangle$ on the plane $\mathbb{R}^2 = J^0(\mathbb{R})$.

There are two relative differential invariants:

$$\Theta_3 = \frac{-9p_2^2p_5 + 45p_2p_3p_4 - 40p_3^3}{54p_2^3}, \quad \Theta_8 = 6\Theta_3 \frac{d^2\Theta_3}{dx^2} - 7\left(\frac{d\Theta_3}{dx}\right)^2$$

of degrees 3 and 8, and of orders 5 and 7 respectively. There is also an $\mathfrak{sl}_3$-invariant differentiation (projective arc, or Study invariant differentiation)

$$\nabla = \frac{1}{\sqrt[3]{\Theta_3}} \frac{d}{dx},$$

and the algebra of $\mathfrak{sl}_3$-differential invariants is generated by the projective curvature

$$\kappa_7 = \frac{\Theta_8^3}{\Theta_3^8} \quad \text{and the derivatives } \nabla^r \kappa_7,\ r > 0.$$

Pseudogroups are infinite-dimensional Lie groups, which can be obtained by integrating Lie equations [9, 17, 57, 86, 55]. Differential invariants and Tresse derivatives are defined for them in the same manner.

Theorem 29 *Algebra $\mathcal{I}$ of differential invariants of pseudogroup G action is finitely generated by algebraic operations and Tresse derivatives.*

This theorem (with a proper assumption of regularity) was formulated and sketched by A. Tresse [98], though important partial cases were considered before by S. Lie [61] (see also [32]). The proof for (finite-dimensional) Lie groups was given by Ovsiannikov [79], for pseudogroups acting on jet-spaces by Kumpera [54]. The general case of pseudogroups G acting on systems of PDEs $\mathcal{E}$ was completed in [50].

Similar to Cartan-Kuranishi theorem one hopes that generic points of $\mathcal{E}$ are regular. This is possible to show in good (algebraic/analytic) situations.

Pseudogroups constitute a special class of Lie equations. With general approach of [50] one does not require their local integrability from the beginning. It is important that passage from formal integrability to the local one is easier for pseudogroups compared to general systems of PDEs. For instance, formally integrable transitive flat pseudogroups are locally integrable [8, 81].

Pseudogroups are basic for solution of equivalence problem. Pseudogroups are also fundamental for establishing special symmetric solutions of PDEs, they can be used to multiply transversal solutions and in some cases (if the pseudogroup is big enough) to integrate PDEs [79, 40, 13, 78, 92, 2, 58].

4.6 Spencer D-cohomology

The Spencer differential

$$D : \mathcal{J}^k(\pi) \otimes \Omega^l(M) \to \mathcal{J}^{k-1}(\pi) \otimes \Omega^{l+1}(M)$$

is uniquely defined by the following conditions:

(i) D is $\mathbb{R}$-linear and satisfies the Leibniz rule:

$$D(\theta \otimes \omega) = D(\theta) \wedge \omega + \pi_{k,k-1}(\theta) \otimes d\omega, \quad \theta \in \mathcal{J}^k(\pi),\ \omega \in \Omega^l(M).$$

(ii) The following sequence is exact:

$$0 \to C^\infty(\pi) \xrightarrow{j_k} \mathcal{J}^k(\pi) \xrightarrow{D} \mathcal{J}^{k-1}(\pi) \otimes \Omega^1(M).$$

The latter operator can be described as follows. Let $x \in M$, $v \in T_xM$, $\theta \in \mathcal{J}^k(\pi)$ and $x_k = \theta(x) \in \pi_k^{-1}(x)$. Since $\tilde{\theta} = \pi_{k,k-1}(\theta) \in C^\infty(\pi_{k-1})$, the value $D_v\theta = i_v \circ D(\theta) \in J_x^{k-1}\pi$ equals $\rho_{k-1}^{\mathrm{v}} \circ (j_{k-1}\tilde{\theta})_*(v)$, where

$$\rho_{k-1}^{\mathrm{v}} : T_{x_{k-1}}(J^{k-1}\pi) \simeq L(x_k) \oplus J_x^{k-1}\pi \to J_x^{k-1}\pi$$

is the projection to the second component (the splitting depends only on x_k). Thus $D(\theta) = 0$ if and only if $\tilde{\theta}(M)$ is an integral manifold of the Cartan distribution on $J^{k-1}(\pi)$ and therefore has the form $j_{k-1}(s)$, which yields $\theta = j_k(s)$ for some $s \in C^\infty(\pi)$.

The above geometric description implies that the Spencer operator D is natural, $D \circ \pi_{k+1,k} = \pi_{k,k-1} \circ D$. Moreover let $\alpha : E_\alpha \to M_\alpha$ and $\beta : E_\beta \to M_\beta$ be two vector bundles and $\Psi : \alpha \to \beta$ be a morphism over a smooth map $\psi : M_\alpha \to M_\beta$, $\psi \circ \alpha = \beta \circ \Psi$, such that $\Psi_x : \alpha^{-1}(x) \to \beta^{-1}(\psi(x))$ are linear isomorphisms for all $x \in M$. Then Ψ generates a map of sections: $\Psi^* : C^\infty(\beta) \to C^\infty(\alpha)$, where $\Psi^*(h)(x) = \Psi_x^{-1}\big(h(\psi(x))\big)$. This in turn generates a map of k-jets: $\Psi_k^* : \mathcal{J}^k(\beta) \to \mathcal{J}^k(\alpha)$ and $\Psi_k^* \circ j_k = j_k \circ \Psi^*$. Then naturality of D means that

$$D \circ \Psi_k^* \otimes \psi^* = \Psi_{k-1}^* \otimes \psi^* \circ D.$$

The above properties of the Spencer differential yield $D^2 = 0$. Hence the following sequence is a complex:

$$0 \to C^\infty(\pi) \xrightarrow{j_k} \mathcal{J}^k(\pi) \xrightarrow{D} \mathcal{J}^{k-1}(\pi) \otimes \Omega^1(M) \xrightarrow{D} \cdots \xrightarrow{D} \mathcal{J}^{k-n}(\pi) \otimes \Omega^n(M) \to 0.$$

It is called the first (naive) *Spencer complex*.

Let $\mathcal{E} = \{\mathcal{E}_k \subset J^k\pi\}$ be a system of linear PDEs. Assume that $\mathcal{E}$ is formally integrable. Then the 1-st Spencer complex can be restricted to $\mathcal{E}$, meaning that $D : \underline{\mathcal{E}_k} \to \underline{\mathcal{E}_{k-1}} \otimes \Omega^1(M)$, where $\underline{\mathcal{E}_k} = C^\infty(\pi_k|_{\mathcal{E}_k})$ denotes the space of sections (non-holonomic solutions of $\mathcal{E}_k$). The resulting complex

$$0 \to \underline{\mathcal{E}_k} \xrightarrow{D} \underline{\mathcal{E}_{k-1}} \otimes \Omega^1(M) \xrightarrow{D} \cdots \xrightarrow{D} \underline{\mathcal{E}_{k-n}} \otimes \Omega^n(M) \to 0 \tag{17}$$

is called the first Spencer complex associated with the system $\mathcal{E}$.

The exact sequences $0 \to g_k \to \mathcal{E}_k \to \mathcal{E}_{k-1} \to 0$ induce the exact sequences of the Spencer complexes and this together with δ-lemma shows that the cohomology of the 1st Spencer complex are stabilizing for sufficiently large k. The stable cohomology are called *Spencer D-cohomology* of $\mathcal{E}$ and they are denoted by $H^i_D(\mathcal{E})$, $i = 0, 1, \dots, n$.

Remark that $H^0_D(\mathcal{E}) = \mathrm{Sol}(\mathcal{E})$ is the space of global smooth solutions of $\mathcal{E}$. Other cohomology group $H^i_D(\mathcal{E})$ describe the solutions spaces of the systems of PDEs corresponding to the place i of the Spencer complex and $H^*_D(\mathcal{E})$ is a module over the de Rham cohomology of the base $H^*(M)$.

Due to the summands g_k the first complex is not formally exact (=exact on the level of formal series). The construction of the second (sophisticated) Spencer complex amends this feature. This 2nd complex is defined as follows.

Pick a vector bundle morphism $\Theta : \mathcal{E}_k \to \mathcal{E}_{k+1}$ that is right-inverse to the projection $\pi_{k+1,k}$: $\pi_{k+1,k} \circ \Theta = \mathrm{id}$. Let $D_\Theta = D \circ \Theta : \underline{\mathcal{E}_k} \otimes \Omega^i(M) \to \underline{\mathcal{E}_k} \otimes \Omega^{i+1}(M)$. Another right-inverse $\Theta' : \mathcal{E}_k \to \mathcal{E}_{k+1}$ gives:

$$D_\Theta - D_{\Theta'} : \underline{\mathcal{E}_k} \otimes \Omega^i(M) \to \delta(g_{k+1} \otimes \Omega^i(M)).$$

Therefore for the quotient $C^i_k = \mathcal{E}_k \otimes \Lambda^i T^*M / \delta(g_{k+1} \otimes \Lambda^{i-1} T^*M)$ the factor-operators (denoted by the same letter D) are well-defined and they constitute the factor complex

$$0 \to \underline{C^0_k} \xrightarrow{D} \underline{C^1_{k-1}} \xrightarrow{D} \cdots \xrightarrow{D} \underline{C^n_{k-n}} \to 0,$$

which is called the 2nd Spencer complex. Its cohomology stabilize for sufficiently large k and coincide with stable cohomology of the 1st Spencer complex. Moreover the second Spencer D-complex is formally exact [88].

Another approach to the Spencer D-cohomology is via the compatibility complex. Let $\Delta_1 : C^\infty(\pi_1) \to C^\infty(\pi_2)$ be a differential operator. Denote by $\Delta_2 : C^\infty(\pi_2) \to C^\infty(\pi_3)$ its compatibility operator, i.e. $\Delta_2 \circ \Delta_1 = 0$ and $\mathrm{Im}[\psi^{\Delta_1}_\infty : J^\infty(\pi_1) \to J^\infty(\pi_2)] = \mathrm{Ker}[\psi^{\Delta_2}_\infty : J^\infty(\pi_2) \to J^\infty(\pi_3)]$.

Denoting Δ_3 the compatibility operator for the operator Δ_2 and so on we get the compatibility complex

$$C^\infty(\pi_1) \xrightarrow{\Delta_1} C^\infty(\pi_2) \xrightarrow{\Delta_2} C^\infty(\pi_3) \xrightarrow{\Delta_3} \cdots .$$

Existence of such complexes was proved by Kuranishi (also Goldschmidt, see [88] and references therein) whenever $\mathcal{E} = \mathrm{Ker}(\Delta_1)$ is formally integrable. Moreover any two such formally exact complexes are homotopically equivalent. Hence the 2nd Spencer D-complex provides us with an explicit construction of such a complex.

However the Spencer D-complexes are not necessary minimal in the sense that ranks of the bundles π_i can be reduced. Important method for constructing minimal compatibility complexes comes from resolutions in commutative algebra. In such a form they can even be generalized to the non-linear situation. For (non-linear) systems of generalized complete intersection type the compatibility complexes were constructed in [52].

Remark 9 Since cohomology of a compatibility complex equal $H^*_D(\mathcal{E})$, this gives a way to calculate non-linear Spencer D-cohomology. To define the non-linear version of Spencer D-complex one can use the machinery of §2.4.

4.7 Calculations of Spencer cohomology

Consider some examples.

(a) If $\mathcal{E} = \mathrm{Ker}\big[\Delta : C^\infty(\pi) \to C^\infty(\pi)\big]$ is a determined system of PDEs, $\mathrm{Char}^{\mathbb{C}}(\Delta) \neq \mathbb{P}^{\mathbb{C}}T^*$, then

$$H^0_D(\mathcal{E}) = \mathrm{Ker}(\Delta) = \mathrm{Sol}(\mathcal{E}),\ H^1_D(\mathcal{E}) = \mathrm{Coker}(\Delta) \simeq \mathrm{Ker}(\Delta^*).$$

(b) Spencer D-cohomology of a system $\mathcal{E}$ of PDEs, defined by the de Rham differential $d : C^\infty(M) \to \Omega^1(M)$, coincide with the de Rham's cohomology of the base manifold: $H^*_D(\mathcal{E}) = H^*_{dR}(M)$.

(c) Let $\nabla : C^\infty(\pi) \to C^\infty(\pi) \otimes \Omega^1(M)$ be a flat connection. Then the Spencer cohomology of the corresponding system $\mathcal{E}$ coincide with the de Rham cohomology of M with coefficients in π: $H^*_D(\mathcal{E}) = H^*_\nabla(\pi)$.

(d) Let M be a complex manifold and π a holomorphic vector bundle over it. Denote by $\Omega^{p,q}(\pi)$ the (p,q)-forms on M with values in π. Then the Spencer D-cohomology of the Cauchy-Riemann equation given by the operator $\bar\partial : \Omega^{p,0}(\pi) \to \Omega^{p,1}(\pi)$ are the Dolbeault cohomology $H^*_{\bar\partial}(M, \Omega^p(\pi))$.

(e) Let $\mathcal{E}$ be a formally integrable system of finite type. Then $\pi_{k+1,k} : \mathcal{E}_{k+1} \to \mathcal{E}_k$ are isomorphisms for large k. Thus the Spencer differential $D : \mathcal{E}_k \simeq \mathcal{E}_{k+1} \to \mathcal{E}_k \otimes \Omega^1(M)$ defines a flat (Cartan) connection ∇ in the vector bundle π_k and the Spencer cohomology equal the de Rham cohomology of this connection: $H^*_D(\mathcal{E}) = H^*_\nabla(\mathcal{E}_k)$.

Finally consider the calculations of Spencer cohomology using the technique of spectral sequences. We will investigate a formally integrable system $\mathcal{E} = \{\mathcal{E}_k \subset J^k\pi\}$ of linear PDEs of maximal order l in a bundle $\pi : E \to M$.

Assume that the base manifold M is itself a total space of a fibre bundle $\varkappa : M \to B$. We say that $\varkappa$ is a *noncharacteristic bundle* if all fibres $F_b = \varkappa^{-1}(b)$, $b \in B$, are strongly noncharacteristic for $\mathcal{E}$ in the sense of §3.2.

A vector field X on M is said to be *vertical* if $\varkappa_*(X) = 0$. A differential form $\theta \in \mathcal{E}_i \otimes \Omega^r(M)$ is called *q-horizontal* if $X_1 \wedge \cdots \wedge X_{q+1} \rfloor \theta = 0$ for any vertical vector fields $X_1, \dots, X_{q+1}$ on M. Denote by $\underline{\mathcal{E}_i} \otimes \Omega^r_q(M)$ the module of q-horizontal elements with $\mathcal{E}_i$-values.

Let $F_{p,q} = \underline{\mathcal{E}_{l-p-q}} \otimes \Omega^{p+q}_q(M)$. Then $\{F_{p,q}\}$ gives a filtration of Spencer complex (17) and $D(F_{p,q}) \subset F_{p,q+1}$. Denote by $\{E^{p,q}_r,\ d^{p,q}_r : E^{p,q}_r \to E^{p+r,q-r+1}_r\}$ the spectral sequence associated with this filtration.

In order to describe the spectral sequence we assume that $\varkappa$ is a noncharacteristic bundle and consider the restriction $\bar\pi_b : E_b \to F_b$ of the bundle π to a fibre F_b. Denote the respective restrictions of $\pi_i : \mathcal{E}_i \to M$ to F_b by $\overline{\mathcal{E}_{i;b}}$ (cf. §3.2 for restrictions of symbolic systems). They satisfy the condition $\overline{\mathcal{E}_{i+1;b}} \subset \overline{\mathcal{E}_{i;b}}^{(1)}$. Due to Cartan-Kuranishi prolongation theorem there exists a number i_0 such that $\overline{\mathcal{E}_{i+1;b}} \subset \overline{\mathcal{E}_{i;b}}^{(1)}$ for $i \geq i_0$.

We call system $\overline{\mathcal{E}}(b) = \{\overline{\mathcal{E}_{i;b}} \subset J^i(\bar\pi_b)\}$ the restriction of $\mathcal{E}$ to the fibre F_b. By Theorem 13 involutivity of $\mathcal{E}$ implies involutivity of $\overline{\mathcal{E}} = \overline{\mathcal{E}}(b)$ for all $b \in B$ (in fact, the theorem concerns only symbolic levels, while the claim involves restrictions of the Weyl tensors).

Similar, $\overline{\mathcal{E}}$ is formally integrable provided that $\mathcal{E}$ is formally integrable (but for the needs of Spencer D-cohomology we can restrict to systems $\overline{\mathcal{E}^{|i\rangle}}$ for $i \geq i_0$).

The following theorem is a generalization of the classical Leray-Serre theorem into the context of Spencer cohomology.

Theorem 30 *[66, 69]. Let $\mathcal{E}$ be a formally integrable system of linear PDEs on a bundle π over M and let $\varkappa : M \to B$ be a noncharacteristic bundle. Assume that the Spencer D-cohomology $H_D^*\big(\overline{\mathcal{E}}(b)\big)$ form a smooth vector bundle over B. Then the above spectral sequence $E_r^{p,q}$ converges to the Spencer D-cohomology $H_D^*(\mathcal{E})$ and the first terms of it equal:*

(0) $E_0^{p,q} = F_{p,q}/F_{p+1,q-1} \simeq \mathcal{E}_{l-p-q} \otimes_{C^\infty(M)} [\Omega^q(\varkappa) \otimes_{C^\infty(B)} \Omega^p(B)]$, *where* $\Omega^q(\varkappa) = \Omega^q(M)/\Omega^q_{q-1}(M)$ *is a module of totally vertical q-forms;*

(a) $E_1^{p,q} \simeq H_D^q(\overline{\mathcal{E}}) \otimes \Omega^p(B)$, *the differential* $d_1^{0,q} : H_D^q(\overline{\mathcal{E}}) \to H_D^q(\overline{\mathcal{E}}) \otimes \Omega^1(B)$ *is a flat connection* ∇ *on the bundle of Spencer cohomology* $H_D^q(\overline{\mathcal{E}})$*;*

(b) $E_2^{p,q} \simeq E_2^{p,q} = H_\nabla^p\big(B, H_D^q(\overline{\mathcal{E}})\big)$, *i.e. the usual* ∇*-de Rham cohomology with coefficients in the sheaf of sections of Spencer* D*-cohomology.*

Assuming that the Spencer cohomology $H_D^q(\mathcal{E})$ are finite dimensional we define the Euler characteristic $\chi(\mathcal{E})$ as

$$\chi(\mathcal{E}) = \sum_{i=0}^{n} (-1)^i \dim H^i(\mathcal{E}).$$

Then the above theorem shows that $\chi(\mathcal{E}) = \chi(\overline{\mathcal{E}}) \cdot \chi_B$, where χ_B is the Euler characteristic of B.

Remark 10 Borel theorem on computation of cohomology of homogeneous spaces together with Leray-Serre spectral sequence constitute the base for computations of de Rham cohomology of smooth manifolds. Borel theorem was generalized to the context of Spencer cohomology in [68], when the symmetry group was assumed compact.

Note that δ-estimate from §4.1 guarantees local exactness of the Spencer complex ([91, 74, 75], Theorem 24 is a partial case). Thus Spencer D-cohomology is the cohomology of the base M with coefficients in the sheaf $\mathrm{Sol}_{\mathrm{loc}}(\mathcal{E})$.

Finite-dimensionality of $H^*(\mathcal{E})$ can be guaranteed if the system $\mathcal{E}$ is elliptic and the manifold M is compact. Another situations is the generalization of the above construction, when the manifold M is foliated (not necessary fibered) and the leaves wind over the manifold densely.

Finally we remark that vanishing of the Spencer cohomology $H_D^q(\mathcal{E}) = 0$ means global solvability of the PDEs corresponding to the operator D at the q-th place of the Spencer complex, provided that compatibility conditions are satisfied.

Bibliography

[1] I. M. Anderson: The variational bicomplex *Preprint* Utah State Univ. (1989); `http://www.math.usu.edu/~fg_mp`

[2] I. M. Anderson, M. E. Fels, C. G. Torre: Group invariant solutions in mathematical physics and differential geometry *In: The geometrical study of differential equations* (Contemp. Math. 285, A.M.S., 2001) 51–60

[3] I. M. Anderson, N. Kamran: The variational bicomplex for hyperbolic second-order scalar partial differential equations in the plane *Duke Math. J.* **87(2)** (1997) 265–319

[4] R. Beals, C. Fefferman: On local solvability of linear partial differential equations *Ann. of Math. (2)* **97** (1973) 482–498

[5] W. Bruns, J. Herzog: *Cohen-Macaulay rings* (Cambridge University Press, Cambridge, U.K., 1993)

[6] D. A. Buchsbaum, D. S. Rim: A generalized Koszul complex, II. Depth and multiplicity *Trans. A.M.S.* **111** (1964) 197–224

[7] R. L. Bryant, S. S. Chern, R. B. Gardner, H. L. Goldschmidt, P. A. Griffiths: *Exterior differential systems* (MSRI Publications 18, Springer-Verlag, 1991)

[8] C. Buttin, P. Molino: Theoreme general d'equivalence pour les pseudogroupes de Lie plats transitifs *J. Diff. Geometry* **9** (1974) 347–354

[9] E. Cartan: Sur la structure des groupes infinis de transformations *In: Oeuvres completes II* (2, Gautier-Villars, Paris, 1953) 571–714

[10] E. Cartan: *Les systèmes différentiels extérieurs et leurs applications géométriques* (Actualités Sci. Ind. 994, (French), Hermann, Paris, 1945)

[11] V. V. Chepyzhov, M. I. Vishik: *Attractors for equations of mathematical physics* (A.M.S. Colloquium Publications 49, Providence, RI, 2002)

[12] Y. Choquet-Bruhat (Foures-Bruhat): Theoreme d'existence pour certains systemes d'equations aux derivees partielles non lineaires *Acta Math.* **88** (French), (1952) 141–225

[13] W. F. Ames, R. L. Anderson, V. A. Dorodnitsyn, E. V. Ferapontov, R. K. Gazizov, N. H. Ibragimov, S. R. Svirshchevskiĭ: *CRC handbook of Lie group analysis of differential equations. Vol. 1. Symmetries, exact solutions and conservation laws* (CRC Press, Boca Raton, 1994)

[14] G. Darboux: *Leçons sur la théorie générale des surfaces et les applications géométriques du calcul infinitésimal* (I–IV partie. Réimpr. de la edition de 1887-1915-1894-1896. Chelsea Publishing Co., Bronx, N.Y., 1972)

[15] N. Dencker: The resolution of the Nirenberg-Treves conjecture *Ann. of Math. (2)* **163(2)** (2006) 405–444

[16] D. Eisenbud: *Commutative algebra with a view toward algebraic geometry* (Springer-Verlag, 1995)

[17] C. Ehresmann: Introduction à la théorie des structures infinitésimales et des pseudogroupes de Lie *In: Colloque de topologie et géométrie différentielle* (no. 11, 16 pp. (French), Strasbourg, 1952)

[18] D. W. Ebner: How Hilbert has found the Einstein equations before Einstein and forgeries of Hilbert's page proofs *Preprint* `arXiv:physics/0610154` (2006)

[19] Yu. Egorov, M. A. Shubin: *Linear PDEs. Foundations of the classical theory* ((transl. from the 1988 Russian original), Partial differential equations I, Encyclopaedia Math. Sci. 30, Springer, 1992)

[20] A.R. Forsyth: *Theory of differential equations* (Vol.1-6., Cambridge University Press, 1900–1906)

[21] S. P. Finikov: *Cartan's Method of Exterior Forms in Differential Geometry* ((Russian), OGIZ, Moscow-Leningrad, 1948)

[22] V. Goldberg, V. Lychagin: On the Blaschke conjecture for 3-webs *J. Geom. Anal.* **16(1)** (2006) 69–115

[23] H. Goldschmidt: Integrability criteria for systems of nonlinear partial differential equations *J. Diff. Geom.* **1(3)** (1967) 269–307

[24] E. Goursat: *Lecons sur l'intégration des équations aux dérivées partielles du premier ordere* (Hermann, Paris, 1891)

[25] E. Goursat: *Lecons sur l'intégration des équations aux dérivées partielles du second ordere* (II, Hermann, Paris, 1898)

[26] M. L. Green: Koszul cohomology and geometry *In: Lectures on Riemann surfaces* (World Sci. Publ. , 1989) 177–200

[27] N. M. Gunter: *Integration of PDEs of the first order* (ONTI (Russian) Leningrad-Moscow, 1934)

[28] V. Guillemin, S. Sternberg: An algebraic model of transitive differential geometry *Bull. A.M.S.* **70** (1964) 16–47

[29] V. Guillemin, D. Quillen, S. Sternberg: The integrability of characteristics *Comm. Pure Appl. Math.* **23(1)** (1970) 39–77

[30] V. Guillemin: The integrability problem for G-structures *Trans. A.M.S.* **116** (1965) 544–560

[31] V. Guillemin: Some algebraic results concerning the characteristics of overdetermined partial differential equations *Amer. Journal of Math.* **90** (1968) 270–284

[32] M. Halpen: *Sur les invariants differentiels* (Thèse, Paris, 1878)

[33] R. Hartshorne: *Algebraic geometry* (Graduate Texts in Mathematics 52, Springer-Verlag, 1977)

[34] L. Hörmander: On the theory of general partial differential operators *Acta Math.* **94** (1955) 161–248

[35] L. Hörmander: *Linear partial differential operators* (3rd edition, Springer-Verlag, 1969)

[36] N. H. Ibragimov, R. L. Anderson: *Lie-Backlund transformations in applications* (SIAM Studies in Applied Mathematics 1, SIAM, Philadelphia, 1979)

[37] M. Janet: *Leçons sur les systèms d'équations aux derivées partielles* (Gauthier-Villers, Paris, 1929)

[38] E. Kähler: *Einfürhung in die Theorie der Systeme von Differentialgleichungen* (B.G. Teubner, Leipzig, 1934)

[39] F. Klein: *Vergleichende Betrachtungen uber neuere geometrische Forschungen* (Erlangen, Andreas Deichert, 1872)

[40] I. S. Krasilschik, V. V. Lychagin, A. M. Vinogradov: *Geometry of jet spaces and differential equations* (Gordon and Breach, 1986)

[41] B. S. Kruglikov: Note on two compatibility criteria: Jacobi-Mayer bracket vs. differential Gröbner basis *Lobachevskii Journ. Math.* **23** (2006) 57–70

[42] B. Kruglikov, O. Lychagina: Finite dimensional dynamics for Kolmogorov-Petrovsky-Piskunov equation *Lobachevskii J. Math.* **19** (2005) 13–28

[43] B. S. Kruglikov, V. V. Lychagin: Mayer brackets and solvability of PDEs – I *Diff. Geom. and its Appl.* **17** (2002) 251–272

[44] B. S. Kruglikov, V. V. Lychagin: Mayer brackets and solvability of PDEs – II *Trans. Amer. Math. Soc.* **358(3)** (2005) 1077–1103

[45] B. S. Kruglikov, V. V. Lychagin: A compatibility criterion for systems of PDEs and generalized Lagrange-Charpit method *In: A.I.P. Conference Proceedings* (Global Analysis and Applied Mathematics: International Workshop on Global Analysis, 729, no. 1, 2004) 39–53

[46] B. S. Kruglikov, V. V. Lychagin: Multi-brackets of differential operators and compatibility of PDE systems *Comptes Rendus Math.* **342(8)** (2006) 557–561

[47] B. S. Kruglikov, V. V. Lychagin: Dimension of the solutions space of PDEs *In: Global Integrability of Field Theories* (Proc. of GIFT-2006, Ed. J.Calmet, W.Seiler, R.Tucker, 2006) 5–25; `arXiv:math.DG/0610789`

[48] B. S. Kruglikov, V. V. Lychagin: On equivalence of differential equations *Acta et Comment. Univ. Tartuensis Math.* **3** (1999) 7–29

[49] B. S. Kruglikov, V. V. Lychagin: Transitive and transversal actions of pseudogroups on submanifolds *Russian Math. (Izv. VUZ, Matematika)* **48(11)** (engl. 27-40), (2004) 30–44

[50] B. S. Kruglikov, V. V. Lychagin: Invariants of pseudogroup actions: Homological methods and Finiteness theorem *Int. J. Geomet. Meth. Mod. Phys.* **3(5–6)** (2006) 1131–1165

[51] B. S. Kruglikov, V. V. Lychagin: Spencer delta-cohomology, restrictions, characteristics and involutive symbolic PDEs *Acta Math. Applicandae* **95** (2007) 31–50; `arXiv:math.DG/0503124`

[52] B. S. Kruglikov, V. V. Lychagin: Compatibility, multi-brackets and integrability of systems of PDEs *Preprint* Univ. Tromsø 2006-49, `arXiv:math.DG/0610930` (2006)

[53] D. Krupka: Global variational principles: foundations and current problems *In: Global analysis and applied mathematics* (AIP Conf. Proc. 729, Amer. Inst. Phys., NY, 2004) 3–18

[54] A. Kumpera: Invariants differentiels d'un pseudogroupe de Lie. Part I; Part II *J. Differential Geometry* **10(2)** (1975) 289–345; *J. Differential Geometry* **10(3)** (1975) 347–416

[55] A. Kumpera, D. Spencer: *Lie equations. Volume 1: General theory* (Princeton University Press and University Tokyo Press, (1972)

[56] M. Kuranishi: On E. Cartan's prolongation theorem of exterior differential systems *Amer. J. Math.* **79** (1957) 1–47

[57] M. Kuranishi: On the local theory of continuous infinite pseudogroups. I-II. *Nagoya Math. J.* **15** (1959) 225–260; *Nagoya Math. J.* **19** (1961) 55–91

[58] A. Kushner, V. Lychagin, V. Roubtsov: *Contact geometry and non-linear differential equations* (Cambridge University Press, 2006)

[59] N. Lerner: When is a pseudo-differential equation solvable? *Ann. Inst. Fourier* **50(2)** (2000) 443–460

[60] H. Lewy: An example of a smooth linear partial differential equation without solution *Ann. of Math. (2)* **66** (1957) 155–158

[61] S. Lie: *Gesammelte Abhandlungen* (Bd. 1-4, Teubner, Leipzig-Kristiania, 1929)

[62] S. Lie: *Theorie der Transformationsgruppen* ((Zweiter Abschnitt, unter Mitwirkung von Prof.Dr.Friederich Engel), Teubner, Leipzig, 1890)

[63] S. Lie, F. Engel: *Theorie der Transformationsgruppen* (vol. II, Begründungstransformationen, Leipzig, Teubner, 1888-1893)
[64] V. V. Lychagin: Homogeneous geometric structures and homogeneous differential equations *In: A. M. S. Transl.* (The interplay between differential geometry and differential equations, V. Lychagin Eds., ser. 2, 167, 1995) 143–164
[65] V. V. Lychagin: Geometric theory of singularities of solutions of nonlinear differential equations *Itogi Nauki i Tekhniki, Problems in geometry* **20** ((Russian),VINITI, Moscow), (1988) 207–247; *Engl. transl.: J. Soviet Math.* **51(6)** (1990) 2735–2757
[66] V. V. Lychagin: Differential Operators on fiber bundles *Uspekhi Mat. Nauk; Engl.Transl.: Russian Math. Surveys* **40(2)** (1985) 187–188
[67] V. Lychagin, O. Lychagina: Finite dimensional dynamics for evolutionary equations *Nonlinear Dynamics* **48** (2007) 29–48
[68] V. V. Lychagin, L. Zilbergleit: Spencer Cohomology and Symmetry Group *Acta Appl. Math.* **41(1-3)** (1995) 227–245
[69] V. V. Lychagin, L. Zilbergleit: Spencer cohomology *In: Geometry in partial differential equations* (World Sci. Publ., 1994) 215–252
[70] T. Lada, J. Stasheff: Introduction to SH Lie algebras for physicists *Int. J. Theor. Phys.* **32(7)** (1993) 1087–1103
[71] B. Malgrange: Equations de Lie. I-II. *J. Differential Geometry* **6** (1972) 503–522; *J. Differential Geometry* **7** (1972) 117–141
[72] B. Malgrange: *Systèmes différentiels involutifs* ((French), Panoramas et Syntheses, Paris, 2005)
[73] Y. Matsushima: On a theorem concerning the prolongation of a differential system *Nagoya Math. J.* **6** (1953) 1–16
[74] B. MacKichan: A generalization to overdetermined systems of the notion of diagonal operators. I. Elliptic operators *Acta Math.* **126** (1971) 83–119
[75] B. MacKichan: A generalization to overdetermined systems of the notion of diagonal operators. II. Hyperbolic operators *Acta Math.* **134(3-4)** (1975) 239–274
[76] Y. Nambu: Generalized Hamiltonian dynamics *Phys. Rev. D* **7** (1973) 2405–2412
[77] L. Nirenberg, J. F. Treves: On local solvability of linear partial differential equations. Part I. Necessary conditions *Comm. Pure Appl. Math.* **23** (1970) 1–38;
L. Nirenberg, J. F. Treves: On local solvability of linear partial differential equations. Part II. Sufficient conditions *Comm. Pure Appl. Math.* **23** (1970) 459–509; *correction ibid.* **24(2)** (1971) 279–288
[78] P. Olver: *Applications of Lie groups to differential equations* (Graduate Texts in Mathematics, 107, Springer-Verlag, New York, 1986)
[79] L. V. Ovsiannikov: *Group analysis of differential equations* (Russian: Nauka, Moscow (1978); Engl. transl.: Academic Press, New York, 1982)
[80] I. G. Petrovsky: *Lectures on partial differential equations* (7th ed., MGU (Russian), Moscow (1984); Engl. transl.: Dover, 1991)
[81] A. S. Pollack: The integrability problem for pseudogroup structures *J. Diff. Geometry* **9** (1974) 355–390
[82] J.-F. Pommaret: *Differential Galois theory* (Gordon and Breach, 1978)
[83] F. A. E. Pirani, D. C. Robinson, W. F. Shadwick: *Local jet bundle formulation of Backlund transformations. With applications to nonlinear evolution equations* (Mathematical Physics Studies 1, Reidel Publ.Co., 1979)

[84] P. K. Rashevsky: *Geometric theory of PDEs* (OGIZ (Russian) Moscow-Leningrad, 1947)
[85] C. Riquier: *Les systèmes d'équations aux derivées partielles* (Gauthier-Villars, Paris, 1910)
[86] I. M. Singer, S. Sternberg: On the infinite groups of Lie and Cartan *J. d'Analyse Math.* **15** (1965) 1–114
[87] D. Spencer: Deformation of structures on manifolds defined by transitive pseudogroups, Part I *Ann. of Math. (2)* **76** (1962) 306–398;
D. Spencer: Deformation of structures on manifolds defined by transitive pseudogroups, Part II *Ann. of Math. (2)* **76** (1962) 399–445
[88] D. C. Spencer: Overdetermined systems of linear partial differential equations *Bull. Amer. Math. Soc.* **75** (1969) 179–239
[89] S. Sternberg: *Lectures on Differential Geometry* (Prentice-Hall, Englewood Cliffs, 1964)
[90] W. J. Sweeney: The δ-Poincaré estimate *Pacific J. Math* **20** (1967) 559–570
[91] W. J. Sweeney: A uniqueness theorem for the Neumann problem *Ann. of Math. (2)* **90** (1969) 353–360
[92] A. V. Bocharov, V. N. Chetverikov, S. V. Duzhin, N. G. Khor'kova, I. S. Krasil'shchik, A. V. Samokhin, Yu. N. Torkhov, A. M. Verbovetsky, A. M. Vinogradov: *Symmetries and conservation laws for differential equations of mathematical physics* (Transl. Math. Monograph 182, A.M.S., Providence, 1999)
[93] G. R. Sell, Y. You: *Dynamics of evolutionary equations* (Appl. Math. Sci. 143, Springer, 2002)
[94] N. Tanaka: On differential systems, graded Lie algebras and pseudogroups *Jour. Math. Kyoto Univ.* **10** (1970) 1–82
[95] N. N. Tarkhanov: *The parametrix method in the theory of differential complexes* ((Russian), Nauka, Novosibirsk (1990); Engl. transl.: Complexes of differential operators, Mathematics and its Applications 340, Kluwer, 1995)
[96] T. V. Thomas: *The differential invariants of generalized spaces* (Cambridge, The University Press, 1934)
[97] J. Thomas: *Differential systems* (American Mathematical Society, 1937)
[98] A. Tresse: Sur les invariants differentiels des groupes continus de transformations *Acta Math.* **18** (1894) 1–88
[99] W. M. Tulczyjew: The Euler-Lagrange resolution. Differential geometrical methods in mathematical physics *Lecture Notes in Math.* **836** (1980) 22–48
[100] E. Vessiot: Sur une théorie nouvelle des problèmes généraux d'integration *Bull. Soc. Math. France* **52** (1924) 336–395
[101] D. Yang: *Involutive hyperbolic differential systems* (Mem. Amer. Math. Soc. 68, no. 370, 1987)

BORIS KRUGLIKOV, VALENTIN LYCHAGIN

Department Mathematics and Statistics, Faculty of Science, University of Tromsø, Tromsø 90-37, Norway

E-mail: `kruglikov@math.uit.no`, `lychagin@math.uit.no`

This page intentionally left blank

Global variational theory in fibred spaces

D. Krupka

Contents

1 Introduction

The purpose of this paper is to present foundations of the theory of global, higher order variational functionals for sections of fibred manifolds, and to review recent general developments in this field. We discuss basic concepts as well as new local and global results obtained during a few last decades.

Connected with the progress in analysis on smooth manifolds, the classical concepts like a *Lagrangian*, a *variation*, the *Euler-Lagrange expressions*, a *symmetry transformation*, etc., have been extended to manifolds. New methods based on a natural use of differential forms, disributions, vector and tensor fields, jets, and on algebraic topology, have been developed. Many new variational functionals and variational principles in physics have been investigated.

The contents and structure of the presented theory differ in many aspects from the classical approach (compare e.g. with Gianessi and Maugeri [21], and Giaquinta and Hildebrandt [22]). Main innovations consist in the use of differential and integral calculus on manifolds, needed for understanding of global structure of integral variational functionals. The corresponding concepts, in particular, the calculus of differential forms and their Lie derivatives, are recalled in Section 2 and Section 3. Further on, in Sections 4, 5, and 6 we give fundamental definitions, and discuss properties of global variational functionals; key concepts are the Lagrangian (Lagrange form), variation, and Euler-Lagrange form, whose components in a fibred chart define the Euler-Lagrange equations. We discuss general properties of the Euler-Lagrange mapping, that assigns to a Lagrangian its Euler-Lagrange form (variational triviality of Lagrangians, and variationality of source equations - the topics closely related with the inverse problem of the calculus of variations). Next we present, within the global context of fibred manifolds, the theory of Emmy Noether on invariant variational functionals. Last part of this work is devoted to some selected open questions.

Our basic references are *Goldshmidt and Sternberg* [23] (the Cartan form, vector valued Euler-Lagrange form, Hamilton theory, the Hamilton-Jacobi equation), *Krupka* [41], [50] (first variation for higher order functionals, Lepage forms, structure of contact forms and horizontalization, the Euler-Lagrange form, invariance), *García* [19] (Poincaré-Cartan form, connections and invariant variational operations, vector-valued Euler-Lagrange form), *Trautman* [92], [93] (invariant variational functionals, Noether theory), and *Dedecker* [15] (geometric concepts in the calculus of variations on grassmannians, regularity).

Later, many other authors contributed to different parts of the theory; for extensive literature on all these topics, and many others, we refer to the works of Grigore, Helein and Wood, Krupková and Prince, Saunders, and Vitolo, published in this book. We do not consider in this paper several specific questions such as variational problems for submanifolds (see e.g. Grigore [29], Grigore and Krupka [30], D. Krupka and M. Krupka [57]), properties of the variational theory over 1-dimensional base manifolds (the *higher order mechanics*), the theory of *harmonic mappings*, and *minimal submanifolds*, variational aspects of the theory of *differential equations*, variational principles in physics, and the variational bicomplex theory.

Throughout this article, the following standard notations and conventions are applied. $\mathbf{R}$ is the field of real numbers, and $\mathbf{R}^n$ is the n-dimensional Euclidean space of ordered n-tuples of real numbers. All manifolds are real and finite-dimensional, and all mappings belong to the category C^∞. We freely use the symbols D and D_i, and $\partial/\partial x^i$ for the derivative of a mapping, and partial derivatives, respectively. As usual in analysis on manifolds, TX is the *tangent bundle* of a manifold X, and Tf is the tangent mapping of a mapping f between two manifolds; i_ξ and ∂_ξ denote the *contraction*, and the *Lie derivative* of differential forms by a vector field ξ on X, respectively.

Standard concepts applied in higher order global analysis and geometry are used without special notice. Y denotes a fixed fibred manifold with base X and projection π, and $n = \dim X$, $m = \dim Y - n$. For any positive integer r, J^rY denotes the *r-jet prolongation* of Y, and $\pi^{r,s} : J^rY \to J^sY$, $\pi^r : J^rY \to X$ are the *canonical jet* projections; we set $J^0Y = 0$. If W is a set in Y, we denote $W^r = (\pi^{r,0})^{-1}(W)$. An element of the set J^rY, i.e., the *r-jet* of a section γ of Y at a point $x \in X$, is denoted $J^r_x\gamma$; the *r-jet prolongation* of γ is the mapping $x \to J^r\gamma(x) = J^r_x\gamma$. For any open set $W \subset Y$, the ring of functions on W^r is denoted by $\Omega^r_0 W$, and the $\Omega^r_0 W$-module of k-forms on W^r is denoted by $\Omega^r_k W$. $\Omega^r_{k,X} W$ ($\Omega^r_{k,Y} W$) is a submodule of $\Omega^r_k W$ formed by π^r-*horizontal* ($\pi^{r,0}$-*horizontal*) forms. $\Omega^r W$ is the exterior algebra of forms on W^r.

Every *fibred chart* on Y, usually denoted by (V, ψ), $\psi = (x^i, y^\sigma)$, where $1 \le i \le n$, $1 \le \sigma \le m$, induces the *associated charts* on J^rY, and X, denoted by (V^r, ψ^r), $\psi^r = (x^i, y^\sigma, y^\sigma_{j_1}, y^\sigma_{j_1 j_2}, \dots, y^\sigma_{j_1 j_2 \dots j_r})$, where $V^r = (\pi^{r,0})^{-1}(V)$, and (U, φ), $\varphi = (x^i)$, where $U = \pi(V)$. Sometimes we use multi-index notation for charts, and write $\psi^r = (x^i, y^\sigma_J)$.

We set

$$\omega_0 = dx^1 \wedge dx^2 \wedge \ldots \wedge dx^n,$$

$$\omega_i = i_{\partial/\partial x^i}\omega_0 = (-1)^{i-1} dx^1 \wedge dx^2 \wedge \ldots \wedge dx^{i-1} \wedge dx^{i+1} \wedge \ldots \wedge dx^n,$$

$$\omega_{i_1 i_2 \ldots i_{k-1} i_k} = i_{\partial/\partial x^{i_k}} \omega_{i_1 i_2 \ldots i_{k-1}}, \quad 1 \le k \le n.$$

δ^i_j are the components of the *Kronecker symbol*, and $\epsilon_{i_1 i_2 \ldots i_n}$ stands for the components of the *Levi-Civita* (alternating) *symbol*. We also use a specific notation for the *alternation operator* applied to a given set of indices $\{i_1, i_2, \ldots, i_p\}$; we denote this operator by Alt$(i_1 i_2 \ldots i_p)$. Then we have $\omega_{i_1 i_2 \ldots i_n} = \epsilon_{i_1 i_2 \ldots i_n}$, and

$$dx^k \wedge \omega_{i_1 i_2 \ldots i_{p+1}} = (p+1)\delta^k_{i_{p+1}} \omega_{i_1 i_2 \ldots i_p} \quad \mathrm{Alt}(i_1 i_2 \ldots i_{p+1}),$$

$$\omega_{i_1 i_2 \ldots i_k} = \frac{1}{(n-k)!} \epsilon_{i_1 i_2 \ldots i_k s_{k+1} s_{k+2} \ldots s_n} dx^{s_{k+1}} \wedge dx^{s_{k+2}} \wedge \ldots \wedge dx^{s_n}.$$

Analogously, the *symmetrization operator* in the set of indices $\{i_1, i_2, \ldots, i_p\}$ is denoted by Sym$(i_1 i_2 \ldots i_p)$.

In general, formal computations in this article, as well as the proofs, are shortened to a minimum.

2 Prolongations of fibred manifolds

The aim of this introductory section is to recall basic concepts of the theory of jet prolongations of fibred manifolds, needed for understanding of the geometric structure of the calculus of variations on fibred manifolds.

2.1 Fibred manifolds

Recall that a *section* of the fibred manifold Y is a mapping $\gamma : W \to Y$, where $W \subset X$ is an open set, such that $\pi \circ \gamma = \mathrm{id}_W$. γ is a section if and only if for every point $x_0 \in W$ there exists a fibred chart (V, ψ), $\psi = (u^i, y^\sigma)$, at $\gamma(x_0) \in W$ such that the associated chart (U, φ), $\varphi = (x^i)$, satisfies $\gamma(U) \subset W$. In these charts, γ has equations of the form $u^i \circ \gamma = x^i$, $y^\sigma \circ \gamma = f^\sigma(x^i)$.

Let Y_1 (Y_2) be a fibred manifold with base X_1 (X_2) and projection π_1 (π_2). A mapping $\alpha : Y_1 \to Y_2$ is called a *morphism* of Y_1 into Y_2, if there exists a mapping $\alpha_0 : X_1 \to X_2$ such that $\pi_2 \circ \alpha = \alpha_0 \circ \pi_1$. If α_0 exists it is unique, and is called the *projection* of the morphism α. A morphism of fibred manifolds $\alpha : Y_1 \to Y_2$, which is a diffeomorphism, is called an *isomorphism*. The projection of an isomorphism of fibred manifolds is a diffeomorphism of their bases.

Let Y be a fibred manifold with base X and projection π. Let V be an open set in Y, considered as a fibred manifold over the open set $\pi(V)$ in X. A morphism of fibred manifolds $\alpha : V \to Y$ is called a *local automorphism* of Y, if $\alpha(V)$ is an open set in Y and α is an isomorphism of V and $\alpha(V)$ over the projection of α.

A morphism of fibred manifolds $\alpha : Y_1 \to Y_2$ is expressed in two fibred charts (V_1, ψ_1), $\psi_1 = (x^i_1, y^\sigma_1)$, and (V_2, ψ_2), $\psi_2 = (x^i_2, y^\sigma_2)$, on Y_1 and Y_2, respectively, by equations of the form

$$x_2^i = f^i(x_1^j), \quad y_2^\sigma = g^\sigma(x_1^j, y_1^\nu). \tag{1}$$

A tangent vector $\xi \in TY$ is said to be π-*vertical*, or simply *vertical*, if $T\pi \cdot \xi = 0$. A differential p-form ρ on Y is said to be π-*horizontal*, or simply *horizontal*, if for each point $y \in Y$, the contraction $i_\xi \rho(y)$ of $\rho(y)$ vanishes whenever ξ is a π-vertical vector from the tangent space $T_y Y$. A *vector field* ξ on Y is said to be π-*projectable*, or simply *projectable*, if there exists a vector field ζ on X such that

$$T\pi \cdot \xi = \zeta \circ \pi. \tag{2}$$

If ζ exists, it is unique, and is called the π-*projection* of ξ.

If α_t is the local one-parameter group of ξ, then it is easily seen that ξ is π-projectable if and only if each point $y \in Y$ has a neighborhood V such that α_t is defined on V for all sufficiently small t and is an isomorphism of Y.

In a fibred chart (V, ψ), $\psi = (x^i, y^\sigma)$, a π-projectable vector field ξ is expressed by

$$\xi = \xi^i \frac{\partial}{\partial x^i} + \Xi^\sigma \frac{\partial}{\partial y^\sigma}, \tag{3}$$

where $\xi^i = \xi^i(x^j)$, $\Xi^\sigma = \Xi^\sigma(x^j, y^\nu)$.

Let $U \in \mathbf{R}^n$ be an open set, let $W \in \mathbf{R}^m$ be an open ball with centre at the origin, and let $\zeta : U \to U \times W$ be the zero section. We define a mapping χ from the set $[0,1] \times U \times W$ to $U \times W$ by

$$\chi(s, (x^i, y^\sigma)) = (x^i, sy^\sigma). \tag{4}$$

Then

$$\chi^* dx^i = dx^i, \quad \chi^* dy^\sigma = y^\sigma ds + s dy^\sigma. \tag{5}$$

For any k-form ρ on $U \times W$, where $k \geq 1$, consider the pull-back $\chi^*\rho$, which is a k-form on $[0,1] \times U \times W$. Obviously, there exists a unique decomposition

$$\chi^*\rho = ds \wedge \rho^{(0)}(s) + \rho'(s) \tag{6}$$

such that the k-forms $\rho^{(0)}(s)$ and $\rho'(s)$ do not contain ds. Note that by (5), $\rho'(s)$ arises from ρ by replacing each factor dy^σ by sdy^σ, and by replacing each coefficient f in ρ by $f \circ \chi$; the factors dx^i remain unchanged. Thus, $\rho'(s)$ obeys

$$\rho'(1) = \rho, \quad \rho'(0) = \pi^*\zeta^*\rho. \tag{7}$$

Define

$$I\rho = \int \rho^{(0)}(s), \tag{8}$$

where the integral on the right-hand side means integration of the coefficients in the form $\rho^{(0)}(s)$ over s from 0 to 1. If $f : U \times W \to \mathbf{R}$ is a function, we define

$$If = 0. \tag{9}$$

The mapping $\rho \to I\rho$ is called the *fibred homotopy operator*.

The proofs of the following two results are standard.

Lemma 1 *For any differential k-form ρ on $U \times W$,*

$$\rho = Id\rho + dI\rho + \pi^* \zeta^* \rho. \tag{10}$$

Theorem 1 (The Volterra-Poincaré lemma) *Let $U \subset \boldsymbol{R}^n$ be an open set, $V \subset \boldsymbol{R}^m$ an open ball with centre 0, ρ a differential form on $U \times V$. If $d\rho = 0$, then there exists a form η on $U \times V$ such that*

$$\rho = d\eta + \pi^* \zeta^* \rho. \tag{11}$$

2.2 Prolongations

Let Y be a fibred manifold with base X and projection π, and let $n = \dim X$, $m = \dim Y - n$. We denote by $J^r Y$, where $r \geq 0$, the *r-jet prolongation* of Y, and by $\pi^{r,s} : J^r Y \to J^s Y$, $\pi^r : J^r Y \to X$ the *canonical jet projections*. An element of the set $J^r Y$, i.e., the *r-jet* of a section γ of Y at a point $x \in X$, is denoted $J^r_x \gamma$; the *r-jet prolongation* of γ is the mapping $x \to J^r \gamma(x) = J^r_x \gamma$. Recall that any *fibred chart* (V, ψ), $\psi = (x^i, y^\sigma)$, on Y, where $1 \leq i \leq n$, $1 \leq \sigma \leq m$, induces the *associated charts* (V^r, ψ^r), $\psi^r = (x^i, y^\sigma, y^\sigma_{j_1}, y^\sigma_{j_1 j_2}, \dots, y^\sigma_{j_1 j_2 \dots j_r})$, on $J^r Y$, and (U, φ), $\varphi = (x^i)$, on X; here $V^r = (\pi^{r,0})^{-1}(V)$, and $U = \pi(V)$. If $W \subset Y$ is an open set, we denote $W^r = (\pi^{r,0})^{-1}(W)$.

Clearly, the concepts of *verticality* of vectors and *horizontality* of forms apply to the canonical jet projections of $J^r Y$. For any open set $W \subset Y$, the ring of functions on W^r is denoted $\Omega^r_0 W$, and the $\Omega^r_0 W$-module of k-forms on W^r is denoted $\Omega^r_k W$. $\Omega^r_{k,X} W$ ($\Omega^r_{k,Y} W$) is a submodule of $\Omega^r_k W$ formed by π^r-*horizontal* ($\pi^{r,0}$-*horizontal*) forms. $\Omega^r W$ is the exterior algebra of forms on W^r.

We introduce an exterior algebra morphism related to the structure of the jet prolongations of a fibred manifold. Let ρ be a differential k-form on W^r, $0 \leq k \leq n$. There exists one and only one π^{r+1}-horizontal k-form $h\rho$ on W^{r+1} such that $J^r \gamma^* \rho = J^{r+1} \gamma^* h\rho$ for all sections γ of W. The existence of $h\rho$ follows from the definition of the pull-back of forms. If $k = 0$ and f is a function, then at any point $J^{r+1}_x \gamma \in (\pi^{r+1,r})^{-1}(W)$,

$$hf(J^{r+1}_x \gamma) = f(J^r_x \gamma). \tag{1}$$

If $k \geq 1$, then for any $J^{r+1}_x \gamma \in (\pi^{r+1,r})^{-1}(W)$, and any tangent vectors $\xi_1, \xi_2, \dots, \xi_k \in TJ^r Y$ at $J^{r+1}_x \gamma$,

$$\begin{aligned} &h\rho(J^{r+1}_x \gamma)(\xi_1, \xi_2, \dots, \xi_k) \\ &\quad = \rho(J^r_x \gamma)(T_x J^r \gamma \cdot T\pi^{r+1} \cdot \xi_1, T_x J^r \gamma \cdot T\pi^{r+1} \cdot \xi_2, \dots, T_x J^r \gamma \cdot T\pi^{r+1} \cdot \xi_k). \end{aligned} \tag{2}$$

If $k = 0$, uniqueness is evident. If $1 \leq k \leq n$, we use a fibred chart (V, ψ), $\psi = (x^i, y^\sigma)$, and express $h\rho$ as $\rho_{i_1 i_2 \dots i_k} dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_k}$; then the condition $J^{r+1} \gamma^* h\rho = 0$ for all γ implies $\rho_{i_1 i_2 \dots i_k} = 0$. If $k > 0$, then $h\rho = 0$ identically. Thus, since $h\rho$ always exists and is unique, (1) and (2) may serve as definitions of $h\rho$.

The mapping $\Omega^r_k W \ni \rho \to h\rho \in \Omega^{r+1}_k W$ is called the π-*horizontalization*, or simply the *horizontalization*. The form $h\rho$ is the *horizontal component* of ρ.

We wish to show that the horizontalization can be considered as a *morphism of exterior algebras*, induced locally by horizontalizations of functions and their exterior derivatives.

Theorem 2 *Let W be an open set in the fibred manifold Y. Then the horizontalization $\Omega^r W \ni \rho \to h\rho \in \Omega^{r+1} W$ is a unique* **R***-linear, exterior-product-preserving mapping such that for any function $f : W^r \to \mathbf{R}$, and any fibred chart (V, ψ), $\psi = (x^i, y^\sigma)$, in W,*

$$hf = f \circ \pi^{r+1,r}, \quad h(df) = d_i f \cdot dx^i, \tag{3}$$

where

$$d_i f = \frac{\partial f}{\partial x^i} + \sum_{k=0}^{r} \sum_{j_1 \le j_2 \le \ldots \le j_k} \frac{\partial f}{\partial y^\sigma_{j_1 j_2 \ldots j_k}} y^\sigma_{j_1 j_2 \ldots j_k i}. \tag{4}$$

The proof that h (1), (2) has the desired properties is standard. To prove uniqueness, note that (3) and (4) imply $hdx^i = dx^i$ and $hdy^\sigma_{j_1 j_2 \ldots j_k} = y^\sigma_{j_1 j_2 \ldots j_k i} dx^i$. Now it is easy to check that any two mappings h_1, h_2, satisfying the assumptions of Theorem 2, which agree on functions and their exterior derivative, coincide.

The function $d_i f : V^{r+1} \to \mathbf{R}$ is called the *i-th formal derivative* of f with respect to the fibred chart (V, ψ). Note that formal derivatives (4) are components of an invariant object, the *horizontal component* of the exterior derivative of a function.

In the following lemma, we list elementary properties of formal derivatives; the formal derivative operator with respect to a fibred chart $(\bar{V}, \bar{\psi})$, $\bar{\psi} = (\bar{x}^i, \bar{y}^\sigma)$, is denoted by $\bar{d}_i$.

Lemma 2 *Let (V, ψ), $\psi = (x^i, y^\sigma)$, be a fibred chart on Y.*

(a) *The coordinate functions $y^\nu_{j_1 j_2 \ldots j_k}$ satisfy*

$$d_i y^\nu_{j_1 j_2 \ldots j_k} = y^\nu_{j_1 j_2 \ldots j_k i}. \tag{5}$$

(b) *If $(\bar{V}, \bar{\psi})$, $\bar{\psi} = (\bar{x}^i, \bar{y}^\sigma)$, is another chart on Y such that $V \cap \bar{V} \neq \emptyset$, then for every function $f : V^r \cap \bar{V}^r \to \mathbf{R}$,*

$$\bar{d}_i f = d_i f \frac{\partial x^j}{\partial \bar{x}^i}. \tag{6}$$

(c) *For any two functions $f, \; g : V^r \to \mathbf{R}$,*

$$d_i (f \cdot g) = g \cdot d_i f + f \cdot d_i g. \tag{7}$$

(d) *For every function $f : V^r \to \mathbf{R}$, and every section $\gamma : U \to Y$,*

$$d_i f \circ J^{r+1}\gamma = \frac{\partial (f \circ J^r \gamma)}{\partial \bar{x}^i}. \tag{8}$$

Remark 1 By (5), $\bar{y}^\sigma_{j_1 j_2 \ldots j_k} = \bar{d}_{j_k} \bar{y}^\sigma_{j_1 j_2 \ldots j_{k-1}}$. Thus, applying (6) to coordinates, we obtain the following *prolongation formula* for coordinate transformations in prolongations of fibred manifolds

$$\bar{y}^\sigma_{j_1 j_2 \ldots j_k} = d_i \bar{y}^\sigma_{j_1 j_2 \ldots j_{k-1}} \frac{\partial x^i}{\partial \bar{x}^{j_k}}. \tag{9}$$

We introduce a vector bundle morphism acting on tangent spaces to the jet prolongations of a fibred manifold. Let $r \geq 0$ be an integer. One can assign to every tangent vector $\xi \in TJ^{r+1}Y$ at a point $J^{r+1}_x\gamma \in J^{r+1}Y$ a tangent vector $h\xi \in TJ^rY$ at $J^r_x\gamma = \pi^{r+1,r}(J^{r+1}_x\gamma) \in J^rY$ by

$$h\xi = T_xJ^r\gamma \circ T\pi^{r+1} \cdot \xi. \tag{10}$$

The mapping $h : TJ^{r+1}Y \to TJ^rY$ defined by this formula is a vector bundle morphism over the jet projection $\pi^{r+1,r}$; we call h the π-*horizontalization*, or simply the *horizontalization*.

It follows from the definition that the tangent vector $h\xi$ is π^{r+1}-horizontal; we sometimes call $h\xi$ the *horizontal component* of ξ. A tangent vector ξ is a π^{r+1}-vertical vector if and only if $h\xi = 0$.

Using a complementary construction, one can assign to every tangent vector $\xi \in TJ^{r+1}Y$ at a point $J^{r+1}_x\gamma \in J^{r+1}Y$ a tangent vector $p\xi \in TJ^rY$ at $J^r_x\gamma$ by

$$p\xi = T\pi^{r+1,r} \cdot \xi - h\xi. \tag{11}$$

We call $p\xi$ the *contact component* of ξ. It is immediate that $p\xi$ is a π^r-vertical vector, and ξ is $\pi^{r+1,r}$-vertical if and only if $h\xi = 0$, $p\xi = 0$.

Let $\xi \in TJ^{r+1}Y$ be a tangent vector at a point $J^{r+1}_x\gamma \in J^{r+1}Y$, and let (V, ψ), $\psi = (x^i, y^\sigma)$, be a fibred chart at the point $y = \gamma(x) \in V$. If ξ has an expression

$$\xi = \xi^i \frac{\partial}{\partial x^i} + \sum_{k=0}^{r+1} \sum_{j_1 \leq j_2 \leq \ldots \leq j_k} \Xi^\sigma_{j_1 j_2 \ldots j_k} \frac{\partial}{\partial y^\sigma_{j_1 j_2 \ldots j_k}}, \tag{12}$$

then

$$h\xi = \xi^i \left(\frac{\partial}{\partial x^i} + \sum_{k=0}^{r} \sum_{j_1 \leq j_2 \leq \ldots \leq j_k} y^\sigma_{j_1 j_2 \ldots j_k i} \frac{\partial}{\partial y^\sigma_{j_1 j_2 \ldots j_k}} \right), \tag{13}$$

and by definition,

$$p\xi = \sum_{k=0}^{r} \sum_{j_1 \leq j_2 \leq \ldots \leq j_k} \left(\Xi^\sigma_{j_1 j_2 \ldots j_k} - y^\sigma_{j_1 j_2 \ldots j_k i}\xi^i \right) \frac{\partial}{\partial y^\sigma_{j_1 j_2 \ldots j_k}}. \tag{14}$$

Let $\alpha : V \to Y$ be a local automorphism of Y, $U = \pi(V)$, and $\alpha_0 : U \to X$ the projection of α. Let $V^r = (\pi^{r,0})^{-1}(V)$. We define a local automorphism $J^r\alpha : V^r \to J^rY$ of J^rY by

$$J^r\alpha(J^r_x\gamma) = J^r_{\alpha_0(x)}(\alpha\gamma\alpha_0^{-1}). \tag{15}$$

$J^r\alpha$ is called the *r-jet prolongation*, or just an *r-prolongation* of α. Note that for every section γ defined on an open subset of U, with values in V, (15) implies

$$J^r\alpha \circ J^r\gamma \circ \alpha_0^{-1} = J^r(\alpha\gamma\alpha_0^{-1}), \tag{16}$$

and *vice versa*. In particular, this formula shows that the r-jet prolongations of local automorphisms carry sections of Y into sections of J^rY.

Lemma 3 (a) *For any s,* $0 \le s \le r$,

$$\pi^r \circ J^r\alpha = \alpha_0 \circ \pi^r, \quad \pi^{r,s} \circ J^r\alpha = J^s\alpha \circ \pi^{r,s}. \tag{17}$$

(b) *If two isomorphisms* α *and* β *of* Y *are composable, then*

$$J^r\alpha \circ J^r\beta = J^r(\alpha \circ \beta). \tag{18}$$

(c) *For any isomorphism* α *of* Y, *and any differential form* ρ *on* J^rY,

$$J^{r+1}\alpha^* h\rho = hJ^r\alpha^*\rho. \tag{19}$$

All these assertions are easy consequences of definitions.

Remark 2 We describe the r-jet prolongation of a isomorphism in terms of fibred charts. Suppose that in two fibred charts on Y, (V,ψ), $\psi = (x^i, y^\sigma)$, and $(\bar{V},\bar{\psi})$, $\bar{\psi} = (\bar{x}^i, \bar{y}^\sigma)$, $\alpha : V \to Y$ is expressed by equations $\bar{x}^i = f^i(x^j)$, $\bar{y}^\sigma = g^\sigma(x^i, y^\nu)$. Since for every $J^r_x\gamma \in V^r$, the transformed point $J^r\alpha(J^r_x\gamma)$ has the coordinates

$$\begin{aligned} &\bar{x}^i \circ J^r\alpha(J^r_x\gamma) = \bar{x}^i \circ \alpha_0(x),\\ &\bar{y}^\sigma \circ J^r\alpha(J^r_x\gamma) = \bar{y}^\sigma \circ \alpha(\gamma(x)),\\ &\bar{y}^\sigma_{j_1 j_2 \dots j_k} \circ J^r\alpha(J^r_x\gamma) = D_{j_1}D_{j_2}\dots D_{j_k}(\bar{y}^\sigma\alpha\gamma\alpha_0^{-1})(\alpha_0(x)), \end{aligned} \tag{20}$$

we easily obtain a recurrent formula. If, for example, $k = 2$, we get

$$\begin{aligned} &\bar{x}^i = f^i(x^j),\\ &\bar{y}^\sigma = g^\sigma(x^i, y^\nu),\\ &\bar{y}^\sigma_{j_1} = d_{k_1}g^\sigma(x^i, y^\nu) \cdot \frac{\partial x^{k_1}}{\partial \bar{x}^{j_1}},\\ &\bar{y}^\sigma_{j_1 j_2} = d_{k_1}d_{k_2}g^\sigma(x^i, y^\nu) \cdot \frac{\partial x^{k_1}}{\partial \bar{x}^{j_1}}\frac{\partial x^{k_2}}{\partial \bar{x}^{j_2}} + d_{k_1}g^\sigma(x^i, y^\nu) \cdot \frac{\partial^2 x^{k_1}}{\partial \bar{x}^{j_1}\partial \bar{x}^{j_2}}. \end{aligned} \tag{21}$$

If the transformation equations for $l < k$ are already given in the form

$$\bar{y}^\sigma_{j_1 j_2 \dots j_l} = g^\sigma_{j_1 j_2 \dots j_l}(x^i, y^\nu, y^\nu_{i_1}, y^\nu_{i_1 i_2}, \dots, y^\nu_{i_1 i_2 \dots i_l}), \tag{22}$$

then the chart expression of $J^r\alpha$ is defined by (21), (22), and

$$\bar{y}^\sigma_{j_1 j_2 \dots j_k} = d_i g^\sigma_{j_1 j_2 \dots j_{k-1}}(x^i, y^\nu, y^\nu_{i_1}, y^\nu_{i_1 i_2}, \dots, y^\nu_{i_1 i_2 \dots i_{k-1}}) \cdot \frac{\partial x^i}{\partial \bar{x}^{j_k}}, \quad 1 \le k \le r. \tag{23}$$

Note that these transformation formulae are polynomial in $y^\nu_{i_1}$, $y^\nu_{i_1 i_2}$, …, $y^\nu_{i_1 i_2 \dots i_r}$.

Let Ξ be π-projectable vector field on Y, ξ its π-projection, α_t the local one-parameter group of Ξ, and $J^r\alpha_t$ the r-jet prolongation of α_t. We define for each point $J^r_x\gamma$ belonging to the domain of $J^r\alpha_t$

$$J^r\Xi(J^r_x\gamma) = \left(\frac{d}{dt}J^r\alpha_t(J^r_x\gamma)\right)_0. \tag{24}$$

Then $J^r\Xi$ is a vector field on J^rY, called *r-jet prolongation* of Ξ. It follows from the definition that $J^r\Xi$ is π^r-projectable ($\pi^{r,s}$-projectable for any s, $0 \leq s \leq r$) and its π^r-projection ($\pi^{r,s}$-projection) is ξ ($J^s\Xi$).

The following lemma describes the local structure of the jet prolongations of projectable vector fields; its proof is based on the use of the chain rule.

Lemma 4 *Let Ξ be a π-projectable vector field on Y, (V, ψ), $\psi = (x^i, y^\sigma)$, a fibred chart on Y, and let Ξ be expressed by*

$$\Xi = \xi^i \frac{\partial}{\partial x^i} + \Xi^\sigma \frac{\partial}{\partial y^\sigma}. \tag{25}$$

Then $J^r\Xi$ is expressed with respect to the associated chart (V^r, ψ^r) by

$$J^r\Xi = \xi^i \frac{\partial}{\partial x^i} + \sum_{k=0}^{r} \sum_{j_1 \leq j_2 \leq \ldots \leq j_k} \Xi^\sigma_{j_1 j_2 \ldots j_k} \frac{\partial}{\partial y^\sigma_{j_1 j_2 \ldots j_k}}, \tag{26}$$

where the components $\Xi^\sigma_{j_1 j_2 \ldots j_k}$ are determined by the recurrent formula

$$\Xi^\sigma_{j_1 j_2 \ldots j_k} = d_{j_k} \Xi^\sigma_{j_1 j_2 \ldots j_{k-1}} - y^\sigma_{j_1 j_2 \ldots j_{k-1} i} \frac{\partial \xi^i}{\partial x^{j_k}}. \tag{27}$$

In the following lemma we discuss the *Lie bracket* operation on r-jet prolongations of projectable vector fields, and the *Lie derivatives* of forms by these vector fields.

Lemma 5 (a) *Let Ξ_1 and Ξ_2 be two π-projectable vector fields. Then the Lie bracket $[\Xi_1, \Xi_2]$ is also π-projectable, and*

$$J^r[\Xi_1, \Xi_2] = [J^r\Xi_1, J^r\Xi_2]. \tag{28}$$

(b) *For any π-projectable vector field Ξ, and any differential form ρ on J^rY,*

$$\partial_{J^{r+1}\Xi} h\rho = h\partial_{J^r\Xi}\rho. \tag{29}$$

To prove (a), we first prove (28) for $r = 1$ in fibred charts. Then we assume that (28)holds for some r. Using the fibred manifold of *non-holonomic* jets J^1J^rY, i.e., the 1-jet prolongation of $\pi^r : J^rY \to X$, we obtain $J^1[J^r\Xi_1, J^r\Xi_2] = [J^1J^r\Xi_1, J^1J^r\Xi_2]$. Finally, we prove that $J^1J^{r-1}\Xi_1 \circ \iota = T\iota \cdot J^r\Xi_1$, $J^1J^{r-1}\Xi_2 \circ \iota = T\iota \cdot J^r\Xi_2$, and $T\iota \cdot (J^r[\Xi_1, \Xi_2] - [J^r\Xi_1, J^r\Xi_2]) = 0$, where ι is the *canonical injection* $J^rY \ni J^r_x\gamma \to J^1_xJ^{r-1}\gamma \in J^1J^{r-1}Y$. (b) follows from Lemma 3.

Now we consider restrictions of jet prolongations of projectable vector fields to jet prolongations of section.

Lemma 6 *Let Ξ_1 and Ξ_2 be two π-projectable vector fields, and suppose that $\Xi_1 \circ \gamma = \Xi_2 \circ \gamma$ for a section γ of Y. Then*

$$J^r\Xi_1 \circ J^r\gamma = J^r\Xi_2 \circ J^r\gamma. \tag{30}$$

To prove Lemma 6, we use the notation of Lemma 4, and write Ξ_1, Ξ_2 in a fibred chart; we get for their components, $\xi^i_1 = \xi^i_2$, $\Xi^\sigma_1 \circ \gamma = \Xi^\sigma_2 \circ \gamma$. Then from Lemma 2, (d),

$$d_i\Xi^\sigma_1(J^1_x\gamma) = d_i\Xi^\sigma_1 \circ J^1\gamma(x) = \frac{\partial(\Xi^\sigma_1 \circ \gamma)}{\partial x^i} = \frac{\partial(\Xi^\sigma_2 \circ \gamma)}{\partial x^i} = d_i\Xi^\sigma_2(J^1_x\gamma). \tag{31}$$

Therefore, by Lemma 4, $J^1\Xi_1 = J^1\Xi_2$. Now we proceed by induction, using Lemma 2 and Lemma 4 again.

Remark 3 Some authors call the forms ω^σ_J the *Cartan forms* (usually in the context of vector bundles); diffeomorphisms, preserving the Cartan forms are called *Lie transformations*, and the vector fields whose local one-parameter groups consist of Lie transformations, are called *Lie fields* (cf. Krasilschik [38]).

3 Differential forms on prolongations of fibred manifolds

Our principal aim in this section is to develop a canonical decomposition theory of differential forms on jet prolongations of fibred manifolds; the tools inducing the decompositions are the canonical jet projections. The proofs are based on the trace decomposition theory (Krupka [52]).

3.1 The first canonical decomposition

Consider the horizontalization $h : TJ^{r+1}Y \to TJ^rY$, introduced in Section 2. h induces a decomposition of each of the modules of q-forms $\Omega^r_q W$, where $q \geq 1$, as follows. Let $\rho \in \Omega^r_q W$ be a form, and let $\xi_1, \xi_2, \dots, \xi_q$ be tangent vectors to $J^{r+1}Y$ at a point $J^{r+1}_x\gamma \in W^{r+1}$. We write for each $i = 1, 2, \dots, q$,

$$T\pi^{r+1,r} \cdot \xi_i = h\xi_i + p\xi_i, \tag{1}$$

and substitute these vectors in the pull-back $(\pi^{r+1,r})^*\rho$ of ρ. Since

$$\begin{aligned} &(\pi^{r+1,r})^*\rho(J^{r+1}_x\gamma)(\xi_1, \xi_2, \dots, \xi_q) \\ &= \rho(J^r_x\gamma)(T\pi^{r+1,r} \cdot \xi_1, T\pi^{r+1,r} \cdot \xi_2, \dots, T\pi^{r+1,r} \cdot \xi_q), \end{aligned} \tag{2}$$

collecting together all terms homogeneous of degree $q - k$ in horizontal components $h\xi_1$, $h\xi_2$, …, $h\xi_q$ of the vectors $\xi_1, \xi_2, \dots, \xi_q$ where $k = 0, 1, 2, \dots, q$, we obtain a q-form $p_k\rho$ on W^{r+1}, defined by

$$\begin{aligned} p_k\rho(J^{r+1}_x\gamma)(\xi_1, \xi_2, \dots, \xi_q) = &\sum_{j_1<j_2<\dots<j_k} \; \sum_{j_{k+1}<j_{k+2}<\dots<j_q} \epsilon^{j_1 j_2 \dots j_k j_{k+1} j_{k+2} \dots j_q} \\ &\cdot\rho(J^r_x\gamma)(p\xi_{j_1}, p\xi_{j_2}, \dots, p\xi_{j_k}, h\xi_{j_{k+1}}, \dots, h\xi_{j_q}), \end{aligned} \tag{3}$$

or equivalently, by

$$\begin{aligned} &p_k\rho(J^{r+1}_x\gamma)(\xi_1, \xi_2, \dots, \xi_q) = \frac{1}{k!(q-k)!}\epsilon^{j_1 j_2 \dots j_k j_{k+1} j_{k+2} \dots j_q} \\ &\cdot\rho(J^r_x\gamma)(p\xi_{j_1}, p\xi_{j_2}, \dots, p\xi_{j_k}, h\xi_{j_{k+1}}, \dots, h\xi_{j_q}) \end{aligned} \tag{4}$$

(summation through *all* values of the indices $j_1, j_2, \dots, j_k, j_{k+1}, j_{k+2}, \dots, j_q$). The form $p_k\rho$ is called the *k-contact component* of the form ρ.

If $(\pi^{r+1,r})^*\rho = p_k\rho$ or, which is the same, if $p_j\rho = 0$ for all $j \neq k$; then the integer k is called the *degree of contactness* of the form ρ. The degree of contactness of the q-form $\rho = 0$ is equal to k for every $k = 0, 1, 2, \dots, q$. We say that ρ is of *degree of contactness* $\geq k$, if $h\rho = 0$, $p_1\rho = 0$, …,$p_{k-1}\rho = 0$.

It is convenient to write

$$h\rho = p_0\rho, \quad p\rho = \sum_{i=1}^{q} p_i\rho, \tag{5}$$

and extend the definition of h and p to functions. If $f : W^r \to \mathbf{R}$ is a function, we define

$$hf = (\pi^{r+1,r})^* f, \quad pf = 0. \tag{6}$$

Now if $q \geq 0$ and $\rho \in \Omega_q^r W$ is a q-form, we call the form $h\rho$ ($p\rho$) the *horizontal* (*contact*) *component* of ρ. Clearly, $(\pi^{r+1,r})^*\rho = h\rho + p\rho$, i.e.,

$$(\pi^{r+1,r})^*\rho = h\rho + p\rho = \sum_{i=0}^{q} p_i\rho. \tag{7}$$

This formula is referred to as the *first canonical decomposition* of the form ρ (note however, the decomposition concerns rather $(\pi^{r+1,r})^*\rho$ than ρ itself).

If $q \geq 1$, the form $h\rho$ is π^{r+1}-horizontal. Moreover, $h\rho$ coincides with the *horizontal component* of ρ; the mapping $\Omega_q^r W \ni \rho \to h\rho \in \Omega_q^{r+1} W$ coincides with the π-*horizontalization*.

For every positive integer k, the mapping $p_k : \Omega^r W \to \Omega^{r+1} W$ satisfies $p_k(\rho, \eta) = p_k\rho + p_k\eta$, $p_k(f\eta) = (f \circ \pi^{r+1,r}) \cdot p_k\eta$ for every ρ, η, and f; p_k is *not* a homomorphism of exterior algebras.

3.2 Contact forms

For every non-negative integer q, we define in this subsection a submodule of *contact* q-forms in the module $\Omega_q^r W$. We first introduce contact q-forms for $q = 0, 1, 2, \ldots, n$, and then extend the definition to arbitary integers $q \geq n + 1$.

Let $q \leq n$. Then we say that a form ρ, defined on W^r, is *contact*, if $h\rho = 0$. From this definition it follows that a function f is contact if and only if $f = 0$.

Remark 1 If $q > n$, then

$$h\rho = 0, \quad p_1\rho = 0, \quad p_2\rho = 0, \quad \ldots, \quad p_{q-n-1}\rho = 0 \tag{1}$$

identically.

Lemma 1 *Let W be an open set in Y, and let $\rho \in \Omega_q^r W$ be a form such that $1 \leq q \leq n$.*

(a) *ρ is contact if and only if*

$$J^r\gamma^*\rho = 0 \tag{2}$$

for every differentiable section γ of Y defined on an open subset of $\pi(W)$.

(b) *ρ is π^r-horizontal if and only if*

$$p\rho = 0. \tag{3}$$

We now discuss examples of contact 1-forms. Let (V, ψ), $\psi = (x^i, y^\sigma)$, be a fibred chart on Y. For every k, $0 \leq k \leq r - 1$, we define differential 1-forms on $V^r \subset J^r Y$ by

$$\omega^\sigma_{j_1 j_2 \ldots j_k} = dy^\sigma_{j_1 j_2 \ldots j_k} - y^\sigma_{j_1 j_2 \ldots j_k j} dx^j. \tag{4}$$

We can also use a multi-index notation, and write $\omega^\sigma_J = dy^\sigma_J - y^\sigma_{Jj}dx^j$ for the form (4), with $J = (j_1 j_2 \dots j_k)$ and $Jj = (j_1 j_2 \dots j_k j)$. It immediately follows from Lemma 1, (a) that the 1-forms (4) are contact.

In the following theorem we summarize basic properties of the forms (4). We define a section δ of $J^r Y$ to be *holonomic*, if there exists a section γ such that $\delta = J^1\gamma$.

Theorem 1 (a) *If (V, ψ), $\psi = (x^i, y^\sigma)$, is a fibred chart on Y, then the forms*

$$dx^i, \quad \omega^\sigma_J, \quad dy^\sigma_I, \tag{5}$$

where $0 \le |J| \le r-1$, and $|I| = r$, define a basis of linear forms on the set V^r.

(b) *If (V, ψ), $\psi = (x^i, y^\sigma)$, and $(\bar V, \bar\psi)$, $\bar\psi = (\bar x^i, \bar y^\sigma)$, are two fibred charts on Y such that $V \cap \bar V \neq \emptyset$, then*

$$\bar\omega^\sigma_J = \sum_{|I| \le |J|} \frac{\partial \bar y^\sigma_J}{\partial y^\nu_I} \omega^\nu_I. \tag{6}$$

(c) *A section $\delta : U \to J^1 Y$ is holonomic if and only if for any fibred chart (V, ψ), $\psi = (x^i, y^\sigma)$, such that $\pi(V) \subset U$,*

$$\delta^* \omega^\sigma_J = 0 \tag{7}$$

for all σ and J such that $1 \le \sigma \le m$ and $0 \le |J| \le r-1$.

(d) *If $q \le n-1$, the forms $d\omega^\sigma_J$ are contact.*

Formula (6) can be obtained by a direct calculation; we have

$$\bar\omega^\sigma_J = \frac{\partial \bar y^\sigma_J}{\partial x^j} dx^j + \sum_{|I| \le |J|} \frac{\partial \bar y^\sigma_J}{\partial y^\nu_I} dy^\nu_I - \bar d_j \bar y^\sigma_J \frac{\partial \bar x^j}{\partial x^k} dx^k = \sum_{|I| \le |J|} \frac{\partial \bar y^\sigma_J}{\partial y^\nu_I} \omega^\nu_I. \tag{8}$$

Condition (c) can be used as a *holonomization condition*.

The following observations are immediate consequences of Theorem 1.

Corollary 1 *The forms $d\omega^\sigma_I$ obey the transformation laws*

$$d\bar\omega^\sigma_J = \sum_{|I| \le |J|} \left(d\frac{\partial \bar y^\sigma_J}{\partial y^\nu_I} \wedge \omega^\nu_I + \frac{\partial \bar y^\sigma_J}{\partial y^\nu_I} \wedge d\omega^\nu_I \right). \tag{9}$$

Corollary 2 *Let q be an integer, $p = q, q-1, q-2, \dots, \max(q-n, 0)$, let k and l be any non-negative integers such that $k + l = p$, and $2k + l \le q$. Then the q=forms in the module $\Omega^r_q W$, locally generated by the forms*

$$\begin{aligned} &\omega^{\sigma_1}_{J_1} \wedge \omega^{\sigma_2}_{J_2} \wedge \dots \wedge \omega^{\sigma_l}_{J_l} \wedge \omega^{\nu_1}_{L_1} \wedge \omega^{\nu_2}_{L_2} \wedge \dots \wedge \omega^{\nu_{k-2}}_{L_{k-2}} \\ &\quad \wedge d\omega^{\nu_{k-1}}_{L_{k-1}} \wedge d\omega^{\nu_k}_{L_k} \wedge dx^{i_{k+l+3}} \wedge dx^{i_{k+l+4}} \dots \wedge dx^{i_q} \end{aligned} \tag{10}$$

constitute a submodule of the module $\Omega^r_q W$.

We list some computation rules for the contact forms (4).

Lemma 2 *The following formulas hold*

$$d(\omega^\sigma_{j_1 j_2 \dots j_k} \wedge \omega_{i_1 i_2}) = \omega^\sigma_{j_1 j_2 \dots j_k i_1} \wedge \omega_{i_2} - \omega^\sigma_{j_1 j_2 \dots j_k i_2} \wedge \omega_{i_1}, \tag{11}$$

$$\omega^{\sigma}_{j_1 j_2 \dots j_k} \wedge \omega_i = \omega^{\sigma}_{j_1 j_2 \dots j_k} \wedge \omega_i \quad \mathrm{Sym}(j_1 j_2 \dots j_k i)$$
$$+ \frac{1}{k+1} d(\omega^{\sigma}_{j_2 j_3 \dots j_k} \wedge \omega_{j_1 i} + \omega^{\sigma}_{j_1 j_3 \dots j_k} \wedge \omega_{j_2 i}$$
$$+ \omega^{\sigma}_{j_1 j_2 j_4 \dots j_k} \wedge \omega_{j_3 i} + \dots + \omega^{\sigma}_{j_1 j_2 \dots j_{k-1}} \wedge \omega_{j_k i}), \tag{12}$$

and

$$\omega^{\sigma}_{j_1 j_2 \dots j_k} \wedge dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_n}$$
$$= -n \delta^{i_1}_{j_1} d\omega^{\sigma}_{j_2 j_3 \dots j_k} \wedge dx^{i_2} \wedge dx^{i_3} \wedge \dots \wedge dx^{i_n} \tag{13}$$
$$\mathrm{Sym}(j_1 j_2 \dots j_k) \quad \mathrm{Alt}(i_1 i_2 \dots i_n).$$

Let now $n+1 \leq q \leq \dim J^r Y$; we define contact q-forms on induction. Let $\rho \in \Omega^r_q W$ be a form. Since $h\rho = 0$, $p_1\rho = 0$, $\dots$, $p_{q-n-1}\rho = 0$, the first canonical decomposition of ρ has the form $(\pi^{r+1,r})^*\rho = p_{q-n}\rho + p_{q-n+1}\rho + \dots + p_q\rho$. ρ is said to be *contact*, if for every point $y_0 \in W$ there exists a fibred chart (V, ψ), $\psi = (x^i, y^\sigma)$, at y_0 and a contact $(q-1)$-form ρ', defined on V^r, such that

$$p_{q-n}(\rho - d\rho') = 0. \tag{14}$$

Contact forms define an Abelian subgroup $\Theta^r_q W$ of the Abelian group $\Omega^r_q W$.

Remark 2 A q-form η such that $p_{q-n}\eta = 0$ is characterized by the property that, in a chart expression, the coefficient at the *local volume element* on X, $dx^1 \wedge dx^2 \wedge \dots \wedge dx^n$, vanishes. Therefore, only products $dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_k}$ with $0 \leq k \leq n-1$ may appear in the chart expression of η.

Lemma 3 *Let* $n+1 \leq q \leq \dim J^r Y$.

(a) *A form* $\rho \in \Omega^r_q W$ *is contact if and only if for every point* $y_0 \in W$ *there exists a fibred chart* (V, ψ), $\psi = (x^i, y^\sigma)$, *at* y_0, *a contact* $(q-1)$*-form* ρ', *and a* q*-form* ρ_0 *on* V^r *of order of contactness* $> q-n$ *such that*

$$\rho = \rho_0 + d\rho'. \tag{15}$$

(b) *The exterior derivative of a contact* q*-form is a contact* $(q+1)$*-form.*

Assertion (a) is a simple restatement of the definition, and (b) follows from formula (15).

Remark 3 Let $\rho \in \Omega^r_q W$ be a form such that $q \leq n-1$. Since

$$(\pi^{r+2,r+1})^* h d\rho = h d h\rho, \tag{16}$$

condition $h\rho = 0$ implies $hd\rho = 0$; in particular, if ρ is contact, then $d\rho$ is also contact. On the other hand, if $q \geq n$, we have

$$(\pi^{r+2,r+1})^* p_{q-n+1} d\rho = p_{q-n+1} d p_{q-n}\rho + p_{q-n+1} d p_{q-n+1}\rho, \tag{17}$$

so it is not true, in general, that the condition $p_{q-n}\rho = 0$ implies $p_{q-n+1}d\rho = 0$. Note, however, that by the definition of the Abelian groups $\Theta^r_q W$, the exterior derivative of a contact q-form is again a contact q-form.

Theorem 2 *Let W be an open set in Y, $q \geq 1$ an integer, $\rho \in \Omega^r_q W$ a form, and let (V, ψ), $\psi = (x^i, y^\sigma)$, be a fibred chart on Y such that $V \subset W$. Assume that on V^r, ρ has a chart expression*

$$\begin{aligned}\rho = \sum_{s=0}^{q} \frac{1}{s!(q-s)!} A^{I_1 I_2 \dots I_s}_{\sigma_1 \sigma_2 \dots \sigma_s i_{s+1} i_{s+2} \dots i_q} dy^{\sigma_1}_{I_1} \wedge dy^{\sigma_2}_{I_2} \wedge \dots \wedge dy^{\sigma_s}_{I_s} \\ \wedge dx^{i_{s+1}} \wedge dx^{i_{s+2}} \wedge \dots \wedge dx^{i_q}.\end{aligned} \tag{18}$$

Then the k-contact component $p_k \rho$ of ρ has on V^{r+1} a chart expression

$$\begin{aligned}p_k \rho = \frac{1}{k!(q-k)!} B^{I_1 I_2 \dots I_k}_{\sigma_1 \sigma_2 \dots \sigma_k i_{k+1} i_{k+2} \dots i_q} \omega^{\sigma_1}_{I_1} \wedge \omega^{\sigma_2}_{I_2} \wedge \dots \wedge \omega^{\sigma_k}_{I_k} \\ \wedge dx^{i_{k+1}} \wedge dx^{i_{k+2}} \wedge \dots \wedge dx^{i_q},\end{aligned} \tag{19}$$

where

$$\begin{aligned}B^{I_1 I_2 \dots I_k}_{\sigma_1 \sigma_2 \dots \sigma_k i_{k+1} i_{k+2} \dots i_s i_{s+1} \dots i_q} = \sum_{s=k}^{q} \binom{q-k}{q-s} A^{I_1 I_2 \dots I_k I_{k+1} I_{k+2} \dots I_s}_{\sigma_1 \sigma_2 \dots \sigma_k \sigma_{k+1} \sigma_{k+2} \dots \sigma_s i_{s+1} i_{s+2} \dots i_q} \\ \cdot y^{\sigma_{k+1}}_{I_{k+1} i_{k+1}} y^{\sigma_{k+2}}_{I_{k+2} i_{k+2}} \dots y^{\sigma_s}_{I_s i_s} \quad \mathrm{Alt}(i_{k+1} i_{k+2} \dots i_s i_{s+1} \dots i_q).\end{aligned} \tag{20}$$

Formula (20) can be derived by computing the value $p_k \rho (J^{r+1}_x \gamma)(\xi_1, \xi_2, \dots, \xi_q)$ at a point $J^{r+1}_x \gamma \in V^{r+1}$ on tangent vectors $\xi_1, \xi_2, \dots, \xi_q$ to $J^{r+1} Y$ at this point, with ρ expressed by (18).

3.3 The second canonical decomposition

We say that a form η is *generated* by a finite family of forms μ_λ, if η is expressible as $\eta = \eta^\lambda \wedge \mu_\lambda$ for some forms η^λ; note that we do not require in this definition μ_λ to be 1-forms, or k-forms for a fixed integer k.

We now give three fundamental theorems on the structure of differential forms on jet prolongations of fibred manifolds. Their proofs rely on algebraic operations, described by the trace decomposition theory, and go outside the scope of this article.

Theorem 3 *Let (V, ψ), $\psi = (x^i, y^\sigma)$, be a fibred chart on Y, $\rho \in \Omega^r_q V$ a q-form. ρ has a unique expression*

$$\rho = \rho_0 + \rho' \tag{1}$$

such that ρ_0 is generated by contact 1-forms ω^ν_J with $|J| \leq r-1$ and contact 2-forms $d\omega^\nu_I$ with $|I| = r-1$, and

$$\begin{aligned}\rho' = {} & A_{i_1 i_2 \dots i_q} dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_q} \\ & + A^{I_1}_{\sigma_1 i_2 i_3 \dots i_q} dy^{\sigma_1}_{I_1} \wedge dx^{i_2} \wedge dx^{i_3} \wedge \dots \wedge dx^{i_q} \\ & + A^{I_1 I_2}_{\sigma_1 \sigma_2 i_3 i_4 \dots i_q} dy^{\sigma_1}_{I_1} \wedge dy^{\sigma_2}_{I_2} \wedge dx^{i_3} \wedge dx^{i_4} \wedge \dots \wedge dx^{i_q} \\ & + \dots + A^{I_1 I_2 \dots I_{q-1}}_{\sigma_1 \sigma_2 \dots \sigma_{q-1} i_q} dy^{\sigma_1}_{I_1} \wedge dy^{\sigma_2}_{I_2} \wedge \dots \wedge dy^{\sigma_{q-1}}_{I_{q-1}} \wedge dx^{i_q} \\ & + A^{I_1 I_2 \dots I_q}_{\sigma_1 \sigma_2 \dots \sigma_q} dy^{\sigma_1}_{I_1} \wedge dy^{\sigma_2}_{I_2} \wedge \dots \wedge dy^{\sigma_q}_{I_q},\end{aligned} \tag{2}$$

where the multi-indices $I_1, I_2, \ldots, I_{q-1}$ *are of length* r*, and all the coefficients* $A^{I_1}_{\sigma_1 i_2 i_3 \ldots i_q}$, $A^{I_1 I_2}_{\sigma_1 \sigma_2 i_3 i_4 \ldots i_q}, \ldots, A^{I_1 I_2 \ldots I_{q-1}}_{\sigma_1 \sigma_2 \ldots \sigma_{q-1} i_q}$ *are traceless.*

The decomposition (1) is called the *canonical decomposition, associated with the chart* (V, ψ), or the *second canonical decomposition* of ρ.

Remark 4 It is easily seen that the decomposition (1) is not invariant.

If J and I are multi-indices such that $|J| \leq r-2$ and $|I| \leq r-1$, then

$$d\omega^\sigma_J = -dy^\sigma_{Jj} \wedge dx^j = -\omega^\sigma_{Jj} \wedge dx^j, \quad d\omega^\sigma_I = -dy^\sigma_{Ij} \wedge dx^j. \tag{3}$$

In particular, the contact 2-forms $d\omega^\sigma_J$ with $|J| \leq r-2$, are expressible as linear combinations of the contact 1-forms ω^σ_J. On the other hand, the contact 2-forms $d\omega^\sigma_I$, where $|I| = r-1$, are *not* expressible as linear combinations of the contact forms ω^σ_J.

The following is the structure theorem for contact q-forms with $1 \leq q \leq n$.

Theorem 5 *Let* $W \subset Y$ *be an open set,* $\rho \in \Omega^r_q W$ *a form, and let* (V, ψ), $\psi = (x^i, y^\sigma)$, *be a fibred chart such that* $V \subset W$.

(a) *Let* $q = 1$. *Then* ρ *is* π*-contact if and only if*

$$\rho = \Phi^J_\sigma \omega^\sigma_J \tag{4}$$

for some functions $\Phi^J_\sigma : V^r \to \boldsymbol{R}$*, where* $|J| \leq r-1$.

(b) *Let* $2 \leq q \leq n$. *Then* ρ *is* π*-contact if and only if*

$$\rho = \omega^\sigma_J \wedge \Phi^J_\sigma + d\omega^\sigma_I \wedge \Psi^I_\sigma, \tag{5}$$

where Φ^J_σ (Ψ^I_σ) *are some* $(q-1)$*-forms* ($(q-2)$*-forms*) *on* V^r*, and* $|J| \leq r-1$, $|I| = r-1$.

We now turn to the case of q-forms such that $q \geq n+1$. We find a formula for forms ρ, satisfying $p_{q-n}\rho = 0$; this gives us, together with the definition, a characterization of contact forms.

Theorem 6 *Let* $W \subset Y$ *be an open set,* q *an integer such that* $n+1 \leq q \leq \dim J^r Y$, $\rho \in \Omega^r_q W$ *a form, and let* (V, ψ), $\psi = (x^i, y^\sigma)$, *be a fibred chart such that* $V \subset W$. *Then* ρ *satisfies* $p_{q-n}\rho = 0$ *if and only if*

$$\begin{aligned}\rho = \sum_{\substack{q-n+1 \leq p+s \\ p+2s \leq q}} & \omega^{\sigma_1}_{J_1} \wedge \omega^{\sigma_2}_{J_2} \wedge \ldots \wedge \omega^{\sigma_p}_{J_p} \\ & \wedge d\omega^{\nu_1}_{I_1} \wedge d\omega^{\nu_2}_{I_2} \wedge \ldots \wedge d\omega^{\nu_s}_{I_s} \wedge \Phi^{J_1 J_2 \ldots J_p I_1 I_2 \ldots I_s}_{\sigma_1 \sigma_2 \ldots \sigma_p \nu_1 \nu_2 \ldots \nu_s},\end{aligned} \tag{6}$$

where $|J_1|, |J_2|, \ldots, |J_p| \leq r-1$, $|I_1|, |I_2|, \ldots, |I_s| = r-1$, *and* $\Phi^{J_1 J_2 \ldots J_p I_1 I_2 \ldots I_s}_{\sigma_1 \sigma_2 \ldots \sigma_p \nu_1 \nu_2 \ldots \nu_s}$ *are some* $(q-p-2s)$*-forms on* V^r.

Note that the summation in expression (6) can also be expressed as

$$\begin{aligned}\rho = \sum_{0 \leq p \leq q} \sum_{q-p-n+1 \leq s \leq 1/2(q-p)} & \omega^{\sigma_1}_{J_1} \wedge \omega^{\sigma_2}_{J_2} \wedge \ldots \wedge \omega^{\sigma_p}_{J_p} \\ & \wedge d\omega^{\nu_1}_{I_1} \wedge d\omega^{\nu_2}_{I_2} \wedge \ldots \wedge d\omega^{\nu_s}_{I_s} \wedge \Phi^{J_1 J_2 \ldots J_p I_1 I_2 \ldots I_s}_{\sigma_1 \sigma_2 \ldots \sigma_p \nu_1 \nu_2 \ldots \nu_s}.\end{aligned} \tag{7}$$

3.4 Contact components and geometric operations

Now we study the differential-geometric operations on forms, such as the wedge product $\wedge$, the exterior derivative d, the contraction i_ζ of a form by a vector ζ, and the Lie derivative ∂_ξ by a vector field ξ. The following formulas are immediate consequences of definitions.

Theorem 7 *Let ρ and η be two differential forms on J^rY, $J_x^{r+1}\gamma \in J^{r+1}Y$ a point, ζ a tangent vector at this point, and ξ a π-projectable vector field on Y. Then*

$$p_k(\rho \wedge \eta) = \sum_{i+j=k} p_i\rho \wedge p_j\eta, \tag{1}$$

$$i_\zeta p_k\rho(J_x^{r+1}\gamma) = p_{k-1}i_{p\zeta}\rho(J_x^{r+1}\gamma) + p_k i_{h\zeta}\rho(J_x^{r+1}\gamma), \tag{2}$$

$$p_k(J^r\alpha^*\rho) = J^{r+1}\alpha^* p_k\rho, \tag{3}$$

$$p_k(\partial_{J^r\xi}\rho) = \partial_{J^{r+1}\xi} p_k\rho, \tag{4}$$

$$(\pi^{r+2,r+1})^* p_k d\rho = p_k dp_{k-1}\rho + p_k dp_k\rho. \tag{5}$$

Remark 5 If $k = 0$, then (1) reduces to the condition $h(\rho \wedge \eta) = h\rho \wedge h\eta$, stating that h is an exterior algebra morphism.

Remark 6 If ζ is a π^{r+1}-vertical, $\pi^{r+1,r}$-projectable vector field, with $\pi^{r+1,r}$-projection ζ_0, then $p\zeta = \zeta_0$, and

$$i_\zeta p_k\rho = p_{k-1}i_{\zeta_0}\rho. \tag{6}$$

3.5 Fibred homotopy operators

In this subsection we specify and refine the concept of a fibred homotopy operator, introduced in Subsection 2.1 for differential forms on fibred manifolds.

Let $U \subset \mathbf{R}^n$ be an open set, let $W \subset \mathbf{R}^m$ be an open ball with centre at the origin, and denote $V = U \times W$. We consider V as a fibred manifold over U with the first Cartesian projection $\pi : V \to U$, and denote by V^s the *s-jet prolongation* of V; explicitly, $V^s = J^s(U \times W)$, that is,

$$V^s = U \times W \times L(\mathbf{R}^n, \mathbf{R}^m) \times L^2_{\text{sym}}(\mathbf{R}^n, \mathbf{R}^m) \times \ldots \times L^s_{\text{sym}}(\mathbf{R}^n, \mathbf{R}^m), \tag{1}$$

where $L^k_{\text{sym}}(\mathbf{R}^n, \mathbf{R}^m)$ is the vector space of k-linear, symmetric mappings from $\mathbf{R}^n$ to $\mathbf{R}^m$. The Cartesian coordinates on V, and the associated jet coordinates on V^s, are denoted by x^i, y^σ, and x^i, y^σ, $y^\sigma_{j_1}$, $y^\sigma_{j_1j_2}$, $\ldots$, $y^\sigma_{j_1j_2\ldots j_s}$, respectively. We denote by $\zeta_s : U \to V^s$ the *zero section*.

We have a mapping χ_s, from the set $[0,1] \times V^s$ to V^s, given by

$$\chi_s(t, (x^i, y^\sigma, y^\sigma_{j_1}, y^\sigma_{j_1j_2}, \ldots, y^\sigma_{j_1j_2\ldots j_s})) = (x^i, ty^\sigma, ty^\sigma_{j_1}, ty^\sigma_{j_1j_2}, \ldots, ty^\sigma_{j_1j_2\ldots j_s}). \tag{2}$$

χ_s defines the *fibred homotopy operator* I_s, assigning to a k-form ρ on V^s, where $k \geq 1$, a $(k-1)$-form $I_s\rho$ on V^s, such that

$$\rho = I_s d\rho + dI_s\rho + (\pi^s)^*\zeta^*\rho. \tag{3}$$

Recall that by definition,

$$I_s\rho = \int I_{s,0}\rho \tag{4}$$

(integration through t from 0 to 1), where $I_{s,0}\rho$ is defined by the decomposition

$$\chi_s^*\rho = dt \wedge I_{s,0}\rho + I'_s\rho \tag{5}$$

such that the k-forms $I_{s,0}\rho$ and $I'_s\rho$ do not contain dt.

Lemma 4 (a) *The mapping χ_s satisfies*

$$\begin{aligned}&\chi_s^* dx^i = dx^i, \quad \chi_s^* dy_J^\sigma = y_J^\sigma dt + t dy_J^\sigma, \quad 0 \le |J| \le s,\\ &\chi_s^*\omega_J^\sigma = y_J^\sigma dt + t\omega_J^\sigma, \quad 0 \le |J| \le s-1.\end{aligned} \tag{6}$$

(b) *Let $q \ge 1$ and let $\rho \in \Omega_q^r V$ be a q-form. Then for every k, $1 \le k \le q$,*

$$I_{s+1}p_k\rho = p_{k-1}I_s\rho. \tag{7}$$

Remark 8 The fibred homotopy operator χ_s, defined by (2), corresponds with the projection $\pi^{s,0} : V^s \to V$. Note, however, that one can also consider different fibred homotopies, associated with the projection $\pi^{s,r} : V^s \to V^r$,

$$\begin{aligned}&\chi(t, (x^i, y^\sigma, y_{j_1}^\sigma, y_{j_1 j_2}^\sigma, \dots, y_{j_1 j_2 \dots j_r}^\sigma, y_{j_1 j_2 \dots j_{r+1}}^\sigma, \dots, y_{j_1 j_2 \dots j_s}^\sigma))\\ &\quad = (x^i, y^\sigma, y_{j_1}^\sigma, y_{j_1 j_2}^\sigma, \dots, y_{j_1 j_2 \dots j_r}^\sigma, t y_{j_1 j_2 \dots j_{r+1}}^\sigma, \dots, t y_{j_1 j_2 \dots j_s}^\sigma).\end{aligned} \tag{8}$$

As a consequence of Lemma 2 we show that every closed contact form ρ on J^rY can locally be expressed as $\rho = d\eta$ for some contact form η.

Lemma 5 *Let $1 \le q \le n$, and let ρ be a contact q-form on J^rY. Then the following two conditions are equivalent:*

(a) $p_1 d\rho = 0$.

(b) *In any fibred chart, ρ has a chart expression*

$$\rho = \rho_0 + d\eta, \tag{9}$$

where the order of contactness of ρ_0 is ≥ 2 and η is a contact form.

The proof is based on the second canonical decomposition theorem.

We can now easily conclude that the following assertion holds.

Theorem 8 *Let q be a positive integer, and let $\rho \in \Omega_q^r W$ be a contact q-form. The following conditions are equivalent:*

(a) $d\rho = 0$.

(b) *For every point $y_0 \in W$ there exists a fibred chart (V, ψ), $\psi = (x^i, y^\sigma)$, at y_0, such that $V \subset W$, and a contact $(q-1)$-form μ, defined on V^r, such that $\rho = d\mu$.*

4 Lagrange structures

Throughout this section, Y is a fibred manifold with *orientable* base manifold X, and projection π. However, orientability of X is not an essential assumption; replacing differential forms by *odd base differential forms*, one can also develop the variational theory for *non-orientable* bases X (cf. Krupka [47], Krupka and Musilová [61]). Variational functionals, defined on fibred manifolds over non-orientable bases, appear in the general relativity theory.

4.1 Variational functionals

Let W be an open set in Y. A *Lagrangian* for Y is a π^r-*horizontal* n-form λ on a subset $W^r \subset J^rY$. The number r is called the *order* of λ. In a fibred chart (V, ψ), $\psi = (x^i, y^\sigma)$, on Y, λ has an expression

$$\lambda = \mathcal{L}\omega_0, \tag{1}$$

where $\mathcal{L} : V^r \to \mathbf{R}$ is the *component* of λ with respect to (V, ψ). $\mathcal{L}$ is called the *Lagrange function* associated with λ and (V, ψ). A pair (Y, λ), where λ is a Lagrangian for Y, is called a *Lagrange structure*.

Let $W \subset Y$ be an open set, and let $\lambda \in \Omega^r_{n,X}W$ be a Lagrangian of order r for π. Let $\Omega \subset \pi(W)$ be a compact, n-dimensional submanifold of X with boundary ∂X (a *piece* of X). Denote by $\Gamma_{\Omega,W}(\pi)$ the set of smooth sections of π over Ω, such that $\gamma(\Omega) \subset W$. Then for any section $\gamma \in \Gamma_{\Omega,W}(\pi)$ of Y, the pull-back $J^r\gamma^*\lambda$ is an n-form on an open subset of the n-*dimensional* base manifold X, which can be integrated over Ω. Thus, λ defines a real function $\Gamma_{\Omega,W}(\pi) \ni \gamma \to \lambda_\Omega(\gamma) \in \mathbf{R}$ by

$$\lambda_\Omega(\gamma) = \int_\Omega J^r\gamma^*\lambda. \tag{2}$$

λ_Ω is called the *variational functional* over Ω, associated with λ; sometimes λ_Ω is also called the *action function*, associated with λ.

The subject of the variational analysis on fibred manifolds is to study the behaviour of variational functionals on set of sections $\Gamma_{\Omega,W}(\pi)$, or on subsets of this set, defined by some additional conditions (*constraints*). Sometimes the integration domain Ω is not fixed, but is arbitrary. Then formula (2) defines a *family* of variational functionals labelled by Ω.

Every n-form $\rho \in \Omega^{r-1}_n W$ defines a Lagrangian of order r for Y, namely the Lagrangian $\lambda = h\rho$, the *horizontal component* of ρ. This Lagrangian is said to be *associated* with ρ. Recall that $h\rho$ is a unique π^r-*horizontal* form on $W^r \subset J^rY$ such that $J^{r-1}\gamma^*\rho = J^r\gamma^*h\rho$ for all sections γ of Y defined on $U = \pi(W)$. In particular, the variational functional λ_Ω is expressible in the form

$$\lambda_\Omega(\gamma) = \int_\Omega J^r\gamma^*h\rho = \int_\Omega J^{r-1}\gamma^*\rho. \tag{3}$$

4.2 Variational derivatives

Let U be an open subset of X, let $\gamma : U \to Y$ be a section. Let ξ be a π-projectable vector field on an open set $W \subset Y$ such that $\gamma(U) \subset W$. If α_t is the local 1-parameter group of

ξ, and $\alpha_{(0)t}$ its π-projection, then

$$\gamma_t = \alpha_t \gamma \alpha_{(0)t}^{-1} \tag{1}$$

is a 1-parameter family of *sections* of Y, depending smoothly on the parameter t. Indeed, since $\pi\alpha_t = \alpha_{(0)t}\pi$, we have

$$\begin{aligned} \pi\gamma_t(x) &= \pi\alpha_t\gamma\alpha_{(0)t}^{-1}(x) = \alpha_{(0)t}\pi\gamma\alpha_{(0)t}^{-1}(x) \\ &= \alpha_{(0)t}\alpha_{(0)t}^{-1}(x) = x \end{aligned} \tag{2}$$

on the domain of γ_t. The family γ_t is called the *variation*, or *deformation*, of the section γ, *induced* by the vector field ξ.

Let $U \subset X$ be an open set, and let $\gamma : U \to Y$ be a section. Recall that a *vector field along* γ is a mapping $\Xi : U \to TY$ such that for every point $x \in U$, $\Xi(x) \in T_{\gamma(x)}Y$. The formula

$$\xi_0 = T\pi \cdot \Xi \tag{3}$$

then defines a vector field ξ_0 on U, called the π-*projection* of Ξ.

The following theorem says that every vector field along a section can be extended to a π-projectable vector field. Moreover, the notion of the r-jet prolongation can be generalized to vector fields along sections.

Theorem 1 *Let γ be a section of Y defined on an open set U, Ξ a vector field along γ.*

(a) *There exists a π-projectable vector field ξ, defined on a neighbourhood of the set $\gamma(U) \subset Y$, such that for each $x \in U$,*

$$\xi(\gamma(x)) = \Xi(x). \tag{4}$$

(b) *Any two π-projectable vector fields ξ_1, ξ_2, defined on a neighbourhood of $\gamma(U)$, such that $\xi_1(\gamma(x)) = \xi_2(\gamma(x)) = \Xi(x)$ for all $x \in U$, satisfy*

$$J^r\xi_1(J^r_x\gamma) = J^r\xi_2(J^r_x\gamma). \tag{5}$$

Assertion (a) can be proved by means of a partition of unity, (b) follows from the structure of prolonged local automorphism of a fibred manifold.

A π-projectable vector field ξ , satisfying condition (a) of Theorem 1, is called a π-*projectable extension* of Ξ. Using (b) we may define

$$J^r\Xi(J^r_x\gamma) = J^r\xi(J^r_x\gamma) \tag{6}$$

for any π-projectable extension ξ of Ξ. Then $J^r\Xi$ is a vector field along the r-jet prolongation $J^r\gamma$ of γ; we call this vector field the r-*jet prolongation* of Ξ.

Variations of sections induce the corresponding changes (variations) of the values of variational functionals. Let $\lambda \in \Omega^r_n W$ be a Lagrangian of order r, $\Omega \subset \pi(W)$ a piece of X. Choose an element $\gamma \in \Gamma_{\Omega,W}(\pi)$ and a π-projectable vector field ξ on W, and consider the variation of γ, induced by ξ (see formula (1)). Since the domain of γ_t contains Ω for all sufficiently small t, the value of the variational functional $\Gamma_{\Omega,W}(\pi) \ni \gamma \to \lambda_\Omega(\gamma) \in \mathbf{R}$

at γ_t is defined, and we get a real-valued function, defined on a neighbourhood $(-\varepsilon, \varepsilon)$ of the point $0 \in \mathbf{R}$,

$$(-\varepsilon, \varepsilon) \ni t \to \lambda_{\alpha_{(0)t}(\Omega)}(\alpha_t \gamma \alpha_{(0)t}^{-1}) = \int\limits_{\alpha_{(0)t}(\Omega)} (J^r(\alpha_t \gamma \alpha_{(0)t}^{-1}))^* \lambda \in \mathbf{R}. \tag{7}$$

It is easily seen that this function is smooth. Since

$$J^r(\alpha_t \gamma \alpha_{(0)t}^{-1})^* \lambda = \left(\alpha_{(0)t}^{-1}\right)^* J^r \gamma^* J^r \alpha_t^* \lambda, \tag{8}$$

where $J^r \alpha_t$ is the local 1-parameter group of the r-jet prolongation $J^r \xi$ of the vector field ξ, we have, using properties of the pull-back operation and the theorem on transformation of the integration domain,

$$\int\limits_{\alpha_{(0)t}(\Omega)} (J^r(\alpha_t \gamma \alpha_{(0)t}^{-1}))^* \lambda = \int\limits_{\Omega} J^r \gamma^* (J^r \alpha_t)^* \lambda. \tag{9}$$

Thus, since Ω is compact, differentiability of (7) follows from the theorem on differentiation of an integral, depending upon a parameter.

Differentiating (7) at $t = 0$ one obtains, using (9) and the definition of the Lie derivative,

$$\left(\frac{d}{dt} \lambda_{\alpha_{(0)t}(\Omega)}(\alpha_t \gamma \alpha_{(0)t}^{-1})\right)_0 = \int\limits_{\Omega} J^r \gamma^* \partial_{J^r \xi} \lambda. \tag{10}$$

Note that this expression can be written as

$$(\partial_{J^r \xi} \lambda)_\Omega(\gamma) = \int\limits_{\Omega} J^r \gamma^* \partial_{J^r \xi} \lambda. \tag{11}$$

The number (11) is called the *variation* of the integral variational functional λ_Ω *at the point* γ, *induced by the vector field* ξ.

This formula shows that the function $\Gamma_{\Omega, W}(\pi) \ni \gamma \to (\partial_{J^r \xi} \lambda)_\Omega(\gamma) \in \mathbf{R}$ is the variational functional (over Ω), associated with the Lagrangian $\partial_{J^r \xi} \lambda$. We call this function the *variational derivative*, or the *first variation* of the variational functional λ_Ω by the vector field ξ.

Formula (11) admits a direct generalization. If ζ is another π-projectable vector field on W, then the *second variational derivative*, or the *second variation*, of the variational function λ_Ω by the vector fields ξ and ζ, is the mapping $\Gamma_{\Omega, W}(\pi) \ni \gamma \to (\partial_{J^r \zeta} \partial_{J^r \xi} \lambda)_\Omega(\gamma) \in \mathbf{R}$, defined by the formula

$$(\partial_{J^r \zeta} \partial_{J^r \xi} \lambda)_\Omega(\gamma) = \int\limits_{\Omega} J^r \gamma^* \partial_{J^r \zeta} \partial_{J^r \xi} \lambda. \tag{12}$$

It is now obvious how *higher-order variational derivatives* are defined.

A section $\gamma \in \Gamma_{\Omega, W}(\pi)$ is called a stable point of the variational functional λ_Ω with respect to its variation ξ, if

$$(\partial_{J^r \xi} \lambda)_\Omega(\gamma) = 0. \tag{13}$$

In practice, one usually requires that a section be a *stable point* with respect to a *family* of its variations, defined by the problem considered.

If ρ is an n-form on $J^{r-1}Y$ and $\lambda = h\rho$ is the associated Lagrangian, then the variation of the integral variational functional λ_Ω at a point γ, induced by the vector field ξ, satisfies

$$\int_\Omega J^r\gamma^* \partial_{J^r\xi} h\rho = \int_\Omega J^{r-1}\gamma^* \partial_{J^{r-1}\xi}\rho. \tag{14}$$

In particular,

$$(\partial_{J^r\xi} h\rho)_\Omega = (h\partial_{J^{r-1}\xi}\rho)_\Omega. \tag{15}$$

4.3 Lepage forms

We introduce in this subsection a class of n-forms ρ on J^rY by imposing certain conditions on the exterior derivative $d\rho$. We need three lemmas.

Lemma 1 *Let $\pi : Y \to X$ be a fibred manifold, and let ζ be a vector field on X. There exists a π-projectable vector field ξ on Y whose π-projection is ζ.*

One can construct ξ from ζ with the help of an atlas, consisting of fibred chart, and a subordinate partition of unity.

Let $s \geq 0$ be an integer, let $\rho \in \Omega^s_n W$ be a form. The pull-back $(\pi^{s+1,s})^*\rho$ has an expression

$$(\pi^{s+1,s})^*\rho = f_0\omega_0 + \sum_{k=0}^{s} f_\sigma^{i,j_1j_2\dots j_k}\omega^\sigma_{j_1j_2\dots j_k} \wedge \omega_i + \eta, \tag{1}$$

where the order of contactness of η is ≥ 2. We show that the *symmetric* component of the coefficient $f_\sigma^{j_{s+1},j_1j_2\dots j_s}$ in this expression is always determined by f_0.

Lemma 2 *The coefficients in the chart expression (1) satisfy*

$$f_\sigma^{j_{s+1},j_1j_2\dots j_s} = \frac{\partial f_0}{\partial y^\sigma_{j_1j_2\dots j_sj_{s+1}}} + \tilde{f}_\sigma^{j_{s+1},j_1j_2\dots j_s}, \tag{2}$$

where

$$\tilde{f}_\sigma^{j_{s+1},j_1j_2\dots j_s} = 0 \quad \mathrm{Sym}(j_1j_2\dots j_sj_{s+1}). \tag{3}$$

To prove (2), note that the form $d(\pi^{s+1,s})^*\rho$ has an expression

$$\begin{aligned}(\pi^{s+1,s})^*d\rho &= df_0 \wedge \omega_0 + \sum_{k=0}^{s} df_\sigma^{i,j_1j_2\dots j_k} \wedge \omega^\sigma_{j_1j_2\dots j_k} \wedge \omega_i \\ &+ \sum_{k=0}^{s} f_\sigma^{i,j_1j_2\dots j_k} d\omega^\sigma_{j_1j_2\dots j_k} \wedge \omega_i + d\eta.\end{aligned} \tag{4}$$

Since $d(\pi^{s+1,s})^*\rho = (\pi^{s+1,s})^*d\rho$ is $\pi^{s+1,s}$-projectable, the coefficient at the forms $dy^\sigma_{j_1j_2\dots j_sj_{s+1}} \wedge \omega_0$ on the right should vanish identically. But all terms containing these forms are

$$\begin{aligned}&\frac{\partial f_0}{\partial y^\sigma_{j_1j_2\dots j_{s+1}}}dy^\sigma_{j_1j_2\dots j_sj_{s+1}}\wedge\omega_0-f^{i,j_1j_2\dots j_s}_\sigma dy^\sigma_{j_1j_2\dots j_sl}\wedge dx^l\wedge\omega_i\\&=\left(\frac{\partial f_0}{\partial y^\sigma_{j_1j_2\dots j_{s+1}}}-f^{j_{s+1},j_1j_2\dots j_s}_\sigma\right)dy^\sigma_{j_1j_2\dots j_sj_{s+1}}\wedge\omega_0,\end{aligned}\tag{5}$$

proving (2).

Lemma 3 *Let $W\subset Y$ be an open set, and let $\rho\in\Omega^s_nW$. The following three conditions are equivalent:*

(a) *$p_1d\rho$ is a $\pi^{s+1,0}$-horizontal $(n+1)$-form.*

(b) *For each $\pi^{s,0}$-vertical vector field ξ on W^s,*

$$hi_\xi d\rho=0.\tag{6}$$

(c) *$(\pi^{s+1,s})^*\rho$ has a chart expression* (1)*, where*

$$\begin{aligned}&\frac{\partial f_0}{\partial y^\sigma_{j_1j_2\dots j_k}}-d_pf^{p,j_1j_2\dots j_k}_\sigma-f^{j_k,j_1j_2\dots j_{k-1}}_\sigma=0\quad\mathrm{Sym}(j_1j_2\dots j_k),\\&\quad 1\le k\le s,\\&\frac{\partial f_0}{\partial y^\sigma_{j_1j_2\dots j_{s+1}}}-f^{j_{s+1},j_1j_2\dots j_s}_\sigma=0\quad\mathrm{Sym}(j_1j_2\dots j_{s+1}).\end{aligned}\tag{7}$$

To prove Lemma 3, we proceed in two steps.

1. Let ξ be a vector field on W^s, and let Ξ be a vector field on W^{s+1} such that $T\pi^{s+1,s}\cdot\Xi=\xi\circ\pi^{s+1,s}$ (Lemma 1). Then $i_\Xi(\pi^{s+1,s})^*d\rho=(\pi^{s+1,s})^*i_\xi d\rho$, and the forms on both sides can be canonically decomposed into their contact components. We have

$$i_\Xi p_1d\rho+i_\Xi p_2d\rho+\ldots+i_\Xi p_{n+1}d\rho=hi_\xi d\rho+p_1i_\xi d\rho+\ldots+p_ni_\xi d\rho.\tag{8}$$

Comparing the horizontal components on both sides we get

$$hi_\Xi p_1d\rho=(\pi^{s+2,s+1})^*hi_\xi d\rho.\tag{9}$$

Let $p_1d\rho$ be $\pi^{s+1,0}$-horizontal. Thus, if ξ is $\pi^{s,0}$-vertical, then Ξ is $\pi^{s+1,0}$-vertical, and we get $hi_\Xi p_1d\rho=(\pi^{s+2,s+1})^*hi_\xi d\rho=0$, which implies, by injectivity of the mapping $(\pi^{s+2,s+1})^*$, that $hi_\xi d\rho=0$. Conversely, let $hi_\xi d\rho=0$ for each $\pi^{s,0}$-vertical vector field ξ. Then by (9), $hi_\Xi p_1d\rho=i_\Xi p_1d\rho=0$ for all $\pi^{s+1,s}$-projectable, $\pi^{s+1,0}$-vertical vector fields Ξ. If in a fibred chart,

$$\Xi=\sum_{k=1}^{s+1}\Xi^\sigma_{j_1j_2\dots j_k}\frac{\partial}{\partial y^\sigma_{j_1j_2\dots j_k}},\quad p_1d\rho=\sum_{k=0}^{s}A^{j_1j_2\dots j_k}_\sigma\omega^\sigma_{j_1j_2\dots j_k}\wedge\omega_0,\tag{10}$$

we get

$$A^{j_1j_2\dots j_k}_\sigma=0,\quad 1\le k\le s,\tag{11}$$

proving $\pi^{s+1,0}$-horizontality of $p_1 d\rho$. Thus, conditions (a) and (b) are equivalent.

2. Express $(\pi^{s+1,s})^*\rho$ in a fibred chart by (1). Then

$$\begin{aligned} p_1 d\rho = {} & \left(\frac{\partial f_0}{\partial y^\sigma} - d_p f_\sigma^p \right) \omega^\sigma \wedge \omega_0 \\ & + \sum_{k=1}^{s} \left(\frac{\partial f_0}{\partial y^\sigma_{j_1 j_2 \dots j_k}} - d_p f_\sigma^{p, j_1 j_2 \dots j_k} - f_\sigma^{j_k, j_1 j_2 \dots j_{k-1}} \right) \omega^\sigma_{j_1 j_2 \dots j_k} \wedge \omega_0 \\ & + \left(\frac{\partial f_0}{\partial y^\sigma_{j_1 j_2 \dots j_{s+1}}} - f_\sigma^{j_{s+1}, j_1 j_2 \dots j_s} \right) \omega^\sigma_{j_1 j_2 \dots j_s j_{s+1}} \wedge \omega_0 . \end{aligned} \tag{12}$$

Since the last term vanishes identically (Lemma 2), we get

$$\begin{aligned} p_1 d\rho = {} & \left(\frac{\partial f_0}{\partial y^\sigma} - d_p f_\sigma^p \right) \omega^\sigma \wedge \omega_0 \\ & + \sum_{k=1}^{s} \left(\frac{\partial f_0}{\partial y^\sigma_{j_1 j_2 \dots j_k}} - d_p f_\sigma^{p, j_1 j_2 \dots j_k} - f_\sigma^{j_k, j_1 j_2 \dots j_{k-1}} \right) \omega^\sigma_{j_1 j_2 \dots j_k} \wedge \omega_0 . \end{aligned} \tag{13}$$

This formula proves equivalence of conditions (a) and (c).

Any form $\rho \in \Omega_n^s W$ satisfying equivalent conditions of Lemma 3 is called a *Lepage form*.

The following basic theorem describes the structure of Lepage forms.

Theorem 2 *Let $W \subset Y$ be an open set. A form $\rho \in \Omega_n^s W$ is a Lepage form if and only if for any fibred chart (V, ψ), $\psi = (x^i, y^\sigma)$, on Y such that $V \subset W$, $(\pi^{s+1,s})^*\rho$ has an expression*

$$(\pi^{s+1,s})^*\rho = \Theta + d\mu + \eta, \tag{14}$$

where

$$\Theta = f_0 \omega_0 + \sum_{k=0}^{s} \left(\sum_{l=0}^{s-k} (-1)^l d_{p_1} d_{p_2} \dots d_{p_l} \frac{\partial f_0}{\partial y^\sigma_{j_1 j_2 \dots j_k p_1 p_2 \dots p_l i}} \right) \omega^\sigma_{j_1 j_2 \dots j_k} \wedge \omega_i , \tag{15}$$

f_0 is a function, defined by the chart expression $h\rho = f_0 \omega_0$, μ is a contact $(n-1)$-form, and the order of contactness of η is ≥ 2.

We prove Theorem 2 in four steps.

1. Decomposing the systems of functions $f_\sigma^{p, j_1 j_2 \dots j_k}$ (1) into their symmetric and complementary parts, we have

$$f_\sigma^{p, j_1 j_2 \dots j_k} = F_\sigma^{p, j_1 j_2 \dots j_k} + G_\sigma^{p, j_1 j_2 \dots j_k}, \tag{16}$$

where

$$\begin{aligned} & F_\sigma^{p, j_1 j_2 \dots j_k} = f_\sigma^{p, j_1 j_2 \dots j_k} \quad \mathrm{Sym}(p j_1 j_2 \dots j_k), \\ & G_\sigma^{p, j_1 j_2 \dots j_k} = 0 \quad \mathrm{Sym}(p j_1 j_2 \dots j_k), 1 \leq k \leq s. \end{aligned} \tag{17}$$

The function $G_\sigma^{p,j_1j_2\dots j_k}$ is the sum of k terms corresponding to symmetries of Young schemes with one row (of k elements) and one column (of two elements).

Then by (7),

$$F_\sigma^{j_k,j_1j_2\dots j_{k-1}} = \frac{\partial f_0}{\partial y^\sigma_{j_1j_2\dots j_k}} - d_pF_\sigma^{p,j_1j_2\dots j_k} - d_pG_\sigma^{p,j_1j_2\dots j_k},$$
$$1 \le k \le s, \tag{18}$$
$$F_\sigma^{j_{s+1},j_1j_2\dots j_s} = \frac{\partial f_0}{\partial y^\sigma_{j_1j_2\dots j_{s+1}}},$$

and we obtain

$$\begin{aligned} F_\sigma^{j_k,j_1j_2\dots j_{k-1}} &= \frac{\partial f_0}{\partial y^\sigma_{j_1j_2\dots j_k}} - d_{p_1}\frac{\partial f_0}{\partial y^\sigma_{j_1j_2\dots j_kp_1}} + d_{p_2}d_{p_1}\frac{\partial f_0}{\partial y^\sigma_{j_1j_2\dots j_kp_1p_2}} \\ &- \dots + (-1)^{s-k}d_{p_{s-k}}\dots d_{p_2}d_{p_1}\frac{\partial f_0}{\partial y^\sigma_{j_1j_2\dots j_kp_1p_2\dots p_{s-k}}} \\ &-(-1)^{s-k}d_{p_1}d_{p_2}\dots d_{p_{s-k+1}}F_\sigma^{p_{s-k+1},j_1j_2\dots j_kp_1p_2\dots p_{s-k}} \\ &-(-1)^{s-k}d_{p_1}d_{p_2}\dots d_{p_{s-k+1}}G_\sigma^{p_{s-k+1},j_1j_2\dots j_kp_1p_2\dots p_{s-k}} \\ &+ \dots - d_{p_1}d_{p_2}d_{p_3}G_\sigma^{p_3,j_1j_2\dots j_kp_1p_2} \\ &+d_{p_1}d_{p_2}G_\sigma^{p_2,j_1j_2\dots j_kp_1} - d_{p_1}G_\sigma^{p_1,j_1j_2\dots j_k}, \quad 1 \le k \le s. \end{aligned} \tag{19}$$

Now we apply the identity

$$\frac{\partial f_0}{\partial y^\sigma_{j_1j_2\dots j_{s+1}}} - f_\sigma^{j_{s+1},j_1j_2\dots j_s} = 0 \quad \mathrm{Sym}(j_1j_2\dots j_{s+1}) \tag{20}$$

(Lemma 2). We get

$$\begin{aligned} F_\sigma^{j_k,j_1j_2\dots j_{k-1}} &= \sum_{l=0}^{s-k+1}(-1)^l d_{p_1}d_{p_2}\dots d_{p_l}\frac{\partial f_0}{\partial y^\sigma_{j_1j_2\dots j_kp_1p_2\dots p_l}} \\ &+ \sum_{l=1}^{s-k+1}(-1)^l d_{p_1}d_{p_2}\dots d_{p_l}G_\sigma^{p_l,j_1j_2\dots j_kp_1p_2\dots p_{l-1}}, \quad 1 \le k \le s, \end{aligned} \tag{21}$$

and

$$F_\sigma^{j_{s+1},j_1j_2\dots j_s} = \frac{\partial f_0}{\partial y^\sigma_{j_1j_2\dots j_{s+1}}}, \tag{22}$$

i.e.,

$$\begin{aligned} F_\sigma^{i,j_1j_2\dots j_k} &= \sum_{l=0}^{s-k}(-1)^l d_{p_1}d_{p_2}\dots d_{p_l}\frac{\partial f_0}{\partial y^\sigma_{j_1j_2\dots j_kip_1p_2\dots p_l}} \\ &+ \sum_{l=1}^{s-k}(-1)^l d_{p_1}d_{p_2}\dots d_{p_l}G_\sigma^{p_l,j_1j_2\dots j_kip_1p_2\dots p_{l-1}}, \quad 0 \le k \le s-1. \end{aligned} \tag{23}$$

These formulas determine the coefficients in the chart expression (1) of $(\pi^{s+1,s})^*\rho$. By (16) and (20),

$$\begin{aligned}(\pi^{s+1,s})^*\rho &= f_0\omega_0\\ &+\sum_{k=0}^{s}\left(\sum_{l=0}^{s-k}(-1)^l d_{p_1}d_{p_2}\dots d_{p_l}\frac{\partial f_0}{\partial y^\sigma_{j_1j_2\dots j_kp_1p_2\dots p_li}}\right)\omega^\sigma_{j_1j_2\dots j_k}\wedge\omega_i\\ &+\sum_{k=0}^{s-1}\left(\sum_{l=1}^{s-k}(-1)^l d_{p_1}d_{p_2}\dots d_{p_l}G_\sigma^{p_l,j_1j_2\dots j_kip_1p_2\dots p_{l-1}}\right)\omega^\sigma_{j_1j_2\dots j_k}\wedge\omega_i\\ &+\sum_{k=1}^{s}\frac{1}{k!}G_\sigma^{i,j_1j_2\dots j_k}\omega^\sigma_{j_1j_2\dots j_k}\wedge\omega_i+\eta.\end{aligned}\tag{24}$$

Thus,

$$(\pi^{s+1,s})^*\rho = \Theta+\nu+\eta,\tag{25}$$

where

$$\Theta = f_0\omega_0+\sum_{k=0}^{s}\left(\sum_{l=0}^{s-k}(-1)^l d_{p_1}d_{p_2}\dots d_{p_l}\frac{\partial f_0}{\partial y^\sigma_{j_1j_2\dots j_kp_1p_2\dots p_li}}\right)\omega^\sigma_{j_1j_2\dots j_k}\wedge\omega_i,\tag{26}$$

and

$$\begin{aligned}\nu &= \sum_{k=1}^{s}G_\sigma^{i,j_1j_2\dots j_k}\omega^\sigma_{j_1j_2\dots j_k}\wedge\omega_i\\ &+\sum_{k=0}^{s-1}\left(\sum_{l=1}^{s-k}(-1)^l d_{p_1}d_{p_2}\dots d_{p_l}G_\sigma^{p_l,j_1j_2\dots j_kip_1p_2\dots p_{l-1}}\right)\omega^\sigma_{j_1j_2\dots j_k}\wedge\omega_i,\end{aligned}\tag{27}$$

and the order of contactness of η is ≥ 2. Writing ν in a more explicit way, we finally have

$$\begin{aligned}\nu &= \sum_{l=1}^{s}(-1)^l d_{p_1}d_{p_2}\dots d_{p_l}G_\sigma^{p_l,ip_1p_2\dots p_{l-1}}\omega^\sigma\wedge\omega_i\\ &+\sum_{k=1}^{s-1}\left(G_\sigma^{i,j_1j_2\dots j_k}+\sum_{l=1}^{s-k}(-1)^l d_{p_1}d_{p_2}\dots d_{p_l}G_\sigma^{p_l,j_1j_2\dots j_kip_1p_2\dots p_{l-1}}\right)\omega^\sigma_{j_1j_2\dots j_k}\wedge\omega_i\\ &+G_\sigma^{i,j_1j_2\dots j_s}\omega^\sigma_{j_1j_2\dots j_s}\wedge\omega_i.\end{aligned}\tag{28}$$

2. We wish to show by the method of undetermined coefficients that there exists a contact $(n-1)$-form μ such that $p_1d\mu=\nu$. Note that

$$dx^k\omega_{ij}=\delta^k_j\omega_i-\delta^k_i\omega_j.\tag{29}$$

Consider an $(n-1)$-form

$$\mu=\frac{1}{2}\sum_{k=0}^{s}H_\sigma^{i_1i_2,j_1j_2\dots j_k}\omega^\sigma_{j_1j_2\dots j_k}\wedge\omega_{i_1i_2},\tag{30}$$

and compute $p_1 d\mu$. We get

$$p_1 d\mu = -\frac{1}{2} d_p H_\sigma^{i_1 i_2} \omega^\sigma \wedge dx^p \wedge \omega_{i_1 i_2}$$

$$-\frac{1}{2} \sum_{k=1}^{s} \left(d_p H_\sigma^{i_1 i_2, j_1 j_2 \dots j_k} + H_\sigma^{i_1 i_2, j_1 j_2 \dots j_{k-1}} \delta_p^{j_k} \right) \omega^\sigma_{j_1 j_2 \dots j_k} \wedge dx^p \wedge \omega_{i_1 i_2} \quad (31)$$

$$-\frac{1}{2} H_\sigma^{i_1 i_2, j_1 j_2 \dots j_s} \omega^\sigma_{j_1 j_2 \dots j_s p} \wedge dx^p \wedge \omega_{i_1 i_2},$$

and using (29),

$$p_1 d\mu = d_p H_\sigma^{pi} \omega^\sigma \wedge \omega_i$$

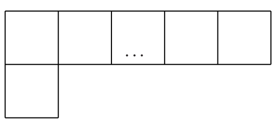
$$+ \sum_{k=1}^{s} (d_p H_\sigma^{pi, j_1 j_2 \dots j_k} + H_\sigma^{j_k i, j_1 j_2 \dots j_{k-1}}) \omega^\sigma_{j_1 j_2 \dots j_k} \wedge \omega_i \quad (32)$$

$$+ H_\sigma^{pi, j_1 j_2 \dots j_s} \omega^\sigma_{j_1 j_2 \dots j_s p} \wedge \omega_i.$$

This expression should now be compared with (28).

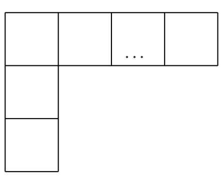
3. To determine the system of functions $H_\sigma^{i_1 i_2, j_1 j_2 \dots j_k}$, $1 \leq k \leq s$, it is sufficient to find independent components of its Young decomposition in the superscripts. Using existing index symmetries, one may easily verify that the independent components correspond with Young diagrams of two groups:

(1) the diagrams with $(k+1)$ columns and 2 rows



(1 diagram), and

(2) the diagrams with k columns and 3 rows



(k diagrams). The contributions of other diagrams vanish identically. Explicitly, a system $H_\sigma^{i_1 i_2, j_1 j_2 \dots j_k}$, antisymmetric in i_1, i_2, and symmetric in $j_1, j_2, \dots, j_k$, has a decomposition

$$H_\sigma^{i_1 i_2, j_1 j_2 \dots j_k} = \frac{k+1}{k+2} ({}^{(1)} H_\sigma^{i_1 i_2, j_1 j_2 \dots j_k} - {}^{(1)} H_\sigma^{i_2 i_1, j_1 j_2 \dots j_k})$$

$$+ \frac{3}{k+2} ({}^{(2)} H_\sigma^{i_1 i_2, j_1 j_2 \dots j_k} + {}^{(2)} H_\sigma^{i_1 i_2, j_2 j_1 j_3 \dots j_k} \quad (33)$$

$$+ \dots + {}^{(2)} H_\sigma^{i_1 i_2, j_k j_2 j_3 \dots j_{k-1} j_1}),$$

where

$$
\begin{aligned}
{}^{(1)}H_\sigma^{i_1i_2,j_1j_2\ldots j_k} &= H_\sigma^{i_1i_2,j_1j_2\ldots j_k} \quad \mathrm{Sym}(i_2j_1j_2\ldots j_k),\\
{}^{(2)}H_\sigma^{i_1i_2,j_1j_2\ldots j_k} &= H_\sigma^{i_1i_2,j_1j_2\ldots j_k} \quad \mathrm{Alt}(i_1i_2j_1).
\end{aligned} \tag{34}
$$

These definitions include the corresponding symmetrization coefficients. Note that if ${}^{(2)}H_\sigma^{i_1i_2,j_1j_2\ldots j_k} = 0$, then the decomposition (33) reduces to

$$
H_\sigma^{i_1i_2,j_1j_2\ldots j_k} = \frac{k+1}{k+2}({}^{(1)}H_\sigma^{i_1i_2,j_1j_2\ldots j_k} - {}^{(1)}H_\sigma^{i_2i_1,j_1j_2\ldots j_k}). \tag{35}
$$

4. Comparing the corresponding coefficients in (28) and (32), we obtain

$$
\begin{aligned}
&H_\sigma^{j_{s+1}i,j_1j_2\ldots j_s} = 0 \quad \mathrm{Sym}(j_1j_2\ldots j_sj_{s+1}), \quad k = s+1,\\
&-d_pH_\sigma^{pi,j_1j_2\ldots j_s} - H_\sigma^{j_si,j_1j_2\ldots j_{s-1}}\\
&\quad = G_\sigma^{i,j_1j_2\ldots j_s} \quad \mathrm{Sym}(j_1j_2\ldots j_s), \quad k = s,\\
&-d_pH_\sigma^{pi,j_1j_2\ldots j_{s-1}} - H_\sigma^{j_{s-1}i,j_1j_2\ldots j_{s-2}} = -d_{p_1}G_\sigma^{p_1,j_1j_2\ldots j_{s-1}i}\\
&\quad +G_\sigma^{i,j_1j_2\ldots j_{s-1}} \quad \mathrm{Sym}(j_1j_2\ldots j_{s-1}), \quad k = s-1,\\
&-d_pH_\sigma^{pi,j_1j_2\ldots j_k} - H_\sigma^{j_ki,j_1j_2\ldots j_{k-1}}\\
&\quad = \sum_{l=1}^{s-k}(-1)^l d_{p_1}d_{p_2}\ldots d_{p_l}G_\sigma^{p_1,j_1j_2\ldots j_kip_1p_2\ldots p_{l-1}}\\
&\quad +G_\sigma^{i,j_1j_2\ldots j_k} \quad \mathrm{Sym}(j_1j_2\ldots j_k), \quad 1\le k\le s-1,\\
&-d_pH_\sigma^{pi,j_1} - H_\sigma^{j_1i}\\
&\quad = \sum_{l=1}^{s-1}(-1)^l d_{p_1}d_{p_2}\ldots d_{p_l}G_\sigma^{p_l,j_1ip_1p_2\ldots p_{l-1}} + G_\sigma^{i,j_1}, \quad k=1,\\
&-d_pH_\sigma^{pi} = \sum_{l=1}^{s}(-1)^l d_{p_1}d_{p_2}\ldots d_{p_l}G_\sigma^{p_l,ip_1p_2\ldots p_{l-1}}, \quad k=0.
\end{aligned} \tag{36}
$$

We wish to solve these equations. The first equation ${}^{(1)}H_\sigma^{pi,j_1j_2\ldots j_s} = 0$ implies that also ${}^{(1)}H_\sigma^{ip,j_1j_2\ldots j_s} = 0$. Therefore, ${}^{(1)}H_\sigma^{pi,j_1j_2\ldots j_s}$ is the sum of components of type (2), which are not determined by this equation. We choose ${}^{(2)}H_\sigma^{pi,j_1j_2\ldots j_s} = 0$, which implies

$$
H_\sigma^{pi,j_1j_2\ldots j_s} = 0. \tag{37}
$$

Then the second equation has the form

$$
{}^{(1)}H_\sigma^{ij_s,j_1j_2\ldots j_{s-1}} = G_\sigma^{i,j_1j_2\ldots j_s}. \tag{38}
$$

Consequently,

$$
{}^{(1)}H_\sigma^{j_si,j_1j_2\ldots j_{s-1}} = G_\sigma^{j_s,j_1j_2\ldots j_{s-1}i}. \tag{39}
$$

Again we may choose ${}^{(2)}H_\sigma^{j_s i, j_1 j_2 \dots j_{s-1}} = 0$ which implies, by (35),

$$H_\sigma^{ij_s, j_1 j_2 \dots j_{s-1}} = \frac{s}{s+1}(G_\sigma^{i, j_1 j_2 \dots j_s} - G_\sigma^{j_s, j_1 j_2 \dots j_{s-1} i}). \tag{40}$$

Consider the third equation (36). Using (40), we get

$$ {}^{(1)}H_\sigma^{ij_{s-1}, j_1 j_2 \dots j_{s-2}} = G_\sigma^{i, j_1 j_2 \dots j_{s-1}} - \frac{1}{s+1} d_p G_\sigma^{p, j_1 j_2 \dots j_{s-1} i} - \frac{s}{s+1} d_p G_\sigma^{i, j_1 j_2 \dots j_{s-1} p}. \tag{41}$$

Thus,

$$\begin{aligned} {}^{(1)}H_\sigma^{i_1 i_2, j_1 j_2 \dots j_{s-2}} &= G_\sigma^{i_1, j_1 j_2 \dots j_{s-2} i_2} - \frac{1}{s+1} d_p G_\sigma^{p, j_1 j_2 \dots j_{s-2} i_2 i_1} \\ &\quad - \frac{s}{s+1} d_p G_\sigma^{i_1, j_1 j_2 \dots j_{s-2} i_2 p}, \\ {}^{(1)}H_\sigma^{i_2 i_1, j_1 j_2 \dots j_{s-2}} &= G_\sigma^{i_2, j_1 j_2 \dots j_{s-2} i_1} - \frac{1}{s+1} d_p G_\sigma^{p, j_1 j_2 \dots j_{s-2} i_1 i_2} \\ &\quad - \frac{s}{s+1} d_p G_\sigma^{i_2, j_1 j_2 \dots j_{s-2} i_1 p}. \end{aligned} \tag{42}$$

Finally,

$$\begin{aligned} H_\sigma^{i_1 i_2, j_1 j_2 \dots j_{s-2}} &= \frac{s-1}{s}(G_\sigma^{i_1, j_1 j_2 \dots j_{s-2} i_2} - G_\sigma^{i_2, j_1 j_2 \dots j_{s-2} i_1}) \\ &\quad + \frac{s-1}{s+1} d_p (G_\sigma^{i_2, j_1 j_2 \dots j_{s-2} i_1 p} - G_\sigma^{i_1, j_1 j_2 \dots j_{s-2} i_2 p}). \end{aligned} \tag{43}$$

Summarizing, we get

$$\begin{aligned} H_\sigma^{pi, j_1 j_2 \dots j_s} &= 0, \\ H_\sigma^{ij_s, j_1 j_2 \dots j_{s-1}} &= \frac{s}{s+1}(G_\sigma^{i, j_1 j_2 \dots j_s} - G_\sigma^{j_s, j_1 j_2 \dots j_{s-1} i}), \\ H_\sigma^{i_1 i_2, j_1 j_2 \dots j_{s-2}} &= \frac{s-1}{s}(G_\sigma^{i_1, j_1 j_2 \dots j_{s-2} i_2} - G_\sigma^{i_2, j_1 j_2 \dots j_{s-2} i_1}) \\ &\quad + \frac{s-1}{s+1} d_p (G_\sigma^{i_2, j_1 j_2 \dots j_{s-2} i_1 p} - G_\sigma^{i_1, j_1 j_2 \dots j_{s-2} i_2 p}). \end{aligned} \tag{44}$$

We use these formulas to state an induction hypothesis. We suppose that equation (36) with $k-1 = s-r$ has a solution defined by

$$\begin{aligned} \frac{1}{s-r+1} H_\sigma^{i_1 i_2, j_1 j_2 \dots j_{s-r}} &= \sum_{m=0}^{r-1} \frac{(-1)^m}{s-r+2+m} \\ &\quad \cdot d_{p_1} d_{p_2} \dots d_{p_m} (G_\sigma^{i_1, j_1 j_2 \dots j_{s-r} i_2 p_1 p_2 \dots p_m} - G_\sigma^{i_2, j_1 j_2 \dots j_{s-r} i_1 p_1 p_2 \dots p_m}). \end{aligned} \tag{45}$$

We obtain on induction that this formula holds for all r.

If $r = s$, (45) gives

$$H_\sigma^{i_1 i_2} = \sum_{m=0}^{s-1} \frac{(-1)^m}{m+2} d_{p_1} d_{p_2} \dots d_{p_m} (G_\sigma^{i_1, i_2 p_1 p_2 \dots p_m} - G_\sigma^{i_2, i_1 p_1 p_2 \dots p_m}). \tag{46}$$

Finally, consider the last equation (36). It can be proved by a direct computation that this equation is satisfied identically.

Therefore, we have found a solution μ of the equation $p_1 d\mu = \nu$, and the proof is complete.

Remark 2 The last formula (36) can also be obtained from the previous one by differentiation.

The n-form Θ defined by (15), is called the *principal component* of the Lepage form ρ with respect to the fibred chart (V, ψ). It can be shown that the splitting (14) ρ of is not coordinate independent.

Theorem 3 *Let ρ be a Lepage form expressed as in Theorem 2. Then the form $(\pi^{s+1,s})^* d\rho$ has an expression*

$$(\pi^{s+1,s})^* d\rho = E + F, \tag{47}$$

where E is a 1-contact, $\pi^{s+1,0}$-horizontal $(n+1)$-form, and F is a form whose order of contactness is ≥ 2. E has the chart expression

$$E = \left(\frac{\partial f_0}{\partial y^\sigma} - \sum_{l=1}^{s+1} (-1)^l d_{p_1} d_{p_2} \dots d_{p_l} \frac{\partial f_0}{\partial y^\sigma_{p_1 p_2 \dots p_l}} \right) \omega^\sigma \wedge \omega_0. \tag{48}$$

Indeed, $E = p_1 d\rho$, and $F = p_2 d\rho + p_3 d\rho + \ldots + p_{n+1} d\rho$. By (14), $E = p_1 d\Theta$. We can express Θ as in (13). Since Θ is a Lepage form, we have

$$p_1 d\rho = \left(\frac{\partial f_0}{\partial y^\sigma} - d_p f^p_\sigma \right) \omega^\sigma \wedge \omega_0, \tag{49}$$

where by (15),

$$f^i_\sigma = \sum_{l=0}^{s} (-1)^l d_{p_1} d_{p_2} \dots d_{p_l} \frac{\partial f_0}{\partial y^\sigma_{p_1 p_2 \dots p_l i}}. \tag{50}$$

Remark 3 Clearly, $h\rho = h\Theta = f_0 \omega_0$; thus, the horizontal component $f_0 \omega_0$ of the principal component Θ of ρ is exactly the Lagrangian, associated with ρ. Both forms Θ and E depend only on $h\rho$. For each $\pi^{s,0}$-projectable vector field ξ, the n-form $hi_\xi d\rho$ depends on the $\pi^{s,0}$-projection of ξ only.

One can determine the chart expression $f_0 \omega_0$ of $h\rho$ explicitly by means of the second canonical decomposition of the form ρ. We get

$$\begin{aligned} f_0 = \epsilon^{i_1 i_2 \dots i_n} (& A_{i_1 i_2 \dots i_n} + A^{I_1}_{\sigma_1 i_2 i_3 \dots i_n} y^{\sigma_1}_{I_1 i_1} + A^{I_1 I_2}_{\sigma_1 \sigma_2 i_3 i_4 \dots i_n} y^{\sigma_1}_{I_1 i_1} y^{\sigma_2}_{I_2 i_2} \\ & + \ldots + A^{I_1 I_2 \dots I_{n-1}}_{\sigma_1 \sigma_2 \dots \sigma_{n-1} i_n} y^{\sigma_1}_{I_1 i_1} y^{\sigma_2}_{I_2 i_2} \dots y^{\sigma_{n-1}}_{I_{n-1} i_{n-1}} \\ & + A^{I_1 I_2 \dots I_n}_{\sigma_1 \sigma_2 \dots \sigma_n} y^{\sigma_1}_{I_1 i_1} y^{\sigma_2}_{I_2 i_2} \dots y^{\sigma_n}_{I_n i_n}). \end{aligned} \tag{51}$$

4.4 Lepage equivalents of a Lagrangian

Let W be an open set in Y, and let $\Omega_n W$ be the *direct limit* of the sets $\Omega_n^s W$ with respect to the canonical injections $\Omega_n^s W \ni \eta \to (\pi^{s+1,s})^* \eta \in \Omega_n^{s+1} W$. $\Omega_n W$ consists of equivalence classes of points of the set $\bigcup \Omega_n^s W$, which have a common successor. Let $\rho \in \Omega_n^r W$, $\eta \in \Omega_n^s W$. The binary relation "*$\rho \sim \eta$ if there exists a positive integer q such that* $(\pi^{q,r+1})^* h\rho = (\pi^{q,s+1})^* h\eta$" is an equivalence on the set $\Omega_n W$. Thus, equivalent n-forms have the same associated Lagrangians, and define the same variational functional.

Let λ be a Lagrangian of order r for Y. A *Lepage form* $\rho \in \Omega_n^s W$, equivalent with λ, is said to be a *Lepage equivalent* of λ.

Theorem 4 *Let $W \subset Y$ be an open set, and let $\lambda \in \Omega_n^r W$ be a Lagrangian of order r, (V, ψ), $\psi = (x^i, y^\sigma)$, a fibred chart on Y such that $V \subset W$. Let λ be expressed by*

$$\lambda = \mathcal{L}\omega_0. \tag{1}$$

A Lepage form $\rho \in \Omega_n^s W$ is a Lepage equivalent of λ if and only the principal component Θ of ρ has an expression

$$\Theta = \mathcal{L}\omega_0 + \sum_{k=0}^{r-1} \left(\sum_{l=0}^{r-1-k} (-1)^l d_{p_1} d_{p_2} \dots d_{p_l} \frac{\partial \mathcal{L}}{\partial y^\sigma_{j_1 j_2 \dots j_k p_1 p_2 \dots p_l i}} \right) \omega^\sigma_{j_1 j_2 \dots j_k} \wedge \omega_i. \tag{2}$$

Θ is defined on W^{2r-1}.

By definition, the horizontal component of a Lepage equivalent of a Lagrangian is equal to the Lagrangian. Thus, Theorem 4 is a direct consequence of the definition of a Lepage equivalent, and of Theorem 3.

Remark 4 For example, the principal Lepage equivalent of a Lagrangian of order 3 is given by

$$\begin{aligned} \Theta = \mathcal{L}\omega_0 &+ \left(\frac{\partial \mathcal{L}}{\partial y^\sigma_i} - d_{p_1} \frac{\partial \mathcal{L}}{\partial y^\sigma_{p_1 i}} + d_{p_1} d_{p_2} \frac{\partial \mathcal{L}}{\partial y^\sigma_{p_1 p_2 i}} \right) \omega^\sigma \wedge \omega_i \\ &+ \left(\frac{\partial \mathcal{L}}{\partial y^\sigma_{j_1 i}} - d_{p_1} \frac{\partial \mathcal{L}}{\partial y^\sigma_{j_1 p_1 i}} \right) \omega^\sigma_{j_1} \wedge \omega_i + \frac{\partial \mathcal{L}}{\partial y^\sigma_{j_1 j_2 i}} \omega^\sigma_{j_1 j_2} \wedge \omega_i. \end{aligned} \tag{3}$$

Corollary 1 *If ρ is a Lepage equivalent of a Lagrangian $\lambda = \mathcal{L}\omega_0$ of order r, then*

$$p_1 d\rho = E_\sigma(\mathcal{L}) \omega^\sigma \wedge \omega_0, \tag{4}$$

where

$$E_\sigma(\mathcal{L}) = \sum_{l=0}^{r} (-1)^l d_{p_1} d_{p_2} \dots d_{p_l} \frac{\partial \mathcal{L}}{\partial y^\sigma_{p_1 p_2 \dots p_l}}. \tag{5}$$

For any Lepage equivalent ρ of a Lagrangian λ, the form $p_1 d\rho$ depends only on λ, and is called *the Euler-Lagrange form* associated with λ. We denote

$$p_1 d\rho = E_\lambda. \tag{6}$$

The components $E_\sigma(\mathcal{L})$ of E_λ, defined by (4), are called *the Euler-Lagrange expressions*. Obviously, E_λ belongs to the set $\Omega^{2r}_{n+1,Y}W$. The mapping $\Omega^r_{n,X}W \ni \lambda \to E_\lambda \in \Omega^{2r}_{n+1,Y}W$, assigning to a Lagrangian λ the associated Euler-Lagrange form E_λ is called *the Euler-Lagrange mapping*.

A basic question is the existence of Lepage equivalents. Note that if (V,ψ) is a fibred chart, then for every Lagrangian λ, defined on V^r, Θ is a Lepage equivalent of λ, defined on V^{2r-1}. Thus, the existence problem means the *global existence*.

Theorem 5 *Every Lagrangian of order r has a Lepage equivalent of order $2r-1$, which is $\pi^{2r-1,r-1}$-horizontal.*

Let $\lambda \in \Omega^r_{n,X}Y$ be a Lagrangian. To prove Theorem 5, we associate to any atlas $\{(V_\iota,\psi_\iota)\}$, $\iota \in I$, on Y, consisted of fibred charts, and to a partition of unity $\{\chi_\iota\}$, subordinate to the covering $\{V_\iota\}$ of Y, a family of Lagrangians $\chi_\iota\lambda$, whose support satisfies $\mathrm{supp}\chi_\iota\lambda \subset V^r_\iota$. We define a Lepage equivalent ρ_ι of $\chi_\iota\lambda$ by formula (2) on V^{2r-1}, and the zero form outside V^{2r-1}. The n-form $\rho = \sum\rho_\iota$, defined on $J^{2r-1}Y$, is a Lepage equivalent of λ.

4.5 The first variation formula

Let $\lambda \in \Omega^r_nW$ be a Lagrangian of order r, Ω a piece of X. Consider the variational functional $\Gamma_{\Omega,W}(\pi) \ni \gamma \to \lambda_\Omega(\gamma) \in \mathbf{R}$, and its variational derivative by a π-projectable vector field ξ on W,

$$(\partial_{J^r\xi}\lambda)_\Omega(\gamma) = \int_\Omega J^r\gamma^*\partial_{J^r\xi}\lambda. \tag{1}$$

Our aim now will be to find an appropriate expression for the integral on the right. We are looking for a decomposition of the n-form $\partial_{J^r\xi}\lambda$ into two terms, such that the first one would depend only on ξ, and the second one only on the values of $J^r\xi$ on the boundary $\partial\Omega$ of Ω, not on the values of ξ inside Ω.

Theorem 6 *Let $\lambda \in \Omega^r_nW$ be a Lagrangian, $\rho \in \Omega^s_nW$ a Lepage equivalent of λ, ξ a π-projectable vector field on W.*

(a) *The Lie derivative $\partial_{J^r\xi}\lambda$ can be expressed by the formula*

$$\partial_{J^r\xi}\lambda = hi_{J^s\xi}d\rho + hdi_{J^s\xi}\rho. \tag{2}$$

(b) *For any section γ of Y,*

$$J^r\gamma^*\partial_{J^r\xi}\lambda = J^s\gamma^*i_{J^s\xi}d\rho + dJ^s\gamma^*i_{J^s\xi}\rho. \tag{3}$$

(c) *For every piece Ω of X and every section γ of Y defined on Ω,*

$$\int_\Omega J^r\gamma^*\partial_{J^r\xi}\lambda = \int_\Omega J^s\gamma^*i_{J^s\xi}d\rho + \int_{\partial\Omega} J^s\gamma^*i_{J^s\xi}\rho. \tag{4}$$

For the proof of this theorem, we need properties of Lepage equivalents of λ, the horizontalization morphism h, and the Stokes' formula for integration of differential forms.

We call formula (2) (and also (3)) the *infinitesimal first variation formula*. Formula (4) is called the *integral first variation formula*.

Remark 5 The forms $hi_{J^r\xi}d\rho$ and $J^r\gamma^* i_{J^r\xi}d\rho$ depend on the Lagrangian $\lambda = h\rho$ only, while the remaining terms on the right-hand side of formulas (2), (3), and (4) depend on the choice of the Lepage equivalent ρ.

Remark 6 Theorem 6 can be used to obtain the corresponding formulae for higher variational derivatives (cf. **4.3**).

4.6 Extremals

Let $U \subset X$ be an open set, $\gamma : U \to Y$ a section, and $\Xi : U \to TY$ a vector field along γ. The *support* of Ξ is the set $\text{supp}\Xi = \text{cl}\{x \in U | \Xi(x) \neq 0\}$ (here cl means *closure*). We know that each smooth vector field Ξ along γ can be smoothly prolonged to a π-projectable vector field ξ defined on a neighbourhood V of the set $\gamma(U) \subset Y$. ξ satisfies

$$\xi \circ \gamma = \Xi. \tag{1}$$

Let $\Omega \subset X$ be a piece of X, $W \subset Y$ an open set, and let $\Gamma_{\Omega,W}(\pi)$ be the set of sections $\gamma : U \to Y$ such that $\Omega \subset U$ and $\gamma(\Omega) \subset W$. Let $\lambda \in \Omega^r_{n,X}W$ be a Lagrangian. We say that a section $\gamma \in \Gamma_{\Omega,W}(\pi)$ is an *extremal* of the variational functional $\Gamma_{\Omega,W}(\pi) \ni \gamma \to \lambda_\Omega(\gamma) \in \mathbf{R}$, or an *extremal* of the Lagrangian λ on Ω, if for all π-projectable vector fields ξ, such that $\text{supp}(\xi \circ \gamma) \subset \Omega$,

$$\int_\Omega J^r\gamma^* \partial_{J^r\xi}\lambda = 0. \tag{2}$$

γ is called an *extremal of the Lagrange structure* (Y, λ), or simply an *extremal*, if it is an extremal of the variational functional λ_Ω for every Ω in the domain of definition of γ. Thus, roughly speaking, the extremals are those sections γ for which the values $\lambda_\Omega(\gamma)$ are not sensitive to small compact deformations of γ inside Ω.

In the following necessary and sufficient conditions for a section to be an extremal, we use the *Euler-Lagrange form* E_λ, associated with a Lagrangian λ, and the components of the Euler-Lagrange form $E_\sigma(\mathcal{L})$.

Theorem 7 *Let $\lambda \in \Omega^r_{n,X}W$ be a Lagrangian, E_λ the Euler-Lagrange form associated with λ, $\gamma : U \to Y$ a section of Y, $\Omega \subset U$ a piece of X, and ρ a Lepage equivalent of λ, defined on W^s. The following conditions are equivalent:*

(a) *γ is an extremal on Ω.*

(b) *For every π-vertical vector field ξ defined on a neighbourhood of $\gamma(U)$, such that* $\text{supp}(\xi \circ \gamma) \subset \Omega$,

$$J^s\gamma^* i_{J^s\xi}d\rho = 0. \tag{3}$$

(c) *For every fibred chart (V, ψ), $\psi = (x^i, y^\sigma)$, such that $V \subset W$, γ satisfies the system of partial differential equations*

$$E_\sigma(\mathcal{L}) \circ J^{2r}\gamma = 0, \quad 1 \leq \sigma \leq m. \tag{4}$$

(d) *The Euler-Lagrange form associated with λ vanishes along $J^{2r}\gamma$, i.e.,*

$$E_\lambda \circ J^{2r}\gamma = 0. \tag{5}$$

Main steps of the proof copy classical considerations. By Theorem 6 (c), for any piece Ω of X and any π-projectable vector field ξ such that $\operatorname{supp}(\xi \circ \gamma) \subset \Omega$,

$$\int_{\Omega} J^r\gamma^* \partial_{J^r\xi}\lambda = \int_{\Omega} J^s\gamma^* i_{J^s\xi} d\rho, \tag{6}$$

because the vector field $J^s\xi$ vanishes over the boundary $\partial\Omega$. Then

$$\int_{\Omega} J^s\gamma^* i_{J^s\xi} d\rho = \int_{\Omega} J^{s+1}\gamma^* (\pi^{s+1,s})^* i_{J^s\xi} d\rho = \int_{\Omega} J^{s+1}\gamma^* i_{J^{s+1}\xi} p_1 d\rho, \tag{7}$$

where $p_1 d\rho = E_\lambda$ is the Euler-Lagrange form.

If Ω is contained in a coordinate neighbourhood, the support $\operatorname{supp}(\xi \circ \gamma) \subset \Omega$ lies in the same coordinate neighbourhood. Writing $\xi = \xi^i \cdot \partial/\partial x^i + \Xi^\sigma \cdot \partial/\partial y^\sigma$ and $p_1 d\rho = E_\sigma(\mathcal{L})\omega^\sigma \wedge \omega_0$, we obtain

$$h i_\xi p_1 d\rho = E_\sigma(\mathcal{L})(\Xi^\sigma - y^\sigma_j \xi^j)\omega_0, \tag{8}$$

and

$$J^s\gamma^* i_{J^s\xi} d\rho = E_\sigma(\mathcal{L}) \circ J^{s+1}\gamma \cdot \left(\Xi^\sigma \circ \gamma - \frac{\partial(y^\sigma \circ \gamma)}{\partial x^j}\xi^j \right) \omega_0. \tag{9}$$

Now supposing that $J^s\gamma^* i_{J^s\xi} d\rho \neq 0$ for some π-vertical variation ξ, the first variation formula

$$\int_{\Omega} J^s\gamma^* i_{J^s\xi} d\rho = \int_{\Omega} E_\sigma(L) \circ J^{s+1}\gamma \cdot \left(\Xi^\sigma \circ \gamma - \frac{\partial(y^\sigma \circ \gamma)}{\partial x^j}\xi^j \right) \omega_0 \tag{10}$$

with sufficiently small Ω would give us a contradiction

$$\int_{\Omega} J^r\gamma^* \partial_{J^r\xi}\lambda \neq 0. \tag{11}$$

Thus, (a) implies (b). The same formulas can be used to complete the proof.

Theorem 7 reduces the problem of finding extremals of variational functionals to the problem of solving the *Euler-Lagrange differential equations* (4). Properties of these non-linear equations depend on the Lagrangian; their *global* structure is defined by condition (3). This condition says that a section γ is an extremal if and only if its jet prolongation of order $2r$ is an *integral mapping* of an ideal of forms generated by the family of n-forms $i_{J^s\xi} d\rho$ (here the vector field ξ is arbitrary). Expressing these forms in a chart as a linear combination of the components of the vector field $J^{2r}\xi$, we get *local generators* of this ideal.

Remark 7 In particular, if the base manifold X is 1-dimensional, the local generators of the forms $i_{J^s\xi} d\rho$ are 1-forms, and we get an ideal defining a *distribution* on the manifold J^sY; the jet prolongations of extremals are then *integral mappings* of this distribution.

5 The structure of the Euler-Lagrange mapping

We consider in this section general properties of the Euler-Lagrange mapping $\Omega^r_{n,X}W \ni \lambda \to E_\lambda \in \Omega^{2r}_{n+1,Y}W$, introduced in subsection **4.4**. The Euler-Lagrange mapping is obviously a morphism of Abelian groups. Our aim is to characterize transformation properties of the Euler-Lagrange mapping with respect to local automorphisms of underlying fibred manifold, and its *kernel* and *image*.

An essential tool we need is the Veinberg-Tonti lagrantian (see e.g. Tonti [91]), that appears here, however, in connection with the fibred homotopy operator, and the Volterra-Poincaré lemma for forms on fibred manifolds. We also need some results on solutions of formal divergence equations.

5.1 The Euler-Lagrange mapping and the fibred automorphisms

Let W be an open subset of the base manifold X, let λ be a Lagrangian of order r for Y, defined on W^r , and let E_λ be the Euler-Lagrange form of λ; E_λ is defined on the set $W^{2r} \subset J^{2r}Y$. Recall that if in a fibred chart (V,ψ),$\psi = (x^i, y^\sigma)$, on Y, λ has an expression

$$\lambda = \mathcal{L}\omega_0, \tag{1}$$

then

$$E_\lambda = E_\sigma(\mathcal{L})\omega^\sigma \wedge \omega_0, \tag{2}$$

where $E_\sigma(\mathcal{L})$ are the Euler-Lagrange expressions, defined by

$$E_\sigma(\mathcal{L}) = \sum_{l=0}^{r}(-1)^l d_{p_1} d_{p_2} \dots d_{p_l} \frac{\partial \mathcal{L}}{\partial y^\sigma_{p_1 p_2 \dots p_l}}. \tag{3}$$

The mapping $\Omega^r_{n,X}W \ni \lambda \to E_\lambda \in \Omega^{2r}_{n+1,Y}W$ is the *Euler-Lagrange mapping*.

Theorem 1 *Let $\lambda \in \Omega^r_n W$ be a Lagrangian.*

(a) *Let W_1 be an open subsets of Y, and let $\alpha : W_1 \to W$ be an isomorphism of fibred manifolds over X. Then*

$$E_{J^r\alpha^*\lambda} = J^{2r}\alpha^* E_\lambda. \tag{4}$$

(b) *For every π-projectable vector field Ξ on W,*

$$E_{\partial_{J^r\Xi}\lambda} = \partial_{J^{2r}\Xi} E_\lambda. \tag{5}$$

The proof is based on the properties of Lepage forms. Let ρ be any Lepage form, and consider the associated Lagrangian $\lambda = h\rho$. If λ is defined on $W^r \subset J^rY$, then we may suppose that ρ is defined on $W^{2r-1} \subset J^{2r}Y$. Then, the Euler-Lagrange form, associated with λ, is $E_\lambda = p_{1d\rho}$. Since the mapping $J^{2r-1}\alpha$ commutes the mapping p_1 and preserves the contact forms, we have

$$p_1 dJ^{2r-1}\alpha^*\rho = p_1 J^{2r-1}\alpha^* d\rho = J^{2r}\alpha^* p_1 d\rho. \tag{6}$$

In particular, if ρ is a Lepage form, also $J^{2r-1}\alpha^*\rho$ is a Lepage form. Then by definition of the Euler-Lagrange form, expression (5) is equal to $E_{hJ^{2r-1}\alpha^*\rho}$. But $J^{2r-1}\alpha$ commutes the the horizontalization h; thus, (5) proves assertion (a). Assertion (b) is an immediate consequence of (a).

5.2 Formal divergence equations

In this subsection we show that the Euler-Lagrange expressions characterize integrability conditions of a class of formal differential equations, the *formal divergence equations* (Krupka [51]). Essential parts of the proofs are based on the trace decompositon theory, and on an analysis of symmetries of tensors by means of the Young projectors.

Let $U \subset \mathbf{R}^n$ be an open set, let $W \subset \mathbf{R}^m$ be an open ball with centre at the origin, and denote $V = U \times W$. We consider V as a fibred manifold over U with the first Cartesian projection $\pi : V \to U$. V^s denotes the *s-jet prolongation* of V; explicitly,

$$V^s = U \times W \times L(\mathbf{R}^n, \mathbf{R}^m) \times L^2_{\mathrm{sym}}(\mathbf{R}^n, \mathbf{R}^m) \times \ldots \times L^s_{\mathrm{sym}}(\mathbf{R}^n, \mathbf{R}^m), \tag{1}$$

where $L^k_{\mathrm{sym}}(\mathbf{R}^n, \mathbf{R}^m)$ is the vector space of k-linear, symmetric mappings from $\mathbf{R}^n$ to $\mathbf{R}^m$. The Cartesian coordinates on V, and the associated jet coordinates on V^s, are denoted by x^i, y^σ, and x^i, y^σ, $y^\sigma_{j_1}$, $y^\sigma_{j_1 j_2}$, $\ldots$, $y^\sigma_{j_1 j_2 \ldots j_s}$, respectively.

Let $f : V^r \to \mathbf{R}$ be a differentiable function. Our aim is to find solutions $g = (g^1, g^2, \ldots, g^n)$ of the *formal divergence equation*

$$d_i g^i = f, \tag{2}$$

whose components g^i are differentiable real-valued functions on V^s, where s is a positive integer. Since the *formal divergence* $d_i g^i$ is defined by

$$d_i g^i = \frac{\partial g^i}{\partial x^i} + \frac{\partial g^i}{\partial y^\sigma} y^\sigma_i + \frac{\partial g^i}{\partial y^\sigma_{i_1}} y^\sigma_{i_1 i} + \frac{\partial g^i}{\partial y^\sigma_{i_1 i_2}} y^\sigma_{i_1 i_2 i} + \ldots + \frac{\partial g^i}{\partial y^\sigma_{i_1 i_2 \ldots i_r}} y^\sigma_{i_1 i_2 \ldots i_r i}, \tag{3}$$

equation (2) is a first order partial differential equation. From this expression we immediately see that every solution $g = g^i$, defined on V^s , such that $s \geq r+1$, satisfies

$$\frac{\partial g^{i_1}}{\partial y^\sigma_{i_2 i_3 \ldots i_{s+1}}} + \frac{\partial g^{i_2}}{\partial y^\sigma_{i_1 i_3 i_4 \ldots i_{s+1}}} + \ldots + \frac{\partial g^{i_s}}{\partial y^\sigma_{i_1 i_2 \ldots i_{s-1} i_{s+1}}} + \frac{\partial g^{i_{s+1}}}{\partial y^\sigma_{i_1 i_2 \ldots i_s}} = 0. \tag{4}$$

We shall use our standard notation. We denote $\omega_0 = dx^1 \wedge dx^2 \wedge \ldots \wedge dx^n$, $\omega_i = i_{\partial/\partial x^i}\omega_0$, and $\omega^\sigma_{j_1 j_2 \ldots j_k} = dy^\sigma_{j_1 j_2 \ldots j_k} - y^\sigma_{j_1 j_2 \ldots j_k l} dx^l$ as before. h and p_k are the π-*horizontalization*, and the k-*contact* mappings.

For any smooth function $f : V^r \to \mathbf{R}$ we define an n-form λ_f on V^r and a system of functions $E_\sigma(f) : V^{2r} \to \mathbf{R}$ by $\lambda_f = f\omega_0$, and

$$\begin{aligned} E_\sigma(f) = {} & \frac{\partial f}{\partial y^\sigma} - d_{i_1} \frac{\partial f}{\partial y^\sigma_{i_1}} + d_{i_1} d_{i_2} \frac{\partial f}{\partial y^\sigma_{i_1 i_2}} - d_{i_1} d_{i_2} d_{i_3} \frac{\partial f}{\partial y^\sigma_{i_1 i_2 i_3}} + \ldots \\ & + (-1)^{r-1} d_{i_1} d_{i_2} \ldots d_{i_{r-1}} \frac{\partial f}{\partial y^\sigma_{i_1 i_2 \ldots i_{r-1}}} + (-1)^r d_{i_1} d_{i_2} \ldots d_{i_r} \frac{\partial f}{\partial y^\sigma_{i_1 i_2 \ldots i_r}}. \end{aligned} \tag{5}$$

λ_f is a *Lagrangian*, and the $(n+1)$-form $E_f = E_\sigma(f)\omega^\sigma \wedge \omega_0$ is the *Euler-Lagrange form* associated with λ_f.

Let us consider a π^r-horizontal $(n-1)$-form η on V^r, expressed as

$$\eta = g^i \omega_i = \frac{1}{(n-1)!} h_{j_2 j_3 \ldots j_n} dx^{j_2} \wedge dx^{j_3} \wedge \ldots \wedge dx^{j_n}. \tag{6}$$

Since

$$\omega_i = \frac{1}{(n-1)!}\epsilon_{ij_2j_3\dots j_n} dx^{j_2} \wedge dx^{j_3} \wedge \dots \wedge dx^{j_n} \tag{7}$$

we have the transformation formulas

$$h_{j_2j_3\dots j_n} = \epsilon_{ij_2j_3\dots j_n} g^i, \quad g^k = \frac{1}{(n-1)!}\epsilon^{kj_2j_3\dots j_n} h_{j_2j_3\dots j_n}. \tag{8}$$

In the following two lemmas we denote by Alt and Sym the *alternation* and *symmetrization* in the corresponding indices.

Lemma 1 *The functions g^i and $h_{j_1j_2\dots j_{n-1}}$ satisfy*

$$\begin{aligned}&\frac{1}{r+1}\epsilon_{il_2l_3\dots l_n}\left(\frac{\partial g^i}{\partial y^\sigma_{k_1k_2\dots k_r}} + \frac{\partial g^{k_1}}{\partial y^\sigma_{ik_2k_3\dots k_r}}\right.\\ &\quad\left. + \frac{\partial g^{k_2}}{\partial y^\sigma_{k_1ik_3k_4\dots k_r}} + \dots + \frac{\partial g^{k_r}}{\partial y^\sigma_{k_1k_2\dots k_{r-1}i}}\right)\\ &\quad = \frac{\partial h_{l_2l_3\dots l_n}}{\partial y^\sigma_{k_1k_2\dots k_r}} - \frac{r(n-1)}{r+1}\frac{\partial h_{il_3l_4\dots l_n}}{\partial y^\sigma_{ik_2k_3\dots k_r}}\delta^{k_1}_{l_2}\\ &\quad \mathrm{Alt}(l_2l_3\dots l_n) \quad \mathrm{Sym}(k_1k_2\dots k_r).\end{aligned} \tag{9}$$

We say that a π^r-horizontal form η, defined on V^r, has a *projectable extension*, or, according to Haková and Krupková [31], is *pertinent*, if there exists a form μ on V^{r-1} such that $\eta = h\mu$.

Let us consider a form η, expressed in two bases of $(n-1)$-forms by (6).

Lemma 2 *Let η be a π^r-horizontal $(n-1)$-form on V^r. The following two conditions are equivalent:*

(a) *η has a $\pi^{r,r-1}$-projectable extension.*

(b) *The components $h_{i_1i_2\dots i_{n-1}}$ satisfy*

$$\begin{aligned}&\frac{\partial h_{i_1i_2\dots i_{n-1}}}{\partial y^\sigma_{j_1j_2\dots j_r}} - \frac{r(n-1)}{r+1}\frac{\partial h_{si_2i_3\dots i_{n-1}}}{\partial y^\sigma_{sj_2j_3\dots j_r}}\delta^{j_1}_{i_1} = 0\\ &\mathrm{Sym}(j_1j_2\dots j_r) \quad \mathrm{Alt}(i_1i_2\dots i_{n-1}).\end{aligned} \tag{10}$$

(c) *The components g^i satisfy condition (4).*

A slightly different version of the following assertion has already been presented.

Lemma 3 *For any function $f : V^r \to \boldsymbol{R}$, there exists an n-form Θ, defined on V^{2r-1}, such that* (a) *$\lambda = h\Theta$, and* (b) *the form $p_1 d\Theta$ is ω^σ-generated.*

We can take for Θ the *principal Lepage equivalent*

$$\Theta_f = f\omega_0 + \sum_{k=0}^{s}\left(\sum_{l=0}^{s-k}(-1)^l d_{p_1}d_{p_2}\dots d_{p_l}\frac{\partial f}{\partial y^\sigma_{j_1j_2\dots j_kp_1p_2\dots p_li}}\right)\omega^\sigma_{j_1j_2\dots j_k}\wedge\omega_i, \tag{11}$$

of the Lagrangian $\lambda = f\omega_0$.

We prove two theorems, describing solutions of the formal divergence equation (2). By a *solution* of this equation we mean any system of functions $g = g^i$, defined on V^s for some non-negative integer s, satisfying this equation.

Lemma 4 *If the formal divergence equation (2) has a solution defined on V^s and $s \geq r+1$, then it has a solution defined on V^{s-1}.*

Theorem 2 *Let $f : V^r \to \mathbf{R}$ be a function. The following two conditions are equivalent:*
(a) *The formal divergence equation has a solution, defined on V^r.*
(b) *The function f satisfies*

$$E_\sigma(f) = 0. \tag{12}$$

We give basic steps of the proof. Suppose that the formal divergence equation (2) has a solution $g = g^i$. Differentiating $d_i g^i$, we get the formulas

$$\frac{\partial d_i g^i}{\partial y^\sigma} = d_i \frac{\partial g^i}{\partial y^\sigma}, \tag{13}$$

and for every $k = 1, 2, \dots, r$,

$$\begin{aligned} \frac{\partial d_i g^i}{\partial y^\sigma_{i_1 i_2 \dots i_k}} &= d_i \frac{\partial g^i}{\partial y^\sigma_{i_1 i_2 \dots i_k}} \\ &+ \frac{1}{k} \left(\frac{\partial g^{i_1}}{\partial y^\sigma_{i_2 i_3 \dots i_k}} + \frac{\partial g^{i_2}}{\partial y^\sigma_{i_1 i_3 \dots i_k}} + \frac{\partial g^{i_3}}{\partial y^\sigma_{i_2 i_1 i_4 \dots i_k}} + \dots + \frac{\partial g^{i_k}}{\partial y^\sigma_{i_2 i_3 \dots i_{k-1} i_1}} \right). \end{aligned} \tag{14}$$

Using these formulas, we can compute the Euler-Lagrange expressions $E_\sigma(f) = E_\sigma(d_i g^i)$ in several steps. After $r - 1$ steps

$$E_\sigma(d_i g^i) = (-1)^r d_{i_1} d_{i_2} \dots d_{i_r d_i} \frac{\partial g^i}{\partial y^\sigma_{i_1 i_2 \dots i_r}}. \tag{15}$$

But since f is defined on V^r, the solution g of equation (2) necessarily satisfies

$$\begin{aligned} &\frac{\partial g^i}{\partial y^\sigma_{i_1 i_2 \dots i_r}} + \frac{\partial g^{i_1}}{\partial y^\sigma_{i i_2 i_3 \dots i_r}} + \frac{\partial g^{i_2}}{\partial y^\sigma_{i_1 i i_3 i_4 \dots i_r}} \\ &\quad + \dots + \frac{\partial g^{i_{r-1}}}{\partial y^\sigma_{i_1 i_2 \dots i_{r-2} i i_r}} + \frac{\partial g^{i_r}}{\partial y^\sigma_{i_1 i_2 \dots i_{r-1} i}} = 0. \end{aligned} \tag{16}$$

Applying this formula in (15) we see that condition (b) is satisfied.

Conversely, suppose that condition (b) is satisfied. We want to show that there exist functions $g^i : V^r \to \mathbf{R}$ such that $d_i g^i = f$, or, in an explicit form,

$$\frac{\partial g^i}{\partial x^i} + \frac{\partial g^{j_1}}{\partial y^\sigma} y^\sigma_{j_1} + \frac{\partial g^{j_2}}{\partial y^\sigma_{j_1}} y^\sigma_{j_1 j_2} + \dots + \frac{\partial g^{j_r}}{\partial y^\sigma_{j_1 j_2 \dots j_{r-1}}} y^\sigma_{j_1 j_2 \dots j_{r-1} j_r} = f. \tag{17}$$

As a consequence of (12), these functions satisfy

$$\begin{aligned} &\frac{\partial g^{i_1}}{\partial y^\sigma_{i_2 i_3 \dots i_{r+1}}} + \frac{\partial g^{i_2}}{\partial y^\sigma_{i_1 i_3 i_4 \dots i_{r+1}}} + \frac{\partial g^{i_3}}{\partial y^\sigma_{i_2 i_1 i_4 i_5 \dots i_{r+1}}} \\ &\quad + \dots + \frac{\partial g^{i_r}}{\partial y^\sigma_{i_2 i_3 \dots i_{r-1} i_1 i_{r+1}}} + \frac{\partial g^{i_{r+1}}}{\partial y^\sigma_{i_2 i_3 \dots i_{r-1} i_r i_1}} = 0. \end{aligned} \tag{18}$$

Let I be the fibred homotopy operator for differential forms on V^{2r}, associated with the projection $\pi : V \to U$. We have

$$\Theta_f = Id\Theta_f + dI\Theta_f + \Theta_0 = Ip_1 d\Theta_f + Ip_2 d\Theta_f + dI\Theta_f + \Theta_0, \tag{19}$$

where Θ_0 is an n-form, projectable on U. In this formula, $p_1 d\Theta_f = 0$ by hypothesis, and we may suppose that $\Theta_0 = d\vartheta_0$ (on U). Moreover $h\Theta_f = hd(I\Theta_f + \vartheta_0) f\omega_0$. Defining functions g^i on V^s, where $s \leq 2r$, by the condition $h(I\Theta_f + \vartheta_0) = g^i\omega_i$, we see we have constructed a solution of the formal divergence equation $d_i g^i = f$. Explicitly,

$$\frac{\partial g^i}{\partial x^i} + \frac{\partial g^{j_1}}{\partial y^\sigma} y^\sigma_{j_1} + \frac{\partial g^{j_2}}{\partial y^\sigma_{j_1}} y^\sigma_{j_1 j_2} + \ldots + \frac{\partial g^{j_{s+1}}}{\partial y^\sigma_{j_1 j_2 \ldots j_s}} y^\sigma_{j_1 j_2 \ldots j_s j_{s+1}} = f. \tag{20}$$

Note, however, that in general, we have not yet proved that the formal divergence equation has a solution defined on V^r. If $s \leq r$, formula (20) shows that condition (a) holds. If $s \geq r+1$, we apply Lemma 4 several times, and obtain a solution of equation (20) defined on V^r. This concludes the proof of assertion (b).

Condition (12) $E_\sigma(f) = 0$ is called the *integrability condition* for the formal divergence equation (2).

Combining Theorem 2 and Lemma 2, we can easily describe solutions of the formal divergence equations $d_i g^i = f$ as some differential forms.

Theorem 3 *Let $f : V^r \to \boldsymbol{R}$ be a function such that $E_\sigma(f) = 0$, let $g = g^i$ be a system of functions, defined on V^r, and let $\eta = g^i\omega_i$. Then the following conditions are equivalent:*

(a) *The system $g = g^i$ is a solution of the formal divergence equation (2).*

(b) *There exists a projectable extension μ of the form η such that*

$$hd\mu = f\omega_0. \tag{21}$$

5.3 The kernel of the Euler-Lagrange mapping

Consider the Euler-Lagrange mapping $\Omega^r_{n,X}W \ni \lambda \to E_\lambda \in \Omega^{2r}_{n+1,Y}W$. The domain and the range of this mapping have the structure of Abelian groups (and real vector spaces), and the Euler-Lagrange mapping is a morphism of these Abelian groups. In this subsection we characterize the *kernel* of this morphism. We shall say that a Lagrangian $\lambda \in \Omega^r_{n,X}W$ is *null*, or *trivial*, if $E_\lambda = 0$.

Theorem 4 *Let $\lambda \in \Omega^r_{n,X}W$ be a Lagrangian of order r. The following three conditions are equivalent:*

(a) *λ is trivial.*

(b) *For every point $y \in W$ there exists a fibred chart (V, ψ), $\psi = (x^i, y^\sigma)$, at y and an $(n-1)$-form $\mu_y \in \Omega^{r-1}_{n-1}V$ such that*

$$\lambda = hd\mu_y. \tag{1}$$

(c) *For every point $y \in W$ there exists a fibred chart (V, ψ), $\psi = (x^i, y^\sigma)$, at y with the following properties: If $\lambda = \mathcal{L}\omega_0$ on V^r in this chart, then there exist functions $f^i : V^r \to \boldsymbol{R}$ such that*

$$\frac{\partial f^i}{\partial y^\sigma_{j_1 j_2 \ldots j_r}} + \frac{\partial f^{j_1}}{\partial y^\sigma_{i j_2 j_3 \ldots j_r}} + \frac{\partial f^{j_2}}{\partial y^\sigma_{j_1 i j_3 \ldots j_r}} + \ldots + \frac{\partial f^{j_r}}{\partial y^\sigma_{j_1 j_2 \ldots j_{r-1} i}} = 0, \tag{2}$$

and

$$\mathcal{L} = d_i f^i. \tag{3}$$

The following are main steps of the proof.

1. Suppose that we have a trivial Lagrangian λ; then $E_\lambda = 0$. If in a fibred chart (V, ψ), $\psi = (x^i, y^\sigma)$, at a point $y \in W$, $\lambda = \mathcal{L}\omega_0$, then by hypothesis, the Euler-Lagrange expressions $E_\sigma(\mathcal{L}) = 0$ vanish. Consequently, by Theorem 3, $\lambda = hd\mu_y$ for some $(n-1)$-form μ_y on V^{r-1}, proving formula (1).

2. Let y be a point of W, let (V, ψ), $\psi = (x^i, y^\sigma)$, be a fibred chart at y, and let $\mu_y \in \Omega^{r-1}_{n-1} V$ be a form satisfying condition (1). Let $\lambda = \mathcal{L}\omega_0$ in this chart. Using the first canonical decomposition of forms, we can easily derive the formula

$$\begin{aligned}&(\pi^{r+2,r+1})^*(hd\mu_y + p_1 d\mu_y + p_2 d\mu_y + \ldots + p_{n-1} d\mu_y + p_n d\mu_y)\\ &\quad = (\pi^{r+2,r+1})^*(dh\mu_y + dp_1\mu_y + dp_2\mu_y + \ldots + dp_{n-1}\mu_y).\end{aligned} \tag{4}$$

In particular, comparing horizontal components, $(\pi^{r+2,r+1})^* hd\mu_y = hdh\mu_y$. By hypothesis, the left-hand side is equal to $\mathcal{L}\omega_0$, and writing $h\mu_y$ as $f^i\omega_i$, the right-hand side is $d_i f^i \omega_0$. Consequently, the Lagrange function $\mathcal{L}$ can be written as, $\mathcal{L} = d_i f^i$, and by projectability,

$$\frac{\partial f^i}{\partial y^\sigma_{j_1 j_2 \ldots j_r}} + \frac{\partial f^{j_1}}{\partial y^\sigma_{i j_2 j_3 \ldots j_r}} + \frac{\partial f^{j_2}}{\partial y^\sigma_{j_1 i j_3 \ldots j_r}} + \ldots + \frac{\partial f^{j_r}}{\partial y^\sigma_{j_1 j_2 \ldots j_{r-1} i}} = 0. \tag{5}$$

3. Suppose that condition (c) is satisfied; then the formal divergence equation (3) has a solution, defined on V^r, and by Theorem 2, $E_\sigma(\mathcal{L}) = 0$. Thus, the Lagrangian λ is trivial.

For Lagrangians of the first order we have a stronger result.

Corollary 1 *A first order Lagrangian $\lambda \in \Omega^1_{n,X} W$ is trivial if and only if there exists an n-form $\eta \in \Omega^0_{n-1} W$ such that*

$$\lambda = h\eta \tag{6}$$

and

$$d\eta = 0. \tag{7}$$

Indeed, we set $\eta_y = d\mu_y$ in Theorem 4. Then $h\eta_y = \lambda$, and since the horizontalization mapping h is in this case injective, there exists a form $\eta \in \Omega^0_{n-1} W$ such that the restriction of η to the domain of definition of η_y coincides with η_y. On this domain, $d\eta = d\eta_y = dd\eta_y = 0$.

5.4 The image of the Euler-Lagrange mapping

A 1-contact, $\pi^{s,0}$-horizontal form $\varepsilon \in \Omega^s_{n+1,Y} W$ is called a *source form* (Takens [90]). From the definition it follows that in a fibred chart (V, ψ), $\psi = (x^i, y^\sigma)$, ε has an expression

$$\varepsilon = \varepsilon_\sigma \omega^\sigma \wedge \omega_0, \tag{1}$$

where the components depend on the jet coordinates x^i, y^σ, $y^\sigma_{j_1}$, $y^\sigma_{j_1 j_2}$, $\dots$, $y^\sigma_{j_1 j_2 \dots j_s}$. We say that a source form ε is *variational*, if $\varepsilon = E_\lambda$ for some Lagrangian $\lambda \in \Omega^r_{n,X} W$. ε is said to be *locally variational*, if there are an open covering $\{V_\iota\}_{\iota \in I}$ of Y and a family $\{\lambda_\iota\}_{\iota \in I}$ of Lagrangians $\lambda_\iota \in \Omega^r_{n,X} V_\iota$ such that for every $\iota \in I$,

$$\varepsilon|_{V_\iota} = \lambda_\iota. \tag{2}$$

We wish to study the image set $\Omega^s_{n+1,Y} W$ of the Euler-Lagrange mapping $\Omega^r_{n,X} W \ni \lambda \to E_\lambda \in \Omega^{2r}_{n+1,Y} W$, consisting of variational forms, and a larger subset of the set $\Omega^s_{n+1,Y} W$, consisting of locally variational forms.

Theorem 5 *A source form $\varepsilon \in \Omega^s_{n+1,Y} W$ is locally variational if and only if there exists a form $F \in \Omega^s_{n+1} W$ of order of contactness ≥ 2 such that $d(\varepsilon + F) = 0$.*

Indeed, if $\varepsilon = E_\lambda$ for some Lagrangian λ, we choose a Lepage equivalent ρ of λ and define F to be $p_2 d\rho + p_2 d\rho + \dots + p_{n+1} d\rho$. Then by the first canonical decomposition, $(\pi^{r+1,r})^* d\rho = E_\lambda + F$. Conversely, if $d(\varepsilon + F) = 0$, then every form ρ such that $\varepsilon + F = d\rho$, is a Lepage form. Then $\varepsilon = p_1 d\rho$, so ε is a locally variational form whose Lagrangian is $h\rho$.

One can derive from Theorem 5 another criterion of local variationality of in fibred charts.

Let ε be a source form, defined on W^s, and let (V, ψ), $\psi = (x^i, y^\sigma)$, be a fibred chart on Y, such that $V \subset W$, and the set $\psi(W)$ is star-shaped. Denote by I the corresponding fibred homotopy operator. Then $I\varepsilon$ is a π^s-horizontal form on V^s, that is, a Lagrangian of order s for Y. We denote $\lambda_\varepsilon = I\varepsilon$, and call λ_ε the *Veinberg-Tonti Lagrangian*, associated with the source form ε.

Recall that $I\varepsilon$ is defined by the fibred homotopy $\chi_s : [0,1] \times V^s \to V^s$, where $\chi_s(t, (x^i, y^\sigma, y^\sigma_{j_1}, y^\sigma_{j_1 j_2}, \dots, y^\sigma_{j_1 j_2 \dots j_s})) = (x^i, ty^\sigma, ty^\sigma_{j_1}, ty^\sigma_{j_1 j_2}, \dots, ty^\sigma_{j_1 j_2 \dots j_s})$. Since the fibred homotopy satisfies $\chi_s^* \varepsilon = (\varepsilon_\sigma \circ \chi_s)(t\omega^\sigma + y^\sigma dt) \wedge \omega_0$, we have, integrating the coefficient in this expression at dt,

$$\lambda_\varepsilon = \mathcal{L}_\varepsilon \omega_0, \tag{3}$$

where

$$\mathcal{L}_\varepsilon = y^\sigma \int \varepsilon_\sigma \circ \chi_s \cdot dt. \tag{4}$$

We find the chart expressions for the principal Lepage equivalent $\Theta_{\lambda_\varepsilon}$, and for the Euler-Lagrange form E_ε of the Veinberg-Tonti Lagrangian λ_ε; these forms are defined by

$$\Theta_\varepsilon = \mathcal{L}_\varepsilon \omega_0 + \sum_{k=0}^{s-1} f_\sigma^{j_1 j_2 \dots j_k i} \omega^\sigma_{j_1 j_2 \dots j_k} \wedge \omega_i, \tag{5}$$

where for every $k = 0, 1, 2, \dots, s-1$,

$$f_\sigma^{j_1 j_2 \dots j_k i} = \sum_{l=0}^{s-k-1} (-1)^l d_{p_1} d_{p_2} \dots d_{p_l} \frac{\partial \mathcal{L}_\varepsilon}{\partial y^\sigma_{j_1 j_2 \dots j_k p_1 p_2 \dots p_l i}}, \tag{6}$$

and

$$E_\varepsilon = E_\sigma(\mathcal{L}_\varepsilon) \omega^\sigma \wedge \omega_0, \tag{7}$$

where

$$E_\sigma(\mathcal{L}_\varepsilon) = \sum_{l=0}^{s} (-1)^l d_{p_1} d_{p_2} \dots d_{p_l} \frac{\partial \mathcal{L}_\varepsilon}{\partial y^\sigma_{p_1 p_2 \dots p_l}}. \tag{8}$$

We need two formulas for the formal derivative operator d_i. Note a specific summation convention adapted in the following Lemma.

Lemma 5 (a) *For every function f on V^p*

$$d_i(f \circ \chi_p) = d_i f \circ \chi_{p+1}. \tag{9}$$

(b) *For every function f on V^s and a collection of functions $g^{p_1 p_2 \dots p_k}$ on V^s, symmetric in all superscripts,*

$$\begin{aligned} & d_{p_1} d_{p_2} \dots d_{p_k} (f \cdot g^{p_1 p_2 \dots p_k}) \\ & \quad = \sum_{i=0}^{s} \binom{k}{i} d_{p_1} d_{p_2} \dots d_{p_i} f \cdot d_{p_{i+1}} d_{p_{i+2}} \dots d_{p_k} g^{p_1 p_2 \dots p_i p_{i+1} p_{i+2} \dots p_k}. \end{aligned} \tag{10}$$

Assertion (a) is an easy consequence of definitions, and formula (10) can be obtained on induction.

Lemma 6 *The Euler-Lagrange expressions of the Veinberg-Tonti Lagrangian λ_ε of a source form $\varepsilon = \varepsilon_\sigma \omega^\sigma \wedge \omega_0$ are*

$$E_\sigma(\mathcal{L}_\varepsilon) = \varepsilon_\sigma - \sum_{k=0}^{s} y^\nu_{q_1 q_2 \dots q_k} \int H^{q_1 q_2 \dots q_k}_{\sigma\nu}(\varepsilon) \circ \chi_{2s} \cdot t dt, \tag{11}$$

where

$$\begin{aligned} H^{q_1 q_2 \dots q_k}_{\sigma\nu}(\varepsilon) &= \frac{\partial \varepsilon_\sigma}{\partial y^\nu_{q_1 q_2 \dots q_k}} - (-1)^k \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{q_1 q_2 \dots q_k}} \\ & - \sum_{l=k+1}^{s} (-1)^l \binom{l}{k} d_{p_{k+1}} d_{p_{k+2}} \dots d_{p_l} \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{q_1 q_2 \dots q_k p_{k+1} p_{k+2} \dots p_l}}. \end{aligned} \tag{12}$$

To prove Lemma 6, we find a formula for the difference $\varepsilon_\sigma - E_\sigma(\mathcal{L}_\varepsilon)$. Consider the Euler-Lagrange form (7). We have

$$\frac{\partial \mathcal{L}_\varepsilon}{\partial y^\sigma} = \int \varepsilon_\sigma \circ \chi_s \cdot dt + y^\nu \int \frac{\partial \varepsilon_\nu}{\partial y^\sigma} \circ \chi_s \cdot t dt, \tag{13}$$

and, by Lemma 5, for every l, $1 \le l \le s$,

$$\begin{aligned} d_{p_l} \dots d_{p_2} d_{p_1} \frac{\partial \mathcal{L}_\varepsilon}{\partial y^\nu_{p_1 p_2 \dots p_l}} &= d_{p_l} \dots d_{p_2} d_{p_1} \left(y^\nu \int \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{p_1 p_2 \dots p_l}} \circ \chi_s \cdot t dt \right) \\ &= \sum_{i=0}^{l} \binom{l}{i} y^\nu_{p_1 p_2 \dots p_i} \cdot \int d_{p_{i+1}} d_{p_{i+2}} \dots d_{p_l} \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{p_1 p_2 \dots p_i p_{i+1} p_{i+2} \dots p_l}} \circ \chi_{s+l-i} \cdot t dt. \end{aligned}$$

(14)

Then by (7), (13), and (14),

$$\begin{aligned} E_\sigma(\mathcal{L}_\varepsilon) &= \int \varepsilon_\sigma \circ \chi_s \cdot dt + y^\nu \int \frac{\partial \varepsilon_\nu}{\partial y^\sigma} \circ \chi_s \cdot tdt \\ &+ \sum_{l=1}^{s} (-1)^l \sum_{i=0}^{l} \binom{l}{i} y^\nu_{p_1 p_2 \dots p_i} \\ &\cdot \int d_{p_{i+1}} d_{p_{i+2}} \dots d_{p_l} \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{p_1 p_2 \dots p_i p_{i+1} p_{i+2} \dots p_l}} \circ \chi_{s+l-i} \cdot tdt. \end{aligned} \tag{15}$$

On the other hand, ε_σ can be expressed as

$$\varepsilon_\sigma = \sum_{i=0}^{l} \int \frac{\partial \varepsilon_\sigma}{\partial y^\nu_{p_1 p_2 \dots p_i}} \circ \chi_s \cdot y^\nu_{p_1 p_2 \dots p_i} \cdot tdt + \int \varepsilon_\sigma \circ \chi_s \cdot dt, \tag{16}$$

hence

$$\begin{aligned} \varepsilon_\sigma - E_\sigma(\mathcal{L}_\varepsilon) &= \int \frac{\partial \varepsilon_\sigma}{\partial y^\nu} \circ \chi_s \cdot y^\nu \cdot tdt - y^\nu \int \frac{\partial \varepsilon_\nu}{\partial y^\sigma} \circ \chi_s \cdot tdt \\ &- \sum_{l=1}^{s} (-1)^l \binom{l}{0} y^\nu \cdot \int d_{p_l} \dots d_{p_2} d_{p_1} \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{p_1 p_2 \dots p_l}} \circ \chi_{s+l} \cdot tdt \\ &+ \sum_{i=1}^{s} \int \frac{\partial \varepsilon_\sigma}{\partial y^\nu_{p_1 p_2 \dots p_i}} \circ \chi_s \cdot y^\nu_{p_1 p_2 \dots p_i} \cdot tdt \\ &- \sum_{l=1}^{s} (-1)^l \sum_{i=1}^{l} \binom{l}{i} y^\nu_{p_1 p_2 \dots p_i} \\ &\cdot \int d_{p_{i+1}} d_{p_{i+2}} \dots d_{p_l} \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{p_1 p_2 \dots p_i p_{i+1} p_{i+2} \dots p_l}} \circ \chi_{s+l-i} \cdot tdt. \end{aligned} \tag{17}$$

Changing summations, the double sum becomes

$$\begin{aligned} &\sum_{l=1}^{s} (-1)^l \sum_{i=1}^{l} \binom{l}{i} y^\nu_{p_1 p_2 \dots p_i} \\ &\quad \cdot \int d_{p_{i+1}} d_{p_{i+2}} \dots d_{p_l} \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{p_1 p_2 \dots p_i p_{i+1} p_{i+2} \dots p_l}} \circ \chi_{s+l-i} \cdot tdt \\ &\quad = \sum_{i=1}^{s} (-1)^i y^\nu_{p_1 p_2 \dots p_i} \cdot \int \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{p_1 p_2 \dots p_i}} \circ \chi_s \cdot tdt \\ &\quad + \sum_{i=1}^{s} (-1)^l \sum_{l=i+1}^{s} \binom{l}{i} y^\nu_{p_1 p_2 \dots p_i} \\ &\quad \cdot \int d_{p_{i+1}} d_{p_{i+2}} \dots d_{p_l} \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{p_1 p_2 \dots p_i p_{i+1} p_{i+2} \dots p_l}} \circ \chi_{s+l-i} \cdot tdt, \end{aligned} \tag{18}$$

and returning to (17) we get formula (11) of Lemma 6.

We call the functions (12) the *Helmholtz expressions*, associated with the source form ε. These functions appeared for the first time in Aldersley [1].

Remark 1 If, for example, $s = 3$, we get

$$H_{\sigma\nu}(\varepsilon) = \frac{\partial \varepsilon_\sigma}{\partial y^\nu} - \frac{\partial \varepsilon_\nu}{\partial y^\sigma} + d_{p_1} \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{p_1}} - d_{p_1} d_{p_2} \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{p_1 p_2}} + d_{p_1} d_{p_2} d_{p_3} \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{p_1 p_2 p_3}},$$

$$H^{p_1}_{\sigma\nu}(\varepsilon) = \frac{\partial \varepsilon_\sigma}{\partial y^\nu_{p_1}} + \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{p_1}} - 2 d_{p_2} \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{p_1 p_2}} + 3 d_{p_2} d_{p_3} \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{p_1 p_2 p_3}},$$

$$H^{p_1 p_2}_{\sigma\nu}(\varepsilon) = \frac{\partial \varepsilon_\sigma}{\partial y^\nu_{p_1 p_2}} - \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{p_1 p_2}} + 3 d_{p_3} \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{p_1 p_2 p_3}},$$

$$H^{p_1 p_2 p_3}_{\sigma\nu}(\varepsilon) = \frac{\partial \varepsilon_\sigma}{\partial y^\nu_{p_1 p_2 p_3}} + \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{p_1 p_2 p_3}}.$$

Suppose we have a *locally variational* source form ε. Then for every point y of Y there exist a fibred chart (V, ψ), $\psi = (x^i, y^\sigma)$, at y, an integer r and a Lagrangian λ of order r on V^r, $\lambda = \mathcal{L}\omega_0$, such that

$$\varepsilon_\sigma = \sum_{l=0}^{r} (-1)^l d_{p_1} d_{p_2} \dots d_{p_l} \frac{\partial \mathcal{L}}{\partial y^\sigma_{p_1 p_2 \dots p_l}}. \tag{19}$$

The source form ε can be considered as defined on W^{2r}, and the Helmholtz expressions $H^{p_1 p_2 \dots p_k}_{\sigma\nu}(\varepsilon)$ (12) for $s = 2r$ are

$$\begin{aligned} H^{q_1 q_2 \dots q_k}_{\sigma\nu}(\varepsilon) &= \frac{\partial \varepsilon_\sigma}{\partial y^\nu_{q_1 q_2 \dots q_k}} - (-1)^k \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{q_1 q_2 \dots q_k}} \\ &\quad - \sum_{l=k+1}^{2r} (-1)^l \binom{l}{k} d_{p_{k+1}} d_{p_{k+2}} \dots d_{p_l} \frac{\partial \varepsilon_\nu}{\partial y^\sigma_{q_1 q_2 \dots q_k p_{k+1} p_{k+2} \dots p_l}}. \end{aligned} \tag{20}$$

where $k = 0, 1, 2, \dots, 2r$.

Lemma 7 *Every locally variational source form* $\varepsilon = \varepsilon_\sigma \omega^\sigma \wedge \omega_0$ *on* W^{2r} *satisfies*

$$H^{q_1 q_2 \dots q_k}_{\sigma\nu}(\varepsilon) = 0 \tag{21}$$

for all $k = 0, 1, 2, \dots, 2r$.

For the proof, we can apply Theorem 5 (Krupková [66]). We can also use Veinberg-Tonti Lagrangians, and Theorem 1 on the properties of the Euler-Lagrange mapping with respect to automorphisms of the underlying fibred manifold (Krupka, Anderson and Duchamp [2], Krupka [45]). Some other proofs are provided by the variational sequence theory (Krbek and Musilová [39], Krupka [55]).

Remark 2 Suppose that $s = 3$ (cf. Remark 1). Then conditions (21)

$$H_{\sigma\nu}(\varepsilon) = 0, \quad H^{p_1}_{\sigma\nu}(\varepsilon) = 0, \quad H^{p_1 p_2}_{\sigma\nu}(\varepsilon) = 0, \quad H^{p_1 p_2 p_3}_{\sigma\nu}(\varepsilon) = 0$$

can be verified, using (19), by a straightforward computation.

We set, using the Helmholtz expressions (12),

$$H_\varepsilon = \frac{1}{2}\sum_{i=0}^{s} H_{\nu\sigma}^{j_1 j_2 \dots j_i}(\varepsilon)\omega^\sigma_{j_1 j_2 \dots j_i} \wedge \omega^\nu \wedge \omega_0. \tag{22}$$

H_ε is the (global) *Helmholtz form* (Anderson Duchamp [2], Krupka [45], [48], Krbek and Musilová [39]).

Summarizing our discussion, we have the following result.

Theorem 6 *A source form ε is locally variational if and only if $H_\varepsilon = 0$.*

Sufficiency follows from Lemma 6, and necessity from Lemma 7.

6 Invariant variational principles

Let X be any manifold, W an open set in X, and let $\alpha : W \to Y$ be a smooth mappings. Recall that a differential form η, defined on a neighbourhood of the set $\alpha(W)$ in X, is said to be *invariant* with respect to α, if $\alpha^*\eta = \eta$ on the set $W \cap \alpha(W)$; we also say in this case that α is an *invariance transformation* of η. A vector field, whose local one-parameter group consists of invariance transformations of η, is said to be a *generator* of invariance transformations of η. In this section we apply these definitions to local automorphisms of a fibred manifold Y. We study properties of integral variational functionals on Y, whose Lagrangians, or Euler-Lagrange forms, are invariant with respect to one-parameter families of local automorphisms.

6.1 Invariant variational functionals

Let λ be a Lagrangian of order r for Y, let $\alpha : W \to Y$ be a local automorphism of Y, and let $J^r\alpha : W^r \to J^rY$ be the r-jet prolongation of α. We say that α is an *invariance transformation* of the Lagrangian λ if $J^r\alpha^*\lambda = \lambda$. A *generator* of invariance transformations of λ is a π-projectable vector field on Y whose local one-parameter group consists of invariance transformations of λ. A variational functional, whose Lagrangian is invariant with respect to α, is said to be invariant with respect to α.

Lemma 1 *Let λ be a Lagrangian of order r for Y.*

(a) *A π-projectable vector field Ξ on Y generates invariance transformations of λ if and only if*

$$\partial_{J^r\Xi}\lambda = 0. \tag{1}$$

(b) *Generators of invariance transformations of λ constitute a subalgebra of the algebra of vector fields on J^rY.*

Equation (1), called the *Noether equation*, represents a relation between λ and the vector field Ξ. Given λ, we can use this equation to determine generators of invariance transformations. Conversely, given a collection of π-projectable vector fields Ξ, one can apply the corresponding Noether equations to determine invariant Lagrangians λ.

The following is known as the (first) *theorem of Emmy Noether*.

Theorem 1 *Let λ be a Lagrangian, ρ a Lepage equivalent of λ, defined on J^sY, and let γ be an extremal. Then for every generator Ξ of invariance transformations of λ,*

$$dJ^s\gamma^* i_{J^s\Xi}\rho = 0. \tag{2}$$

The proof is based on the first variation formula, and is trivial. Indeed, we have

$$J^r\gamma^*\partial_{J^r\Xi}\lambda = J^s\gamma^* i_{J^s\Xi}d\rho + dJ^s\gamma^* i_{J^s\Xi}\rho, \tag{3}$$

and since the left-hand side vanishes, by invariance, and the first summand on the righ-hand side also vanishes, because γ is an extremal, we get (2) as required.

Note that (global) condition (2) can also be written in a different way, by means of locally defined principal Lepage equivalents Θ_λ of the Lagrangian λ. From the structure theorem on Lepage forms we know that, locally, $\rho = \Theta_\lambda + d\nu + \mu$, where ν is a contact form, and μ is a contact form of order of contactness ≥ 2. Then $dJ^s\gamma^* i_{J^s\Xi}\rho = dJ^s\gamma^*(i_{J^s\Xi}\Theta_\lambda + i_{J^s\Xi}d\nu + i_{J^s\Xi}\mu)$. But the form $i_{J^s\Xi}\mu$ is contact; moreover, $i_{J^s\Xi}d\nu = \partial_{J^s\Xi}\nu - di_{J^s\Xi}\nu$, from which we deduce that

$$J^s\gamma^* i_{J^s\Xi}\mu = 0, \quad dJ^s\gamma^* i_{J^s\Xi}d\nu = dJ^s\gamma^*\partial_{J^s\Xi}\nu - dJ^s\gamma^* di_{J^s\Xi}\nu = 0. \tag{4}$$

Consequently, under the hypothesis of Theorem 1, condition

$$dJ^s\gamma^* i_{J^s\Xi}\Theta_\lambda = 0 \tag{5}$$

is satisfied over coordinate neighbourhoods of fibred charts on Y.

6.2 Invariant Euler-Lagrange forms

Let $\alpha : W \to Y$ be a local automorphism of Y, and let ε be a source form on J^sY. We say that α is an *invariance transformation* of ε, if $J^s\alpha^*\varepsilon = \varepsilon$. A *generator* of invariance transformations of ε is a π-projectable vector field on Y whose local one-parameter group consists of invariance transformations of ε.

Lemma 2 *Let ε be a source form of order s for Y.*

(a) *A π-projectable vector field Ξ on Y generates invariance transformations of ε if and only if*

$$\partial_{J^s\Xi}\varepsilon = 0. \tag{1}$$

(b) *Generators of invariance transformations of ε constitute a subalgebra of the algebra of vector fields on J^sY.*

Equation (1) is a geometric version of what is known as the *Noether-Bessel-Hagen equation* for *variational* source forms.

Let λ be a Lagrangian of order r for Y, and let E_λ be the Euler-Lagrange form of λ. Combining Lemma 2 with the identity

$$J^{2r}\alpha^* E_\lambda = E_{J^r\alpha^*\lambda}, \tag{2}$$

where α is any local automorphism of Y, we easily obtain the following assertion.

Lemma 3 *Let λ be a Lagrangian of order r.*

(a) *Every invariance transformation of λ is an invariance transformation of the Euler-Lagrange form E_λ.*

(b) *For every invariance transformation α of E_λ, the Lagrangian $\lambda - J^r\alpha^*\lambda$ is variationally trivial.*

We can generalize the Noether's theorem to invariance transformations of the Euler-Lagrange form. However, since the proof is based on the theorem on the kernel of the Euler-Lagrange mapping, the assertion we obtain is of local character. We denote by Θ_λ the principal Lepage equivalent of λ.

Theorem 2 *Let λ be a Lagrangian of order r, let γ be an extremal, and let Ξ be a generator of invariance transformations of the Euler-Lagrange form E_λ. Then for every point $y_0 \in Y$ there exists a fibred chart (V, ψ) at y_0 and an $(n-1)$-form η, defined on V^{r-1}, such that*

$$dJ^{2r-1}\gamma^*(i_{J^{2r-1}\Xi}\Theta_\lambda + \eta) = 0 \tag{3}$$

on $U = \pi(V)$.

Indeed, under the hypothesis of Theorem 2, from the formula $\partial_{J^{2r}\Xi}E_\lambda = E_{\partial_{J^r\Xi}\lambda}$ we obtain $E_{\partial_{J^r\Xi}\lambda} = 0$, thus, the Lagrangian $\partial_{J^r\Xi}\lambda$ belongs to the kernel of the Euler-Lagrange mapping. Thus, $\partial_{J^r\Xi}\lambda = hd\eta$ over sufficiently small open sets V in Y such that (V, ψ) is a fibred chart. Then, however, from the infinitesimal first variation formula over V,

$$J^r\gamma^*\partial_{J^r\Xi}\lambda = J^{2r-1}\gamma^* i_{J^{2r-1}\Xi}d\Theta_\lambda + dJ^{2r-1}\gamma^* i_{J^{2r-1}\Xi}\Theta_\lambda, \tag{4}$$

reduces to

$$J^r\gamma^* hd\eta = dJ^{2r-1}\gamma^* i_{J^{2r-1}\Xi}\Theta_\lambda. \tag{5}$$

Since $J^r\gamma^* hd\eta = J^r\gamma^* d\eta = dJ^r\gamma^*\eta$, this proves formula (3).

Remark 1 If $r = 1$ in Theorem 2, then the principal Lepage equivalent Θ_λ is globally well-defined. Moreover, it follows from the properties of the Euler-Lagrange mapping that the form η may be taken as a globally well-defined form on Y.

6.3 Jacobi vector fields

Let λ be a Lagrangian of order r for Y, and let γ be an extremal of λ; thus, we suppose that γ satisfies the Euler-Lagrange equations

$$E_\lambda \circ J^{2r}\gamma = 0. \tag{1}$$

Let $\alpha : W \to Y$ be a local automorphism of Y with projection α_0, and let $J^r\alpha : W^r \to J^rY$ be the r-jet prolongation of α. We say that α is a *symmetry* of γ, if the section $\alpha\gamma\alpha_0^{-1}$ is also a solution of the Euler-Lagrange equations, i.e.,

$$E_\lambda \circ J^{2r}(\alpha\gamma\alpha_0^{-1}) = 0. \tag{2}$$

We say that a π-projectable vector field *generates symmetries* of γ, if its local one-parameter group consists of symmetries of γ; we also say in this case that Ξ is a *Jacobi vector field* along γ.

The following lemma can be proved by means of differential-geometric operations with the Euler-Lagrange form.

Lemma 4 *An invariance transformation of the Euler-Lagrange form E_λ is a symmetry of every extremal γ.*

Let γ be any section of Y, let α be an invariance transformation. To prove Lemma 4, we need two simple observations. First, neglecting the details on the domains of definition, we have, for every point $J^s_x\gamma$, belonging to the domain $J^s\alpha$, $J^s\alpha(J^s_x\gamma) = J^s_{\alpha_0(x)}(\alpha\gamma\alpha_0^{-1})$. Then, $(J^s\alpha \circ J^s\gamma)(x) = (J^s(\alpha\gamma\alpha_0^{-1}) \circ \alpha_0)(x)$, and we have on the domain of $\alpha\gamma\alpha_0^{-1}$

$$J^s\alpha \circ J^s\gamma \circ \alpha_0^{-1} = J^s(\alpha\gamma\alpha_0^{-1}). \tag{3}$$

Second, let s be the order of the Euler-Lagrange form E_λ. Then for any π-projectable vector field Z,

$$J^s(\alpha\gamma\alpha_0^{-1})^* i_{J^s\mathrm{Z}} E_\lambda = (\alpha_0^{-1})^*(J^s\gamma)^*(J^s\alpha)^* i_{J^s\mathrm{Z}} E_\lambda. \tag{4}$$

But for every point $J^s_x\delta$ from the domain of the form (4) and all tangent vectors $\xi_1, \xi_2, \dots, \xi_n$ at this point,

$$\begin{aligned}
&(J^s\alpha)^* i_{J^s\mathrm{Z}} E_\lambda(J^s_x\delta)(\xi_1, \xi_2, \dots, \xi_n) \\
&\quad = E_\lambda(J^s\alpha(J^s_x\delta))(J^s\mathrm{Z}, TJ^s\alpha \cdot \xi_1, TJ^s\alpha \cdot \xi_2, \dots, TJ^s\alpha \cdot \xi_n) \\
&\quad = E_\lambda(J^s\alpha(J^s_x\delta))(TJ^s\alpha \cdot TJ^s\alpha^{-1} \cdot J^s\mathrm{Z}, TJ^s\alpha \cdot \xi_1, TJ^s\alpha \cdot \xi_2, \dots, TJ^s\alpha \cdot \xi_n) \\
&\quad = (J^s\alpha)^* E_\lambda(J^s_x\delta))(TJ^s\alpha^{-1} \cdot J^s\mathrm{Z}, \xi_1, \xi_2, \dots, \xi_n) \\
&\quad = i_{TJ^s\alpha^{-1}\cdot J^s\mathrm{Z}}(J^s\alpha)^* E_\lambda(J^s_x\delta))(\xi_1, \xi_2, \dots, \xi_n).
\end{aligned} \tag{5}$$

But $TJ^s\alpha^{-1} \cdot J^s\mathrm{Z} = J^s(T\alpha^{-1} \cdot \mathrm{Z} \circ \alpha)$, so we have

$$(J^s\alpha)^* i_{J^s\mathrm{Z}} E_\lambda = i_{J^s(T\alpha^{-1}\cdot\mathrm{Z}\circ\alpha)}(J^s\alpha)^* E_\lambda. \tag{6}$$

We use formulas (4) and (6) to prove Lemma 4. Let now γ be an extremal, and let $(J^s\alpha)^* E_\lambda = E_\lambda$. Then

$$\begin{aligned}
&J^s(\alpha\gamma\alpha_0^{-1})^* i_{J^s\mathrm{Z}} E_\lambda = (\alpha_0^{-1})^*(J^s\gamma)^*(J^s\alpha)^* i_{J^s\mathrm{Z}} E_\lambda \\
&\quad = (\alpha_0^{-1})^*(J^s\gamma)^* i_{J^s(T\alpha^{-1}\cdot\mathrm{Z}\circ\alpha)} E_\lambda = 0,
\end{aligned} \tag{7}$$

because γ is an extremal.

Theorem 3 *Let λ be a Lagrangian of order r, let s be the order of the Euler-Lagrange form E_λ, and let γ be an extremal. Then a π-projectable vector field Ξ generates symmetries of γ if and only if*

$$E_{\partial_{J^r\Xi}\lambda} \circ J^s\gamma = 0. \tag{8}$$

We prove necessity. Let x be a point, belonging to the domain of γ, let α_t be the local one-parameter group of Ξ, and let $\alpha_{0,t}$ be the projection of α_t. Choose some vectors $\xi_0, \xi_1, \xi_2, \dots, \xi_n$ at the point $J^s_x\gamma$. Then

$$\begin{aligned}
&E_{(J^r\alpha_t)^*\lambda}(J^s_x\gamma)(\xi_0, \xi_1, \xi_2, \dots, \xi_n) = (J^s\alpha_t)^* E_\lambda(J^s_x\gamma)(\xi_0, \xi_1, \xi_2, \dots, \xi_n) \\
&\quad = E_\lambda(J^s\alpha_t(J^s_x\gamma))(TJ^s\alpha_t \cdot \xi_0, TJ^s\alpha_t \cdot \xi_1, TJ^s\alpha_t \cdot \xi_2, \dots, TJ^s\alpha_t \cdot \xi_n).
\end{aligned} \tag{9}$$

By continuity, the point x has a neighbourhood U such that the real function

$$(t, x) \to E_\lambda(J^s\alpha_t(J^s_x\gamma))(TJ^s\alpha_t \cdot \xi_0, TJ^s\alpha_t \cdot \xi_1, TJ^s\alpha_t \cdot \xi_2, \ldots, TJ^s\alpha_t \cdot \xi_n) \quad (10)$$

is defined on the set $(-\varepsilon, \varepsilon) \times U$ for some $\varepsilon > 0$.

Suppose that Ξ generates symmetries of γ. Then $E_\lambda \circ J^s\alpha_t \circ J^s\gamma \circ \alpha^{-1}_{0,t} = 0$ on $(-\varepsilon, \varepsilon) \times U$, that is, $E_\lambda \circ J^s\alpha_t \circ J^s\gamma = 0$. Consequently, since the restriction of the tangent mapping $TJ^s\alpha_t$ to the point $J^s_x\gamma$ is a linear isomorphism, we have $E_{(J^r\alpha_t)^*\lambda}(J^s_x\gamma)(\xi_0, \xi_1, \xi_2, \ldots, \xi_n) = 0$, i.e., $E_{(J^r\alpha_t)^*\lambda} \circ J^s\gamma = 0$, proving (8).

Conversely, from (8) we find that $E_{(J^r\alpha_t)^*\lambda}(J^s_x\gamma) = E_\lambda(J^s_x\gamma)$. But γ is an extremal, so we have $E_{(J^r\alpha_t)^*\lambda}(J^s_x\gamma) = 0$.

7 Remarks

We present in this last part of the work some complements and comments on general theory of integral variational functionals in fibred spaces.

7.1 Examples of Lepage forms, generalizations

The concept of a Lepage form unifies several known examples of forms, used in the first order calculus of variations of multiple integrals, with different properties; the most well-known are the Cartan form, the Poincaré-Cartan form, the Carathéodory form, and the so called *fundamental form* (see Betounes [5], Crampin and Saunders [14], Dedecker [15], García [19], Goldschmidt and Sternberg [23], Gotay [25], Horak and Kolář [33], Krupka [41], [42], [45], Olver [79], Rund [82], Saunders [86], Shadwick [87], Sniatycki [88]). For a second order Lepage form, which can be considered as a direct generalization of the Cartan form, we refer to Krupka [50]. Lepage equivalents of a fixed Lagrangian define the same variational functional, the Euler-Lagrange form, as well as the Noether currents; they lead to different Hamilton equations.

Some authors studied possibilities to extend properties of Lepage forms to differential forms of higher degree (Krupka and Sedenková [62], Krupková [65], [67], with the idea to use them in the theory the inverse problem of the calculus of variations, or for computation of classes in the variational sequence theory.

Let Y be a fibred manifold of dimension $n + m$, with n-dimensional base X and projection π. We discuss known examples of Lepage forms in the following three cases:

(a) $n = 1$, m and r arbitrary (*higher order fibred mechanics*),

(b) $r = 1$, n and m arbitrary (*first order field theory*),

(c) $r = 2$, n and m arbitrary (*second order field theory*).

(a) Higher order fibred mechanics For $n = 1$ we usually denote a fibred chart on Y by (V, ψ), $\psi = (t, q^\sigma)$, and the associated chart on J^rY by (V^r, ψ^r), $\psi^r = (t, q^\sigma_{(0)}, q^\sigma_{(1)}, \ldots, q^\sigma_{(r)})$, and $\omega^\sigma_{(k)} = dq^\sigma_{(k)} - q^\sigma_{(k+1)}dt$, $0 \le k \le r - 1$. For $r = 1$ we write $q^\sigma_{(0)} = q^\sigma$, $q^\sigma_{(1)} = \dot{q}^\sigma$, and $\omega^\sigma = \omega^\sigma_{(0)}$. The formal derivative with respect to t is denoted by d/dt.

Lemma 1 *Every Lagrangian* $\lambda \in \Omega^1_{1,X}W$ *has a unique Lepage equivalent* Θ_λ. *If* λ *is* *expressed in a fibred chart by* $\lambda = \mathcal{L}dt$*, then*

$$\Theta_\lambda = \mathcal{L}dt + \frac{\partial \mathcal{L}}{\partial \dot{q}^\sigma}\omega^\sigma. \tag{1}$$

The form Θ_λ is called the *Cartan equivalent* of λ, or just the *Cartan form*. Introducing the function

$$\mathcal{H} = -\mathcal{L} + \frac{\partial \mathcal{L}}{\partial \dot{q}^\sigma}\dot{q}^\sigma, \tag{2}$$

we can write Θ_λ in the *Hamiltonian form*

$$\Theta_\lambda = -\mathcal{H}dt + \frac{\partial \mathcal{L}}{\partial \dot{q}^\sigma}dq^\sigma. \tag{3}$$

The form Θ_λ (1) was first considered by Cartan; expression (3), with $\partial\mathcal{L}/\partial\dot{q}^\sigma$ replaced by independent coordinates p_σ, goes back to Whitaker.

Lemma 2 *Every Lagrangian* $\lambda \in \Omega^r_{1,X}W$ *has a unique Lepage equivalent* Θ_λ*. If* λ *is expressed in a fibred chart by* $\lambda = \mathcal{L}dt$*, then* Θ_λ *is defined by*

$$\Theta_\lambda = \mathcal{L}dt + \sum_{k=0}^{r-1}\left(\sum_{l=0}^{r-k-1}(-1)^l \frac{d^l}{dt^l}\frac{\partial \mathcal{L}}{\partial q^\sigma_{(k+1+l)}}\right)\omega^\sigma_{(k)}. \tag{4}$$

The Lepage equivalent Θ_λ *is of order* $2r-1$*.*

(b) First order field theory Now the positive integer $n = \dim X$ is arbitrary.

Lemma 3 *Every Lagrangian* $\lambda \in \Omega^1_{n,X}W$ *has a unique Lepage equivalent* $\Theta_\lambda \in \Omega^1_{n,Y}W$ *whose order of contactness is* ≤ 1*. If* λ *is expressed in a fibred chart by* $\lambda = \mathcal{L}\omega_0$*, then*

$$\Theta_\lambda = \mathcal{L}\omega_0 + \frac{\partial \mathcal{L}}{\partial y^\sigma_i}\omega^\sigma \wedge \omega_i. \tag{5}$$

The form Θ_λ (5) was considered by different authors (see e.g. Sniatycki, Goldschmidt and Sternberg, Krupka, García and Perez-Rendon; following García [19], we call Θ_λ the *Poincaré-Cartan equivalent* of the Lagrangian λ, or just the *Poincaré-Cartan form*.

Let ρ be an n-form on Y. Since $d\rho$ is defined on Y, ρ is always a Lepage form, and is a Lepage equivalent of the Lagrangian of order 1, $\lambda = h\rho \in \Omega^1_{n,X}W$. Since the horizontalization $\Omega^0_n W \ni \rho \to h\rho \in \Omega^1_{n,X}W$ is in this case an *injection*, for any Lagrangian λ in the *image* of h one can reconstruct the preimage of λ, which is indeed unique.

Lemma 4 *Let* ρ *be an* n*-form on* Y*, and let* $h\rho = \mathcal{L}\omega_0$*. Then* $(\pi^{1,0})^*\rho$ *can be written as*

$$(\pi^{1,0})^*\rho = \mathcal{L}\omega_0 + \sum_{k=0}^{n}\frac{1}{k!(n-k)!}\frac{1}{k!}\frac{\partial^k \mathcal{L}}{\partial y^{\sigma_1}_{j_1}\partial y^{\sigma_2}_{j_2}\dots\partial y^{\sigma_k}_{j_k}}\epsilon_{j_1 j_2\dots j_k i_{k+1} i_{k+2}\dots i_n} \tag{6}$$
$$\cdot\omega^{\sigma_1}\wedge\omega^{\sigma_2}\wedge\dots\wedge\omega^{\sigma_k}\wedge dx^{i_{k+1}}\wedge dx^{i_{k+2}}\wedge\dots\wedge dx^{i_n},$$

and is a Lepage equivalent of $h\rho$*.*

Formula (6) gives us an expression for the inverse mapping of the horizontalization $\Omega_n^0 W \ni \rho \to h\rho \in \Omega_{n,X}^1 W$. In particular, formula (6) shows that a Lagrangian $\lambda \in \Omega_{n,X}^r W$ may have a Lepage equivalent belonging to the module $\Omega_n^{r-1} W$. We extend the inverse mapping of the horizontalisation h, given by (6), to the whole set $\Omega_{n,X}^1 W$ of the first order Lagrangians. We define a Lepage equivalent Z_λ of any Lagrangian $\lambda = \mathcal{L}\omega_0$ by formula (6); Z_λ is defined on V^1, and is called the *fundamental Lepage equivalent* of the Lagrangian.

Since each term in the fundamental Lepage equivalent Z_λ (6) is invariant, restricting the summation to terms of order of contactness $\leq p$, we get again a Lepage equivalent of λ.

Distinguished properties of the fundamental Lepage equivalent Z_λ are summarized in the following Lemma. Let E_λ be the Euler-Lagrange form of λ.

Lemma 5 *The mapping $\lambda \to \mathrm{Z}_\lambda$ of $\Omega_{n,X}^1 W$ into $\Omega_n^1 W$ has the following properties:*

(a) *If $\rho \in \Omega_n^0 W$, then $\mathrm{Z}_{h\rho} = (\pi^{1,0})^* \rho$.*

(b) *Z_λ is $\pi^{1,0}$-projectable if and only if E_λ is $\pi^{2,1}$-projectable.*

(c) *Z_λ is closed if and only if $E_\lambda = 0$.*

(d) *For any automorphism $\alpha : W \to Y$, $J^1\alpha^* \mathrm{Z}_\lambda = \mathrm{Z}_{J^1\alpha^*\lambda}$.*

From Lemma 5 we easily deduce that for every π-projectable vector field ξ, the fundamental Lepage form Z_λ satisfies $\partial_{J^1\xi} \mathrm{Z}_\lambda = \mathrm{Z}_{\partial_{J^1\xi}\lambda}$, and the Poincaré-Cartan form Θ_λ satisfies $J^1\alpha^*\Theta_\lambda = \Theta_{J^1\alpha^*\lambda}$ and $\partial_{J^1\xi}\Theta_\lambda = \Theta_{\partial_{J^1\xi}\lambda}$; the Euler-Lagrange form E_λ satisfies $\partial_{J^2\xi} E_\lambda = E_{\partial_{J^1\xi}\lambda}$. Finally, for a first order Lagrangian $\lambda \in \Omega_{n,X}^1 W$, the Euler-Lagrange form E_λ vanishes if and only if there exists an n-form $\rho \in \Omega_n^0 W$ such that $\lambda = h\rho$ and $d\rho = 0$.

Another example is provided by the following. Let $\lambda \in \Omega_{n,X}^1 W$ be a nowhere zero first order Lagrangian. Then the n-form

$$\begin{aligned} \rho = \frac{1}{\mathcal{L}^{n-1}} &\left(\mathcal{L} dx^1 - \frac{\partial \mathcal{L}}{\partial y_1^{\sigma_1}} \omega^{\sigma_1} \right) \wedge \left(\mathcal{L} dx^2 - \frac{\partial \mathcal{L}}{\partial y_2^{\sigma_2}} \omega^{\sigma_2} \right) \\ &\wedge \ldots \wedge \left(\mathcal{L} dx^n - \frac{\partial \mathcal{L}}{\partial y_n^{\sigma_n}} \omega^{\sigma_n} \right), \end{aligned} \tag{7}$$

where $\mathcal{L}$ is defined by the chart expression $\lambda = \mathcal{L}\omega_0$, is a Lepage equivalent of λ. This form is called *Carathéodory form*.

(c) Lepage equivalents of second order Lagrangians It can be easily seen by checking invariance that for second order Lagrangians, the principal Lepage equivalent is globally well-defined. This proves the following lemma.

Lemma 6 *Let $W \subset Y$ be an open set, and let $\lambda \in \Omega_{n,X}^2 W$ be a Lagrangian of order 2. There exists a Lepage equivalent $\Theta_\lambda \in \Omega_n^{2r-1} W$ of λ such that for any fibred chart (V, ψ), $\psi = (x^i, y^\sigma)$, such that $V \subset W$, $\lambda = \mathcal{L}\omega_0$, and*

$$\Theta_\lambda = \mathcal{L}\omega_0 + \left(\frac{\partial \mathcal{L}}{\partial y_j^\sigma} - d_p \frac{\partial L}{\partial y_{pi}^\sigma} \right) \omega^\sigma \wedge \omega_j + \frac{\partial \mathcal{L}}{\partial y_{ij}^\sigma} \omega_i^\sigma \wedge \omega_j. \tag{8}$$

7.2 Variational sequence

The variational sequence is a main tool, which can serve to discover new information about the local structure of different variational concepts and constructions, and to characterize differences between local and global properties of these concepts. A typical problem which should be considered in this context, is the structure of variationality conditions (Helmholtz form), and the *local* and *global* inverse problems of the calculus of variations.

Research in this direction originated from different sources; one of the most significant was, in the author's opinion, the work of Th. Lepage, indicating that there should be a close correspondence between the exterior derivative of differential forms on one side, and the Euler-Lagrange mapping of the calculus of variations on the other side. A conclusion, derived from this approach, was the theory of Lepage forms (Krupka [41], [50]); in particular, it was shown that the Euler-Lagrange mapping can be globally interpreted as an assignment, sending an n-form (the Lagrangian) to an $(n+1)$-form (the Euler-Lagrange form). These studies, supported by the ideas of the variational bicomplex theory (cf. Anderson and Duchamp [2], Dedecker and Tulczyjew [16], Takens [90], Tulczyjew [94], Vinogradov [95], Vinogradov, Krasilschik and Lychagin [96]), gave rise to the concept of a finite order (cohomological) exact sequence of forms, the *variational sequence*, in which the Euler-Lagrange mapping is included as one arrow (Krupka [55]). We are not concerned with the variational bicomplex theory here (for this topic see Vitolo [98]).

For further results in this field we refer to Anderson and Thompson [4], Brajerčík and Krupka [6], [8], Dedecker and Tulczyjew [16], Francaviglia, Palese and Vitolo [18], Grassi [27], Grigore [29], Kolář and Vitolo [37], Krbek and Musilová ([39], [40]), Krupka [45], [48], [54], Krupka and Sedenková [62], Krupka, Krupková, Prince, Sarlet [58], [59], Krupková [68], Musilová [76], Pommaret [80], Stefanek [89], and Vitolo [97].

Let $\Omega^r_{0,c} = \{0\}$, and let $\Omega^r_{k,c}$ be the sheaf of *contact* k-forms on J^rY. We set

$$\Theta^r_k = \Omega^r_{k,c} + d\Omega^r_{k-1,c}, \tag{1}$$

where $d\Omega^r_{k-1,c}$ is the image sheaf of $\Omega^r_{k-1,c}$ by the exterior derivative d. It can be shown that we get an exact sequence of soft sheaves $0 \to \Theta^r_1 \to \Theta^r_2 \to \Theta^r_3 \to \ldots$, where the morphisms are the exterior derivative, i.e., a subsequence of the *De Rham sequence* $0 \to \mathbf{R} \to \Omega^r_0 \to \Omega^r_1 \to \Omega^r_2 \to \Omega^r_3 \to \ldots$. The quotient sequence

$$0 \to \mathbf{R} \to \Omega^r_0 \to \Omega^r_1/\Theta^r_1 \to \Omega^r_2/\Theta^r_2 \to \Omega^r_3/\Theta^r_3 \to \ldots \tag{2}$$

which is also exact, is called the *r-th order variational sequence* on Y. We denote the quotient mappings in (2) by $E_k : \Omega^r_k/\Theta^r_k \to \Omega^r_{k+1}/\Theta^r_{k+1}$. The following is a basic property of this sequence.

Theorem 1 *The variational sequence is an acyclic resolution of the constant sheaf* $\mathbf{R}$ *over* Y.

Denote (2) symbolically as $0 \to \mathbf{R} \to \mathcal{V}^r$. Let $\Gamma(Y, \mathcal{V}^r)$ be the cochain complex $0 \to \Gamma(Y, \mathbf{R}) \to \Gamma(Y, \Omega^r_0) \to \Gamma(Y, \Omega^r_1) \to \Gamma(Y, \Omega^r_2) \to \ldots$ of global sections of (2). We get as a corollary to the abstract De Rham theorem the following identification of the cohomology groups $H^k(\Gamma(Y, \mathcal{V}^r))$ of this complex with the De Rham cohomology groups of the manifold Y

$$H^k(\Gamma(Y, \mathcal{V}^r)) = H^kY. \tag{3}$$

Now we discuss some consequences of the theory of variational sequences. To understand the meaning of the cohomology groups (3), one should compute the classes entering the quotient spaces in (2). Note that the quotient spaces Ω_k^r/Θ_k^r are determined *up to an isomorphism*. Thus, the classes admit various equivalent characterizations. A simple analysis shows that the elements of Ω_n^r/Θ_n^r can be identified, in fibred charts, with some n-forms $\mathcal{L}\omega_0$, i.e., with some *Lagrangians* for Y. The elements of $\Omega_{n+1}^r/\Theta_{n+1}^r$ can be identified with $(n+1)$-forms $\varepsilon_\sigma \omega^\sigma \wedge \omega_0$, i.e., with *source forms*. More precisely, we have the following result.

Theorem 2 *The sheaf Ω_n^r/Θ_n^r is isomorphic with a subsheaf of the sheaf of Lagrangians $\Omega_{n,X}^{r+1}$, $\Omega_{n+1}^r/\Theta_{n+1}^r$ is isomorphic with a subsheaf of the sheaf of source forms $\Omega_{n+1,Y}^{2r+1}$, and the quotient mapping $E_n : \Omega_n^r/\Theta_n^r \to \Omega_{n+1}^r/\Theta_{n+1}^r$ is the Euler-Lagrange mapping.*

Now it is clear what kind of results are described by the variational sequence. Assume that a Lagrangian $\lambda = [\rho]$ satisfies $E_n(\lambda) = 0$. Then by exactness of (2), there always exists a class $[\eta]$ such that $E_{n-1}([\eta]) = [\rho] = [d\eta]$. This means that, locally, ρ decomposes into the sum of a closed form and a contact form. Condition

$$E_n(\lambda) = 0 \tag{4}$$

is the *local variational triviality condition*, and may be explicitly expressed with the help of (3). If in addition, $H^n Y = \{0\}$, (3) says that η may be chosen *globally defined* on $J^r Y$. The local variational triviality condition strongly determines the structure of the Lagrangians whose Euler-Lagrange forms vanish identically.

Analogously, assume that we have a source form $\varepsilon = [\rho]$ which satisfies the *local variationality condition*

$$E_{n+1}(\varepsilon) = 0. \tag{5}$$

Then there exists a class $[\eta]$ such that $E_n([\eta]) = [\rho] = [d\eta]$. Thus, locally, ρ can be expressed as the sum of a closed form and a contact form. If in addition, $H^{n+1}Y = \{0\}$, (3) guarantees that η may be chosen *globally defined* on $J^r Y$. The local variationality condition strongly determines the structure of such source forms, which can, at least locally, be treated as the Euler-Lagrange forms of suitable Lagrangians.

If ε is a source form, then $E_{n+1}(\varepsilon) = [d\varepsilon]$ is the *Helmholtz form* and E_{n+1} is the *Helmholtz mapping*. In the well-known sense, the vanishing of the Helmholtz form is a necessary and sufficient condition for existence of (local) Lagrangians for ε (the *Helmholtz conditions*).

7.3 Fibered mechanics: Local and global triviality, local and global variationality

In this subsection, we discuss differences between local and global variational triviality of Lagrangians, and local and global variationality of source forms for some examples of fibred manifolds Y over 1-*dimensional* bases X (*higher order fibred mechanics*). Recall that these concepts are defined by the Euler-Lagrange mapping $\Omega_{n,X}^r W \ni \lambda \to E_1(\lambda) \in \Omega_{n+1,Y}^{2r} W$, where W runs through open subsets of Y.

A Lagrangian $\lambda \in \Omega_1^r Y$ is locally variationally trivial if and only if $E_1(\lambda) = 0$. If λ is locally variationally trivial and $H^1 Y = 0$, then λ is globally variationally trivial. A source form $\varepsilon \in \Omega^s_{n+1,Y}W$ is locally variational if and only if $E_2(\lambda) = 0$, (in the notation of **7.2**). If ε is locally variational and $H^2Y = 0$, then ε is globally variational.

Consider simple examples (Anderson and Duchamp [2], Krupka [54]). Denote by $Q = \mathbf{R}^m, S^m, T, M, K$ the real m-dimensional Euclidean space, the m-dimensional sphere, the 2-dimensional torus $S^1 \times S^1$, the Möbius strip, and the Klein bottle, respectively. Then $H^0Q = \mathbf{R}$, and

$$\begin{aligned}
&H^i\mathbf{R}^m, \quad 1 \le i \le m, \\
&H^1S^1 = \mathbf{R}, \quad H^iS^m = 0, \quad H^mS^m = \mathbf{R}, \quad m \ge 2, \quad 1 \le i \le m-1, \\
&H^1T = \mathbf{R} \oplus \mathbf{R}, \quad H^2T = \mathbf{R}, \\
&H^1M = \mathbf{R}, \quad H^2M = 0, \\
&H^1K = \mathbf{R}, \quad H^2K = 0.
\end{aligned} \tag{1}$$

Since $\dim X = 1$, if we restrict ourselves to *connected* base manifolds, we have essentially two possibilities: (a) $X = \mathbf{R}$, and (b) $X = S^1$.

(a) Let $X = \mathbf{R}$. Assume that $Y = \mathbf{R} \times Q$. Then by the Künneth formula, $H^1(\mathbf{R} \times Q) = H^1Q$. Thus if $Q = \mathbf{R}^m$, or $Q = S^m$, $m \ge 2$, then local variational triviality always implies global variational triviality. Analogously, $H^2(\mathbf{R} \times Q) = H^2Q$. Thus if $Q = \mathbf{R}^m$, or $Q = S^m$, $m \neq 2$, or $Q = M$, $Q = K$, local variationality automatically implies global variationality. If Y is a vector bundle over $\mathbf{R}$, then local variational triviality implies global variational triviality.

(b) Let $X = S^1$. Assume that $Y = S^1 \times Q$. Then $H^1(S^1 \times Q) = H^1Q \oplus H^0Q$. Since H^0Q is always nontrivial, $H^1(S^1 \times Q) \neq 0$, and in this case local variational triviality does not imply global variational triviality. Therefore, one should examine every case independently. Similarly, $H^2(S^1 \times Q) = H^2Q \oplus H^1Q$, and if $Q = \mathbf{R}^m$, or $Q = S^m$, where $m \ge 3$, then local variationality always implies global variationality. If we consider M (K) as a fibered manifold over S^1 with fiber $\mathbf{R}$ (S^1), then in both cases, local variationality implies global variationality. If Y is a vector bundle over S^1, then local variationality implies global variationality.

7.4 Regularity and generalizations of the Hamilton theory

The theory of *Hamilton equations* for *Hamilton extremals*, presented below, has some specific features: (a) It is a theory of *Lagrangian* type, there is no *dual* concept of a Hamiltonian (such a concept arises *locally*), (b) the theory *extends* the Euler-Lagrange theory in the sense that every solution of the associated Euler-Lagrange equations is a Hamilton extremal, (c) no additional assumption, such as existence of a distinguished *time coordinate*, is imposed, and (d) in classical mechanics, where the Lagrangian satisfies the standard *regularity condition* (regularity of the Hessian matrix), the theory gives the well-known Hamilton theory.

First motivation for this geometric theory has been formulated by Dedecker [15] whose understanding of regularity was much weider than the classical one. The classical concept was considered on fibred manifolds by Goldschmidt and Sternberg [23], and generalized by Krupka and Stepanková [63], and García and Munoz [20] (see also Gotay [24], Horak and Kolář [33], Kolář [35], Krupka and Musilová [60], Shadwick [87]); Krupka [46], [49] and Krupka and Stepanková [62] interpreted the classical regularity condition *locally*, and introduced *adapted Legendre coordinates* instead of a global *Legendre morphism*. For a new, wider concept, extending regularity from Lagrangians to source forms, we refer to Krupková [65] - [68].

We know that every form $\rho \in \Omega_n^r W$ defines a Lagrangian $h\rho \in \Omega_{n,X}^{r+1} W$, the corresponding variational functional, and the Euler-Lagrange form $E_{h\rho}$. ρ also defines another variational functional

$$\Gamma_\Omega(J^r W) \ni \delta \to \rho_\Omega(\delta) = \int_\Omega \delta^* \rho \in \mathbf{R}, \tag{1}$$

whose domain are sections of $J^r Y$. Note that the sections in the set $\Gamma_\Omega(J^r W)$ are not, in general, holonomic, i.e., are not necessarily of the form $\delta = J^r \gamma$.

To be more precise, consider $J^r Y$ as fibred over X by the projection π^r. Then prolonging $J^r Y$ we get the 1-jet prolongation $J^1 J^r Y$ of the fibred manifold $J^r Y$, and the associated horizontalization, denoted $\tilde{h}$, defined by

$$\tilde{h} f = f \circ (\pi^r)^{1,0}, \quad \tilde{h} dx^i = dx^i, \quad \tilde{h} dy^\sigma_{j_1 j_2 \ldots j_p} = y^\sigma_{j_1 j_2 \ldots j_p, k} dx^k, \quad 0 \le p \le r. \tag{2}$$

Note that in this formula the associated coordinates on $J^1 J^r Y$ are denoted by $x^i, y^\sigma_{j_1 j_2 \ldots j_p, k}$ where $0 \le p \le r$. Correspondingly, let $\Lambda = \tilde{h}\rho$ be the Lagrangian; we call Λ the *extended Lagrangian*. Clearly, Λ is a $(\pi^r)^1$-horizontal form on $J^1 J^r Y$. We set

$$H_\rho = E_\Lambda, \tag{3}$$

and call H_ρ the *Hamilton form* of ρ. Thus, the Hamilton form is defined to be the Euler-Lagrange form of the extended Lagrangian. The corresponding Euler-Lagrange equations are called the *Hamilton equations*, and their solutions δ are called the *Hamilton extremals*.

As an illustration of this general scheme, consider the case $n = 1$, $r = 1$ (*first order fibred mechanics*). Denote by $(t, q^\sigma, \dot{q}^\sigma)$ some fibred coordinates on $J^1 Y$, and by $(t, q^\sigma, \dot{q}^\sigma, q_1^\sigma, \dot{q}_1^\sigma)$ the associated coordinates on $J^1 J^1 Y$. Assume that we have a first order Lagrangian $\lambda \in \Omega^1_{1,X} W$, $\lambda = \mathcal{L} dt$, and consider the *Cartan form* $\Theta = \mathcal{L} dt + (\partial \mathcal{L} / \partial \dot{q}^\sigma) \omega^\sigma$. Then $\lambda = h\Theta$ and $\Lambda = \tilde{h}\Theta = \tilde{\mathcal{L}} dt$, where

$$\tilde{\mathcal{L}} = \mathcal{L} + \frac{\partial \mathcal{L}}{\partial \dot{q}^\sigma}(q_1^\sigma - \dot{q}^\sigma), \tag{4}$$

and

$$\begin{aligned} H_\Theta = & \left(\frac{\partial \mathcal{L}}{\partial q^\sigma} + \frac{\partial^2 \mathcal{L}}{\partial q^\sigma \partial \dot{q}^\nu}(q_1^\nu - \dot{q}^\nu) - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}^\sigma} \right) dq^\sigma \wedge dt \\ & - \frac{\partial^2 \mathcal{L}}{\partial \dot{q}^\sigma \partial \dot{q}^\nu}(q_1^\nu - \dot{q}^\nu) d\dot{q}^\sigma \wedge dt. \end{aligned} \tag{5}$$

If λ is *regular*, i.e., if

$$\det \left(\frac{\partial^2 \mathcal{L}}{\partial \dot{q}^\sigma \partial \dot{q}^\nu} \right) \neq 0, \tag{6}$$

then H_Θ can be computed in the *Legendre coordinates* $p_\sigma = \partial\mathcal{L}/\partial\dot{q}^\sigma$. Denoting $\mathcal{H} = -\mathcal{L} + p_\sigma \dot{q}^\sigma$, we get

$$H_\Theta = -\left(\frac{\partial\mathcal{H}}{\partial q^\sigma} + \frac{dp_\sigma}{dt}\right) dq^\sigma \wedge dt + \left(-\frac{\partial\mathcal{H}}{\partial p_\sigma} + \frac{dq^\sigma}{dt}\right) dp_\sigma \wedge dt. \tag{7}$$

The coefficients in H_Θ are exactly the left-hand sides of the well-known *Hamilton equations*.

It is interesting that an analogous situation arises for the second order *Hilbert Lagrangian* of the general relativity theory. Let X be any n-dimensional manifold (*space-time*), $Y = \mathrm{Met}X$ the fibred manifold of *regular* tensors of degree $(0, 2)$ over X, $\lambda = \mathcal{L}\omega_0$, where $\mathcal{L} = R\sqrt{|\det(g_{ij})|}$, and R is the scalar curvature invariant (considered as a function on $J^2\mathrm{Met}X$). For any coordinates (x^i) on X, we have the associated coordinates (x^i, g_{ij}) on $\mathrm{Met}X$, and $(x^i, g_{jk}, g_{ij,k}, g_{ij,kl})$ on $J^2\mathrm{Met}X$. We can compute the principal Lepage equivalent Θ_λ. It is easily seen that Θ_λ is a *global* n-form on $J^1\mathrm{Met}X$ defined by

$$\begin{aligned}\Theta_\lambda = {} & \sqrt{|\det(g_{ij})|}\, g^{ip}(\Gamma^j_{ip}\Gamma^k_{jk} - \Gamma^j_{ik}\Gamma^k_{jp})\,\omega_0 \\ & + \sqrt{g}\,(g^{jp}g^{iq} - g^{pq}g^{ij})(dg_{pq,j} + \Gamma^k_{pq}dg_{jk}) \wedge \omega_i\end{aligned} \tag{8}$$

(cf. **7.1**, Lemma 6). The corresponding Hamilton form (3) that is a global form on $J^1J^1\mathrm{Met}X$, can be derived by a routine calculation. One can check that the Lepage form (8) satisfies a *regularity condition* (different from the standard one for quadratic Lagrangians); on the basis of this new regularity, one may also obtain the corresponding Legendre transformation, and the Hamilton equations (see [63]) for the metric field $g = g_{jk}dx^j \otimes dx^k$, equivalent with the Euler-Lagrange equation of the Hilbert Lagrangian, i.e., with the vacuum Einstein equations.

In this context, general structure and properties of the Hamilton equations in field theory have not been understood yet.

7.5 The inverse problem: First order field theory and variational energy-momentum tensors

In this subsection, we present a solution of the equation $E_{n+2}(\tau) = 0$ for the *first order* source forms τ in field theory (Haková and Krupková [31], Krupka [53]). For higher order forms only partial results are known (Krupková [65], [66]).

The results on the structure of variational first order source forms can be applied to energy-momentum tensors, known in the general relativity, and field theory. One should distinguish between (a) Noether type energy-momentum tensors, that arises when the underlying variational functionals are invariant with respect to a given Lie group, and (b) variational energy-momentum tensors, connected with variationality of the underlying source forms (of field equations). For different aspect of the geometric theory, we refer to Fernandes, García, and Rodrigo [17], Gotay and Marsden [26].

Assume that we have a source form on J^1Y, $\tau = \tau_\sigma\omega^\sigma \wedge \omega_0$. Recall that τ is said to be *variational*, if it coincides with the Euler-Lagrange form of a Lagrangian for Y. *Local variationality* is equivalent with the Helmholtz conditions

$$\frac{\partial\tau_\sigma}{\partial y^\nu_j} + \frac{\partial\tau_\nu}{\partial y^\sigma_j} = 0, \qquad \frac{\partial\tau_\sigma}{\partial y^\nu} - \frac{\partial\tau_\nu}{\partial y^\sigma} + d_j\frac{\partial\tau_\nu}{\partial y^\sigma_j} = 0. \tag{1}$$

One can prove equivalence of the following four conditions: (a) τ is locally variational, (b) there exists a unique $(n+1)$-form α on Y such that $\tau = p_1\alpha$, and $d\alpha = 0$, (c) there exists an n-form η on Y such that $\tau = E_\lambda$, and $\lambda = h\eta$, (4) τ is (globally) variational.

In any fibred chart, variationality of τ is equivalent to the existence of functions $A_{\sigma_1\sigma_2\ldots\sigma_k,i_{k+1}i_{k+2}\ldots i_n}$, $0 \le k \le n$, defined on V, such that

$$\begin{aligned}\tau = \Bigg(& -\sum_{k=1}^{n} \frac{1}{(k-1)!} \frac{\partial A_{\nu\sigma_2\ldots\sigma_k,i_{k+1}i_{k+2}\ldots i_n}}{\partial x^{i_1}} y^{\sigma_2}_{i_2} y^{\sigma_3}_{i_3} \ldots y^{\sigma_k}_{i_k} \\ & + \sum_{k=1}^{n} \frac{1}{(k-1)!} \left(-\frac{\partial A_{\nu\sigma_2\ldots\sigma_k,i_{k+1}i_{k+2}\ldots i_n}}{\partial y^{\sigma_1}} + \frac{\partial A_{\sigma_1\sigma_2\ldots\sigma_k,i_{k+1}i_{k+2}\ldots i_n}}{\partial y^{\nu}} \right) \\ & \cdot y^{\sigma_1}_{i_1} y^{\sigma_2}_{i_2} \ldots y^{\sigma_k}_{i_k} \Bigg) \omega^\nu \wedge dx^{i_1} \wedge dx^{i_2} \wedge \ldots \wedge dx^{i_n}.\end{aligned} \tag{2}$$

If τ is variational, then τ has a first order Lagrangian, where

$$\begin{aligned}\mathcal{L} = \Bigg(& A_{i_1i_2\ldots i_n} + \frac{1}{1!} A_{\sigma_1,i_2i_3\ldots i_n} y^{\sigma_1}_{i_1} + \frac{1}{2!} A_{\sigma_1\sigma_2,i_3i_4\ldots i_n} y^{\sigma_1}_{i_1} y^{\sigma_2}_{i_2} \\ & + \ldots + \frac{1}{(n-1)!} A_{\sigma_1\sigma_2\ldots\sigma_{n-1},i_n} y^{\sigma_1}_{i_1} y^{\sigma_2}_{i_2} \ldots y^{\sigma_{n-1}}_{i_{n-1}} \\ & + \frac{1}{n!} A_{\sigma_1\sigma_2\ldots\sigma_n} y^{\sigma_1}_{i_1} y^{\sigma_2}_{i_2} \ldots y^{\sigma_n}_{i_n} \Bigg) \varepsilon^{i_1i_2\ldots i_n},\end{aligned} \tag{3}$$

and the coefficients $A_{i_1i_2\ldots i_n}$, $A_{\sigma_1,i_2i_3\ldots i_n}$, $A_{\sigma_1\sigma_2,i_3i_4\ldots i_n}$, $\ldots$, $A_{\sigma_1\sigma_2\ldots\sigma_{n-1},i_n}$, $A_{\sigma_1\sigma_2\ldots\sigma_n}$ depend on x^i, y^σ only. α is given by

$$\begin{aligned}\alpha = {} & \tau_\sigma \omega^\sigma \wedge \omega_0 + \sum_{k=1}^{n} \frac{1}{k!(k+1)!} \frac{\partial^k \tau_\sigma}{\partial y^{\nu_1}_{j_1} \partial y^{\nu_2}_{j_2} \ldots \partial y^{\nu_k}_{j_k}} \omega^\sigma \wedge \omega^{\nu_1} \\ & \wedge \omega^{\nu_2} \wedge \ldots \wedge \omega^{\nu_k} \wedge \omega_{j_1j_2\ldots j_k},\end{aligned} \tag{4}$$

where $\omega_{j_1j_2\ldots j_k} = i_{\partial/\partial x^{j_k}} \omega_{j_1j_2\ldots j_{k-1}}$.

Note that conditions (1) can be applied to the problem of finding *variational energy momentum tensors*, known from the general relativity theory. We say that a source form τ is an *energy-momentum tensor* for a source form ε, if the source form $\varepsilon - \tau$ is variational.

Consequently, with necessary changes in the notation, the general structure of first order variational energy momentum tensors for the vacuum Einstein equations is described by formula (2).

7.6 Invariance: Natural variational principles, principal bundles

We briefly discuss in this subsection geometric structure of variational functionals, which are known as *invariant with respect to diffeomorphisms*, and generalizations of these variational functionals.

Such variational functionals arise when the underlying fibred manifolds belong to a given *category*, and a given Lagrangian defines a variational functional for *each* object of this category (*not* for a fixed fibred manifold only). We call Lagrangians of this kind *natural*.

To make this scheme more precise, one needs the concepts of a *category*, *covariant functor* between two categories, *natural transformation*, *differential invariant*, *natural bundle*, etc.; for generalities on jets, categories of fibre bundles, and their jet prolongations, differential invariants and natural bundles, we refer to Kolář, Michor, and Slovák [36], Krupka and Janyska [56], D. Krupka, M. Krupka [57], Nijenhuis [77], and Saunders [85].

The most common examples of a covariant functor are the *tangent functor* T, assigning to a manifold X its *tangent bundle* TX, and to a morphism $f : X_1 \to X_2$ of manifolds the *tangent mapping* $Tf : TX_1 \to TX_2$, and the *tensor functors*, derived from T.

Historically, one of the well-known examples of a natural Lagrangian is the *Hilbert Lagrangian* for the Einstein equations (that is, the scalar curvature function on a pseudoriemannian manifold, considered as a Lagrangian on an appropriate jet bundle). Many others can be found in classical sources (see e.g. Lovelock [70], [71], Rund [83], Rund and Lovelock [84]). A differential geometric discussion of several examples based on jets can be found in Horak and Krupka [34], [44], P. Musilová and Krupka [74], P. Musilová and J. Musilová [75], and Novotný [78]. The *theory* as presented below follows the approach, explained in Krupka [43], [47], [50], and Krupka and Trautman [64].

Denote by $\mathcal{D}_n$ the category of smooth manifolds and their diffeomorphisms, and by $\mathcal{P}B_n(G)$ the cateory of principal bundles over n-dimensional manifolds, with structure group G; the *morphisms* in $\mathcal{P}B_n(G)$ are considered to be homomorphisms of principal bundles, whose projections belong to the category $\mathcal{D}_n$. Fibre bundles, associated with objects of the category $\mathcal{P}B_n(G)$, and homomorphisms of these fibre bundles, form a category, denoted by $\mathcal{F}B_n(G)$.

Suppose that we have a covariant functor $\tau : \mathcal{D}_n \to \mathcal{P}B_n(G)$ (a *lifting functor*, or just a *lifting*). τ assigns to every n-dimensional manifold X and every diffeomorphism $f : U \to X$, where U is an open set in X, a morphism of fibred manifolds $\tau f : \tau U \to \tau X$, commuting with the projection of τX. Let Q be a manifold, endowed with a left action of the Lie group G. Q defines the fibre bundle with fibre Q, *associated* with τX, whose *type fibre* is Q, and a morphism of fibre bundles $\tau_Q f : \tau_Q U \to \tau_Q X$, commuting with the projection of $\tau_Q X$. We get a covariant functor $\tau_Q : \mathcal{D}_n \to \mathcal{F}B_n(G)$, called a *Q-lifting* from the category $\mathcal{D}_n$ to the category $\mathcal{F}B_n(G)$, associated with τ, or simply a *lifting*.

We introduce an important example, the *higher order frame lifting*. By the r-th *differential group* L^r_n of $\mathbf{R}^n$ we mean the Lie group of invertible r-jets with source and target at the origin $0 \in \mathbf{R}^n$; by an r-*frame* at a point $x \in X$ we mean an invertible r-jet with source $0 \in \mathbf{R}^n$ and target x. The set $\mathcal{F}^r X$ of all r-frames has a natural structure of a principal bundle with structure group L^r_n. Every diffeomorphism $\alpha : U \to X$, where U is an open set in X, defines, by means of composition of diffeomorphisms, a morphism of principal bundles $\mathcal{F}^r\alpha : \mathcal{F}^r U \to \mathcal{F}^r X$; the correspondence $X \to \mathcal{F}^r X$, $\alpha \to \mathcal{F}^r\alpha$ is a covariant functor from the category $\mathcal{D}_n$ to $\mathcal{P}B_n(L^r_n)$, called the r-*frame lifting* ($\mathcal{F}^1 X$ is the bundle of *linear frames*, and $\mathcal{F} = \mathcal{F}^1$ is the standard *frame lifting*).

Let Q be a space of tensors on the vector space $\mathbf{R}^n$; elements of Q are called *tensors of type* Q over $\mathbf{R}^n$. Q is endowed with the *tensor action* $(g, p) \to g \cdot p$ of the general linear group $\mathrm{Gl}_n(\mathbf{R})$. The *Q-lifting* is the correspondence, assigning to an n-dimensional manifold X the tensor bundle $\tau_Q X$ of tensors of type Q over X, and to any isomorphism f of manifolds the corresponding isomorphism $\tau_Q f$ of tensor bundles. Let $T^r_n Q$ be the set of r-jets with source at $0 \in \mathbf{R}^n$ and target in Q. Then the mapping

$$L_n^{r+1} \times T_n^r Q \ni (J_0^{r+1}\alpha, J_x^r \zeta) \to J_0^{r+1}\alpha \cdot J_x^r \zeta = J_0^r((D\alpha \cdot \zeta) \circ \alpha^{-1}) \in T_n^r Q \quad (1)$$

is a left action of the group L_n^{r+1} on $T_n^r Q$ (Krupka [43]). Computing the $(r+1)$-jet on the right-hand side in components, one can easily see that this formula represents, formally, the transformation rules for components of tensors of type Q and their derivatives up to order r.

Lemma 1 (a) *Formula (1) defines on the r-jet prolongation $J^r \tau_Q X$ the structure of a fibre bundle with fibre $T_n^r Q$, associated with the principal bundle $\mathcal{F}^{r+1} X$.*

(b) *The correspondence $X \to J^r \tau_Q X$, $f \to J^r \tau_Q f$ is a covariant functor from the category $\mathcal{D}_n$ to $\mathcal{F}B_n(L_n^{r+1})$.*

The functor $J^r \tau_Q$ is called the *r-jet prolongation of the lifting* τ_Q.

Let λ be a Lagrangian of order r for $\tau_Q X$. We say that λ is *natural*, if for every diffeomorphism $\alpha : U \to X$, α_Q is an invariance transformation of λ, i.e.,

$$(J^r \alpha_Q)^* \lambda = \lambda \quad (2)$$

on the corresponding open set. We are now in a position to prove the following result, a version of the *second theorem of Emmy Noether*, stating that the Euler-Lagrange expressions of a natural Lagrangian λ and the currents $i_{J^r \tau_Q \xi} \Theta_\lambda$ satisfy certain equations.

Theorem 1 *Let X be an n-dimensional manifold, and let λ be a natural Lagrangian of order r on $J^r \tau_Q X$. Let ρ be a Lepage equivalent of λ. Then for every section γ and every vector field ξ on X*

$$J^{2r}\gamma^* i_{J^{2r}\tau_Q \xi} E_\lambda + dJ^{2r-1}\gamma^* i_{J^{2r-1}\tau_Q \xi} \Theta_\lambda = 0. \quad (3)$$

In particular, each extremal γ satisfies the conservation laws

$$dJ^{2r-1}\gamma^* i_{J^{2r-1}\tau_Q \xi} \Theta_\lambda = 0. \quad (4)$$

Indeed, let ξ be any vector field on X, and let α_t^ξ be the local one-parameter group of ξ. Applying τ_Q to α_t^ξ, we get invariance transformations $J^r \tau_Q \alpha_t^\xi$ of λ, and a generator $\tau\xi$ of the one-parameter group α_t^ξ. Then by definition, the Lie derivative $\partial_{\tau\xi}\lambda$ vanishes, so the infinitesimal first variation formula

$$\partial_{\tau\xi}\lambda = i_{\tau\xi} d\Theta_\lambda + d i_{\tau\xi}\Theta_\lambda \quad (5)$$

for these vector fields reduces to $i_{J^{2r-1}\tau_Q \xi} d\Theta_\lambda + d i_{J^{2r-1}\tau_Q \xi}\Theta_\lambda = 0$, that is, to (4).

Note that Lemma 1 as well as Theorem 1 can be easily formulated for liftings $J^r \tau_Q$, where Q is an *arbitrary* manifold, endowed with an action of the general linear group, or, more generally, with an action of any differential group L_n^s (Krupka [43], [47]). In particular, the same assertions are valid for variational functionals, whose underlying spaces are *frame bundles*.

For various aspects of the theory of variational functionals on frame bundles and general principal bundles, invariant with respect to the structure group, we refer to Brajerčík, Krupka [7], [9], Castrillon Lopez and Munoz Masque [10], Castrillon Lopez, Munoz Masque and Ratiu [11], Cendra, Ibort and Marsden [12], [13], Munoz Masque and Rosado Maria [72], [73], and Prieto [81].

Acknowledgement

The author appreciates the support of the Czech Science Foundation (grant 201/06/0922), and the Czech Ministry of Education, Youth, and Sports (grant MSM 6198959214). He is also grateful to the Mathematics Department and the Institute for Advanced Study at La Trobe University, Melbourne, where a part of the research has been completed.

Bibliography

[1] S. J. Aldersley: Higher Euler operators and some of their applications *J. Math. Phys.* **20** (1979) 522–531

[2] I. Anderson and T. Duchamp: On the existence of global variational principles *Am. J. Math.* **102** (1980) 781–867

[3] I. Anderson and J. Pohjanpelto: Variational principles for differential equations with symmetries and conservation laws, 1. Second order scalar equations *Math. Ann.* **299** (1994) 191–222;
2. Polynomial differential equations *Math. Ann.* **301** (1995) 627–653

[4] I. Anderson and G. Thompson: *The Inverse Problem of the Calculus of Variations for Ordinary Differential Equations* (Memoirs of the AMS 98, No. 473, 1992)

[5] D. Betounes: Extension of the classical Cartan form *Phys. Rev.* **D 29** (1984) 599–606

[6] J. Brajerčík and D. Krupka: Cohomology and local variational principles *In: Proc. XVth Internat. Workshop on Geometry and Physics (Puerto de la Cruz, Tenerife, Spain, Sept. 2006)* (to appear)

[7] J. Brajerčík and D. Krupka: Noether currents and order reduction on frame bundles *In: Proc. 40th Sympos. on Finsler Geometry "In the Memory of our Teachers"* (Hokkaido Tokai University, Sapporo, Society of Finsler Geometry, 2006) 34–37

[8] J. Brajerčík and D. Krupka: Variational principles for locally variational forms *J. Math. Phys.* **46** (052903), (2005) 1–15

[9] J. Brajerčík and D. Krupka: Variational principles on the frame bundles *In: Differential Geometry and its Applications, Proc. Conf., Prague, August 2004* (Charles University, Prague, 2005) 559–569

[10] M. Castrillon Lopez and J. Munoz Masque: Gauge invariant variationally trivial problems on T^*M *J. Math. Phys.* **40** (1999) 821–829

[11] M. Castrillon Lopez, J. Munoz Masque and T. Ratiu: Trivial lagrangians on connections and invariance under automorphisms *In: Steps in Differential Geometry, L. Kozma, P. T. Nagy and L. Tamássy, Eds.; Proc. Colloq. on Diff. Geom., Debrecen, Hungary, July 2000* 77–83

[12] H. Cendra, A. Ibort and J. E. Marsden: Variational principles on principal fibre bundles: A geometric theory of Clebsch potentials and Lin constraints *J. Geom. Phys.* **4** (1987) 183–205

[13] H. Cendra and J. E. Marsden: Lin constraints, Clebsch potentials and variational principles *Physica* **27D** (1987) 63–89

[14] M. Crampin and D. J. Saunders: The Hilbert-Caratheodory forms for parametric multiple integral problems in the calculus of variations *Acta Appl. Math.* **76** (2003) 37–55

[15] P. Dedecker: On the generalization of symplectic geometry to multiple integrals in the calculus of variations *In: Lecture Notes in Math. 570* (Springer, Berlin, 1977) 395–456

[16] P. Dedecker and W. M. Tulczyjew: Spectral sequences and the inverse problem of the calculus of variations, Internat. Colloq., Aix-en-Provence, 1979 *In: Differential Geometric Methods in Mathematical Physics, Lecture Notes in Math. 836* (Springer, Berlin, 1980) 498–503

[17] A. Fernandes, P. L. García, and C. Rodrigo: Stress-energy-momentum tensors in higher order variational calculus *J. Geom. Phys.* **34** (2000) 41–72

[18] M. Francaviglia, M. Palese and R. Vitolo: Symmetries in finite order variational sequences *Czech. Math. J.* **52** (2002) 197–213

[19] P. L. García: The Poincaré-Cartan invariant in the calculus of variations *Symposia Mathematica* **14** (1974) 219–246

[20] P. L. García and J. M. Masqué: Le probléme de la régularité dans le calcul des variations du second ordre *C. R. Acad. Sci. Math.* **301** (1985) 639–642

[21] F. Gianessi and A. Maugeri (Eds): *Variational Analysis and Applications* (Springer, New York, 2005)

[22] M. Giaquinta and S. Hildebrandt: *Calculus of Variations I, II* (Springer, Berlin, 1996)

[23] H. Goldschmidt and S. Sternberg: The Hamilton-Cartan formalism in the calculus of variations *Ann. Inst. H. Poincaré* **23** (1973) 203–267

[24] M. Gotay: A multisymplectic framework for classical field theory and the calculus of variations, I. Covariant Hamiltonian formalism *In: M. Francaviglia, D.D. Holm, Eds.: Mechanics, Analysis, and Geometry: 200 Years After Lagrange* (North Holland, Amsterdam, 1990) 203–235

[25] M. Gotay: An exterior differential systems approach to the Cartan form *In: Geométrie Symplectique et Physique Mathematique, P. Donato at all., Eds.* (Birkhauser, Boston, 1991)

[26] M. J. Gotay and J.E. Marsden: Stress-energy-momentum tensor and the Belinfante-Rosenfeld formula *Contemporary Mathematics* **132** (1992) 367–392

[27] M. Grassi: Local vanishing of characteristic cohomology *Duke Math. J.* **102** (2000) 307–328

[28] D. R. Grigore: Lagrangian formalism on Grassmann manifolds *In: Handbook of Global Analysis* (Amsterdam: Elsevier, 2007) 327–373

[29] D. R. Grigore: The variational sequence on finite jet bundle extensions and the Lagrange formalism *Diff. Geom. Appl.* **10** (1999) 43–77

[30] D. R. Grigore and D. Krupka: Invariants of velocities and higher order Grassman bundles *J. Geom. Phys.* **24** (1998) 244–264

[31] A. Haková and O. Krupková: Variational first-order partial differential equations *J. Differential Equations* **191** (2003) 67–89

[32] F. Helein and J. C. Wood: Harmonic maps *In: Handbook of Global Analysis* (Amsterdam: Elsevier, 2007) 417–491

[33] M. Horak and I. Kolář: On the higher order Poincaré-Cartan forms *Czechoslovak Math. J.* **38** (1983) 467–475

[34] M. Horak and D. Krupka: First order invariant Einstein-Cartan structures *Internat. J. Theoret. Phys.* **17** (1978) 573–583

[35] I. Kolář: A geometric version of the higher order Hamilton formalism in fibered manifolds *J. Geom. Phys.* **1** (1984) 127–137

[36] I. Kolář, P. Michor and J. Slovák: *Natural Operations in Differential Geometry* (Springer-Verlag, Berlin, 1993)

[37] I. Kolář and R. Vitolo: On the Helmholtz operator for Euler morphisms *Math. Proc. Camb. Phil. Soc.* **135** (2003) 277–290

[38] J. Krasilschik: Geometry of differential equations: A concise introduction *Acta Appl. Math.* **72** (2002)

[39] M. Krbek and J. Musilová: Representation of the variational sequence by forms *Acta Applicandae Mathematicae* **88** (2005) 177–199

[40] M. Krbek and J. Musilová: Representation of the variational sequence *Rep. Math. Phys.* **51** (2003) 251–258

[41] D. Krupka: A geometric theory of ordinary first order variational problems in fibered manifolds, I. Critical sections *J. Math. Anal. Appl.* **49** (1975) 180–206;
II. Invariance *J. Math. Anal. Appl.* **49** (1975) 469–476

[42] D. Krupka: A map associated to the Lepagean forms of the calculus of variations in fibered manifolds *Czech. Math. J.* **27** (1977) 114–118

[43] D. Krupka: A setting for generally invariant Lagrangian structures in tensor bundles *Bull. Acad. Polon. Sci. , Ser. Math. Astronom. Phys.* **22** (1974) 967–972

[44] D. Krupka: A theory for generally invariant Lagrangians for the metric fields I *Internat. J. Theoret. Phys.* **17** (1978) 359–368;
II *Internat. J. Theoret. Phys.* **15** (1976) 949–959

[45] D. Krupka: Lepagean forms in higher order variational theory *In: Modern Developments in Analytical Mechanics, Proc. IUTAM-ISIMM Sympos., Turin, June 1982* (Academy of Sciences of Turin, 1983) 197–238

[46] D. Krupka: On the higher order Hamilton theory in fibered spaces *In: Proc. Conf. on Diff. Geom. Appl., Vol. 2, Nové Mesto na Morave, Sept. 1983* (J. E. Purkyne Univ., Brno, 1984) 167–184

[47] D. Krupka: Natural Lagrangian structures *In: Semester on Differential Geometry, Banach Center, Warsaw, 1979* (Banach Center Publications **12**, 1984) 185–210

[48] D. Krupka: On the local structure of the Euler-Lagrange mapping of the calculus of variations *In: Proc. Conf. on Diff. Geom. Appl, Nove Mesto na Morave, Sept. 1980* (Charles Univ., Prague, 1981) 181-188; `arXiv:math-ph/0203034`

[49] D. Krupka: Regular lagrangians and Lepagean forms *In: Differential Geometry and Its Applications, Proc. Conf., D. Krupka and A. Svec, Eds., Brno, Czechoslovakia, 1986* (D. Reidel, Dordrecht, 1986) 111–148

[50] D. Krupka: *Some Geometric Aspect of the Calculus of Variations in Fibered Manifolds* (Folia Fac. Sci. Nat. UJEP Brunensis **14**, 1973); `arXiv:math-ph/0110005`

[51] D. Krupka: The total divergence equation *Lobachevskii Journal of Mathematics* **23** (2006) 71–93

[52] D. Krupka: Trace decompositions of tensor spaces *Lin. and Multilin. Alg.* **54** (2006) 235–263

[53] D. Krupka: Variational principles for energy-momentum tensors *Rep. Math. Phys.* **49** (2002) 259–268

[54] D. Krupka: Variational sequences in mechanics *Calc. Var.* **5** (1997) 557–583

[55] D. Krupka: Variational sequences on finite order jet spaces *In: Differential Geometry and its Applications, Proc. Conf., Brno, Czechoslovakia, 1989* (World Scientific, Singapore, 1990) 236–254

[56] D. Krupka and J. Janyska: *Lectures on Differential Invariants* (J. E. Purkyne Univ., Faculty of Science, Brno, Czechoslovakia, 1990); 193 pp.

[57] D. Krupka and M. Krupka: Jets and Contact Elements *In: Proceedings of the Seminar on Differential Geometry, D. Krupka, ed.* (Mathematical Publications, Silesian Univ. in Opava, Opava, Czechia, 2000) 39–85

[58] D. Krupka, O. Krupková, G. Prince and W. Sarlet: Contact symmetries and variational sequences *In: Differential Geometry and Its Applications, Proc. Conf., Edited by J. Bures, O. Kowalski, D. Krupka, J. Slovák, Prague, August 2004* (Charles University in Prague, Czech Republic, 2005) 605–615

[59] D. Krupka, O. Krupková, G. Prince and W. Sarlet: Contact symmetries of the Helmholtz form *Diff. Geom. Appl.* (to appear)

[60] D. Krupka and J. Musilová: Hamilton extremals in higher order mechanics *Arch. Math. (Brno)* **20** (1984) 21–30

[61] D. Krupka and J. Musilová: *Calculus of Odd Base Forms on Differential Manifolds* (Folia Fac. Sci. Nat. Univ. Purk. Brunensis, Physica **24**, Brno, Czechoslovakia, 1983)

[62] D. Krupka and J. Sedenková: Variational sequences and Lepage forms *In: Differential Geometry and Its Applications, Proc. Conf., Edited by J. Bures, O. Kowalski, D. Krupka, J. Slovák, Prague, August 2004* (Charles University in Prague, Czech Republic, 2005) 617–627

[63] D. Krupka and O. Stepanková: On the Hamilton form in second order calculus of variations *In: Proc. of the Meeting "Geometry and Physics", Florence, October 1982* (Pitagora Editrice Bologna, 1983) 85–101

[64] D. Krupka and A. Trautman: General invariance of Lagrangian structures *Bull Acad. Polon. Sci, Ser. Math. Astronom. Phus* **22** (1974) 207–211

[65] O. Krupková: Hamiltonian field theory *J. Geom. Phys.* **43** (2002) 93–132

[66] O. Krupková: Hamiltonian field theory revisited: A geometric approach to regularity *In: Steps in Differential Geometry, Proc. Colloq. Diff. Geom., Debrecen, July 2000, L. Kozma, P.T. Nagy, and L. Tamássy, Eds.* (University of Debrecen, Debrecen, 2001) 187–207

[67] O. Krupková: Lepagean 2-forms in higher order Hamiltonian mechanics, I. Regularity *Arch. Math. (Brno)* **22** (1986) 97–120;
II. Inverse problem *Arch. Math. (Brno)* **23** (1987) 155–170

[68] O. Krupková: *The Geometry of Ordinary Variational Equations, Lecture Notes in Math. 1678* (Springer, Berlin, 1997)

[69] O. Krupková and G. Prince: Second Order Ordinary Differential Equations in Jet Bundles and the Inverse Problem of the Calculus of Variations *In: Handbook of Global Analysis* (Amsterdam: Elsevier, 2007) 837–904

[70] D. Lovelock: Degenerate Lagrange densities involving geometric objects *Arch. Ration. Mech. Anal.* **36** (1970) 293–304

[71] D. Lovelock: The Euler-Lagrange expressions and degenerate Lagrange densities *J. Australian Math. Soc.* **14** (1972) 482–495

[72] J. Munoz Masque and M. Eugenia Rosado Maria: Invariant variational problems on linear frame bundles *J. Phys. A: Math. Gen.* **25** (2002) 2013–2036

[73] J. Munoz Masque and M. Eugenia Rosado Maria: Integrability of the field equations of invariant variational problems in linear frame bundles *J. Geom. Phys.* **49** (2004) 119–155

[74] P. Musilová and D. Krupka: Differential invariants of immersions of manifolds with metric fields *Rep. Math. Phys.* **51** (2003) 307–313

[75] P. Musilová and J. Musilová: Differential invariants of immersions of manifolds with metric fields *Commun. Math. Phys.* **249** (2), (2004) 319–329

[76] J. Musilová: Variational sequence in higher order mechanics *In: Differential Geometry and Applications, Proc. Conf., 1995, Brno, Czech Republic* (Masaryk University, Brno, 1996) 611–624

[77] A. Nijenhuis: Natural bundles and their general properties, Geometric objects revisited *In: Differential Geometry, in honour of K. Yano* (Kinokuniya, Tokyo, 1972) 317–334

[78] J. Novotný: On generally invariant Lagrangians for the metric field and other tensor fields *Internat. J. Theoret. Phys.* **17** (1978) 667-684

[79] P. Olver: Equivalence and the Cartan form *Acta Appl. Math.* **31** (1993) 99–136

[80] J. F. Pommaret: Spencer sequence and variational sequence *Acta Appl. Math.* **41** (1995) 285–296

[81] C. T. Prieto: Variational formulation of Chern-Simons theory for arbitrary Lie groups *J. Geom. Phys.* **50** (2004) 138–161

[82] H. Rund: A Cartan form for the field theory of Caratheodory in the calculus of variations *In: Lecture Notes in Pure and App. Math. 100* (Springer, Berlin, 1985) 455–469

[83] H. Rund: Invariant theory of variational problems for geometric objects *Tensor N.S.* **18** (1967) 239–258

[84] H. Rund and D. Lovelock: Variational principles in the general theory of relativity *Uber Deutsch. Math.-Verein.* **74** (1972) 1–65

[85] D. J. Saunders: Jet manifolds and natural bundles *In: Handbook of Global Analysis* (Amsterdam: Elsevier, 2007) 1035–1068

[86] D. J. Saunders: *The Geometry of Jet Bundles* (Cambridge Univ. Press, 1989)

[87] W. F. Shadwick: The Hamiltonian formulation of regular r-th order Lagrangian field theories *Letters in Math. Phys.* **6** (1982) 409–416

[88] J. Sniatycki: On the geometric structure of classical field theory in Lagrangian formulation *Proc. Camb. Phil. Soc.* **68** (1970) 475–484

[89] J. Stefanek: A representation of the variational sequence in higher order mechanics *In: Differential Geometry and Applications, Proc. Conf., 1995, Brno, Czech Republic* (Masaryk University, Brno, 1996) 469–478

[90] F. Takens: A global version of the inverse problem of the calculus of variations *J. Differential Geometry* **14** (1979) 543–562

[91] E. Tonti: Variational formulation of nonlinear differential equations: I *Bull. Acad. Roy. Belg. C. Sci* **55** (1969) 137–165;
II *Bull. Acad. Roy. Belg. C. Sci* **55** (1969) 262–278

[92] A. Trautman: Invariance of Lagrangian systems *In: General Relativity, Papers in honor of J. L. Synge* (Oxford, Clarendon Press, 1972) 85–99

[93] A. Trautman: Noether equations and conservation laws *Comm. Math. Phys.* **6** (1967) 248–261

[94] W. M. Tulczyjew: The Euler-Lagrange resolution *In: Internat. Coll. on Diff. Geom. Methods in Math. Phys., Aix-en-Provence, 1979; Lecture Notes in Math. 836* (Springer, Berlin, 1980) 22–48

[95] A. M. Vinogradov: A spectral sequence associated with a non-linear differential equation, and the algebro-geometric foundations of Lagrangian field theory with constraints *Soviet Math. Dokl.* **19** (1978) 790–794

[96] A. M. Vinogradov, I. S. Krasilschik and V. V. Lychagin: *Introduction to the Geometry of Non-linear Differential Equations* (Russian) (Nauka, Moscow, 1986)

[97] R. Vitolo: Finite order Lagrangian bicomplexes *Math. Proc. Cambridge Phil. Soc.* **125** (1999) 321–333

[98] R. Vitolo: Variational sequences *In: Handbook of Global Analysis* (Amsterdam: Elsevier, 2007) 1115–1163

DEMETER KRUPKA

Department of Algebra and Geometry, Palacky University,
779 00 Olomouc, Czech Republic

and

Department of Mathematics, La Trobe University,
Melbourne, Victoria 3086, Australia

E-mail: krupka@inf.upol.cz

Handbook of Global Analysis
Demeter Krupka and David Saunders

Second Order Ordinary Differential Equations in Jet Bundles and the Inverse Problem of the Calculus of Variations

O. Krupková and G. E. Prince

Contents

1 Introduction

Second order ordinary differential equations on finite dimensional manifolds appear in a wide variety of applications in mathematics, physics and engineering. In differential geometry they describe the autoparallel curves of a linear connection, the geodesics of the metric in Riemann and Finsler geometries and the integral curves of the Reeb field on a contact manifold. In the calculus of variations they are the Euler-Lagrange equations in the single independent variable case. In classical mechanics they are Newton's equations of motion and the Euler-Lagrange equations of a mechanical Lagrangian. In general relativity and its variants they describe worldlines of free particles. In classical electrodynamics they describe the paths of charged particles.

The first observation to be made here is that the calculus of variations puts up an umbrella over many of these cases, bringing with it the manifest benefits of the integrability theorems of Noether, Liouville, and Jacobi. In the case of the autoparallel curves of a linear connection, there is an obvious inverse problem: *"are these curves the geodesics of some metric?"* (This question has both Riemannian and Finslerian versions.) When this is true additional, geometric benefits flow. Clearly, we should generally ask *" when are the solutions of a system second order ordinary differential equations those of a system of Euler-Lagrange equations (on the same manifold)?"*

A second observation, or rather reservation, arises: just how important is the existence of a variational principle for a system of these equations? This question is indicated by the fact that, at least locally, every regular system of n second order ordinary differential equations on an n dimensional manifold M provides a Reeb field on $\mathbb{R}\times TM$. On the other hand the variationality of such a system is not even a universal local property. Since the classical integrability theorems are available for Reeb fields, why bother with variationality?

To answer this rhetorical question we remark that variational equations play a fundamental role, not only in physics but also in the theory of differential equations alone. Essentially, for *regular Lagrangians* they have a fundamental alternative-Hamilton equations-that are first order equations, equivalent with the Euler-Lagrange equations, appearing as equations for integral curves of a vector field on a prolongation of the configuration manifold ("phase space"). And, of course, all the known integration methods for variational equations in classical mechanics based on symmetries and first integrals (the Noether and Liouville theorems), as well as the powerful Hamilton-Jacobi integration method, benefit from this representation. They can be used to solve differential equations once a Lagrangian is known. Next, the existence of a Lagrangian is of high importance in physics: in particular, without a Lagrangian there is no quantisation.

The restriction of the use of the classical Hamilton and Hamilton-Jacobi theory to the class of regular Lagrangians was a motivation for Dirac to start to study singular differential equations, that is, such that cannot be put into the normal form

$$\ddot{x}^a = f^a(t, x, \dot{x}).$$

Surprisingly, it turns out that singular equations have unexpected properties: even in the smooth case the Cauchy initial problem may have more solutions or no solution at all, and, as a significant complication, integration methods so useful for solving regular equations, cannot be applied, or their use may produce incorrect results.

Conversely, motivations and ideas coming from the calculus of variations and physics recently resulted in a systematic study, based on differential geometry and global analysis, of general, not necessarily regular, second and higher order differential equations. As a benefit one obtains a setting for a general geometric theory of differential equations *where the class of variational equations appears as a special case*, and many results and techniques, known only within the variational calculus are extended and generalised to differential equations in general. In this setting, of course, the question of whether the given equations come from a Lagrangian, or more generally, are equivalent with some variational equations, plays a crucial role.

In this article we deal with the geometry of second order ordinary differential equations (SODEs) and its intimate relationship to the corresponding inverse problem in the calculus of variations which we spelt out above. While yet there is no complete solution to the inverse problem in general, there have been significant advances since the seminal papers of Helmholtz [47] in 1887 and Douglas [30] in 1941: Helmholtz finding conditions for variationality of a system of SODEs in "covariant form",

$$B_{ab}(t, x, \dot{x})\ddot{x}^b + A_a(t, x, \dot{x}) = 0,$$

and Douglas solving the $n = 2$ "contravariant" case (that concerns regular equations in normal form). But more importantly, the underlying geometry of both generic and variational SODEs has developed enormously, bringing deeper understanding to the many areas of application in which SODEs appear. We leave the reader to decide whether the search for variationality is a fruitful one.

2 Second-order differential equations on fibred manifolds

The study of second and higher-order differential equations on manifolds needs to consider *jet bundles* as appropriate underlying spaces. Systems of differential equations on manifolds and their solutions then can be modelled by global objects defined on manifolds of jets and studied by tools of global analysis and differential geometry. We shall be interested in systems of second-order ordinary differential equations whose solutions are curves in a smooth manifold M. In this case, it is appropriate to consider the fibred manifold $\mathbb{R} \times M \to \mathbb{R}$, whose sections are graphs of curves into M, and its jet prolongations. We start with a short exposition of the concepts of fibred spaces, jets of mappings, and jet fields and jet connections going back to Ehresmann [31, 32, 33], and the calculus of horizontal and contact forms in fibred manifolds due to Krupka [58, 62]; for more details the reader can consult e.g. [82, 129].

For a global representation of systems of second-order differential equations we use *dynamical forms* (that are certain 2-forms on the second jet bundle). Regular equations then are described either by regular second-order dynamical forms or by *semisprays* (second-order vector fields). The core of this section is a study of exterior differential systems related with second order differential equations, and geometric classification of general systems of SODEs, due to Krupková [74, 77, 78, 79, 82, 84, 87].

2.1 Fibred manifolds and their jet prolongations

Let M be a manifold of dimension m, and consider the projection $\pi_0 : \mathbb{R} \times M \to \mathbb{R}$. π_0 is an example of a *fibred manifold* over $\mathbb{R}$. For every point $t_0 \in \mathbb{R}$ the fibre $\pi_0^{-1}(t_0)$ is diffeomorphic to M. In this context, M is called *configuration space*, and the total space $\mathbb{R} \times M$ of the fibred manifold π_0 is called *extended configuration space*. On $\mathbb{R} \times M$ we shall always consider local coordinates of the form (t, x^a), where t is the global coordinate on $\mathbb{R}$ and (x^a), $1 \leq a \leq m$, are local coordinates on M.

If $c : \mathbb{R} \to M$ is a curve in M defined on an open interval $I \subset \mathbb{R}$, we denote by γ its graph, i.e. $\gamma(t) = (t, c(t))$. Since $\pi_0 \circ \gamma = \mathrm{id}_I$, γ is a *(local) section* of the fibred manifold π_0.

Let $s \geq 1$ be an integer. Two sections γ_1, γ_2 of π_0, defined on an open set $I \subset \mathbb{R}$, are called *s-equivalent* at a point $t_0 \in I$, if $\gamma_1(t_0) = x = \gamma_2(t_0)$, and if there is a chart (t, x^a) around x such that the derivatives of the components γ_1^a, γ_2^a of these sections at the point t_0 coincide up to the order s, i.e., if

$$\frac{d^k\gamma_1^a}{dt}(t_0) = \frac{d^k\gamma_2^a}{dt}(t_0), \quad 1 \leq k \leq s,$$

for every $a = 1, \ldots, m$. The equivalence class containing a section γ is called the *s-jet* of γ at t_0 and is denoted by $J^s_{t_0}\gamma$. The union of all s-jets at all points of $\mathbb{R}$ is a manifold of dimension $1 + m(s + 1)$, denoted by $J^s\pi_0$ and called the *s-jet prolongation* of the fibred manifold π_0. For our case when $\pi_0 : \mathbb{R} \times M \to \mathbb{R}$ it can be shown that $J^s\pi_0$ as a fibred manifold over $\mathbb{R}$ identifies with $\mathbb{R} \times T^sM \to \mathbb{R}$, where T^sM is the *manifold of s-velocities* in the sense of Ehresmann; in particular, $T^1M = TM$.

One has a family of natural fibred projections from $J^s\pi_0$ onto $\mathbb{R}$ and $J^k\pi_0$, $0 \leq k \leq s-1$, denoted by π_s and $\pi_{s,k}$, respectively. Here, for the sake of simplicity, we have used the notation $J^0\pi_0 = \mathbb{R} \times M$ that will be used later when appropriate.

In what follows, the *first jet prolongation of the extended configuration space* $\mathbb{R} \times M$ will play a crucial role. We denote $E = \mathbb{R} \times TM$ and call it *evolution space*.

Every section γ of π_0 naturally prolongs to a section $J^s\gamma$ of π_s defined by

$$J^s\gamma(t) = J^s_t\gamma, \quad \forall t \in I,$$

and called the *s-jet prolongation of* γ. A general section of π_s, however, need not be of this form; we say that a section δ of the fibred manifold π_s is *holonomic* if $\delta = J^s\gamma$ for a section γ of π_0. If $\gamma(t) = (t, c(t))$, then $J^s\gamma(t) = (t, c(t), \dot{c}(t), \ddot{c}(t), \ldots, c^{(s)}(t))$.

If (t, x^a) are local coordinates on $\mathbb{R} \times M$, we denote the associated coordinates on $\mathbb{R} \times TM$ by $(t, x^a, \dot{x}^a)$ or (t, x^a, u^a), and the corresponding coordinates on $\mathbb{R} \times T^2M$ by $(t, x^a, \dot{x}^a, \ddot{x}^a)$. For a general s we also write (t, x^a_k), $0 \leq k \leq s$.

2.2 Calculus on jet bundles

On fibred manifolds and their jet prolongations there arise many specific geometric objects, such as vector fields, differential forms, distributions, etc., which are *adapted to the fibred and prolongation structures*.

A vector field X on $J^s\pi_0$ is called π_s*-vertical* if $T\pi_s.X = 0$, and π_s*-projectable* if there exists a vector field X_0 on the base $\mathbb{R}$ such that $T\pi_s.X = X_0 \circ \pi_s$. Note that considering a vertical vector field on the extended configuration space $\mathbb{R} \times M$ corresponds to considering a "time-dependent" vector field on M, while a (non-vertical) projectable vector field on the extended configuration space does not have a counterpart on M.

In local coordinates, projectable vector fields have their $\partial/\partial t$ component dependent on t only, and vertical vector fields have this component equal to zero.

Local flows of projectable vector fields transfer sections into sections; consequently, π_0-projectable vector fields on the extended configuration space can be naturally prolonged to vector fields on $J^s\pi_0$. The procedure is as follows: Let X be a π_0-projectable vector field, X_0 its projection, and denote $\{\phi_u\}$ and $\{\phi_{0u}\}$) the corresponding local one-parameter groups. For every u, the mapping ϕ_u is an isomorphism of the fibred manifold π_0, i.e. $\pi_0 \circ \phi_u = \phi_{0u} \circ \pi_0$. Then for every section γ, the composition $\bar{\gamma} = \phi_u \circ \gamma \circ \phi_{0u}^{-1}$ is again a section and we can define the *s-jet prolongation* of ϕ_u by

$$J^s\phi_u(J^s_t\gamma) = J^s_{\phi_{0u}(t)}(\phi_u\gamma\phi_{0u}^{-1}).$$

Then $J^s\phi$ is a local flow corresponding to a vector field on $J^s\pi_0$, denoted by J^sX and called the *s-jet-prolongation* of X. The vector field J^sX is both π_s-projectable and $\pi_{s,k}$-projectable for $0 \leq k < s$, and its π_s-projection, (resp. $\pi_{s,k}$-projection) is X_0 (resp. X, resp. J^kX, $1 \leq k \leq s-1$). In local coordinates, where

$$X = X^0(t)\,\frac{\partial}{\partial t} + X^a(t,x)\,\frac{\partial}{\partial x^a},$$

we have

$$J^sX = X^0(t)\,\frac{\partial}{\partial t} + X^a(t,x)\,\frac{\partial}{\partial x^a} + \sum_{k=1}^{s} X^a_k\,\frac{\partial}{\partial x^a_k}\,,$$

where the functions X^a_k are defined by the recurrence formula

$$X^a_k = \frac{dX^a_{k-1}}{dt} - x^a_k \frac{dX^0}{dt}, \quad 1 \leq k \leq s. \tag{1}$$

By a *distribution* on $J^s\pi_0$ we mean a mapping $\mathcal{D}$ assigning to every point $z = J^s_t\gamma \in J^s\pi_0$ a vector subspace $\mathcal{D}(z)$ of $T_z J^s\pi_0$. The dimension of $\mathcal{D}(z)$ is then called *rank* of the distribution $\mathcal{D}$ at z. We say that $\mathcal{D}$ has *constant rank* if the function $z \to \dim \mathcal{D}(z)$ is constant. A distribution of a constant rank on $J^s\pi_0$ is thus a subbundle of the tangent bundle over $J^s\pi_0$. The rank of a distribution is a lower semi-continuous function. Consequently, a distribution of *constant* rank is smooth (i.e. is spanned by smooth vector fields) if and only if its annihilator is smooth (i.e. is spanned by smooth differential 1-forms), however, *a distribution of nonconstant rank with continuous annihilator is not continuous* (and vice versa).

A distribution on $J^s\pi_0$ is called *vertical* if it is spanned by π_s-vertical vector fields, and *weakly horizontal* if it is complementary to a vertical distribution. The rank of a weakly horizontal distribution is ≥ 1 at each point, and may be non-constant. The π_s-vertical bundle over $J^s\pi_0$ is then called the *maximal vertical distribution*; by a *horizontal distribution* we understand a distribution that is complementary to the maximal vertical distribution. By definition, horizontal distributions have constant rank equal to one.

Given a distribution on $J^s\pi_0$, we shall deal with integral mappings that are *sections* of the fibred manifold π_s, and speak about *integral sections*. An integral section may be holonomic or not.

For differential forms on jet bundles the properties of *horizontality* and different kinds of *contactness* are fundamental. The following concepts have been introduced and systematically studied by D. Krupka since the early $1970's$ [58, 59, 62, 63].

Let $\bigwedge^q(J^s\pi_0)$, $q \geq 0$, denote the module of q-forms on $J^s\pi_0$ over the ring of functions (for $q = 0$ we have smooth functions on $J^s\pi_0$).

A form $\eta \in \bigwedge^q(J^s\pi_0)$ is called π_s*-horizontal* if $i_X\eta = 0$ for every π_s-vertical vector field X on $J^s\pi_0$. A form $\eta \in \bigwedge^q(J^s\pi_0)$ is called $\pi_{s,k}$*-horizontal*, $0 \leq k < s$, if $i_X\eta = 0$ for every $\pi_{s,k}$-vertical vector field X on $J^s\pi_0$. The module of π_s-horizontal (resp. of $\pi_{s,k}$-horizontal) q-forms on $J^s\pi_0$ is a submodule of $\bigwedge^q(J^s\pi_0)$ and is denoted by $\bigwedge^q_{\mathbb{R}}(J^s\pi_0)$ (resp. $\bigwedge^q_{J^k\pi_0}(J^s\pi_0)$). In our situation we get from the definition that η is π_s-horizontal if and only if in coordinates it is represented by $\eta = f\,dt$, where $f = f(t, x^a, \cdots, x^a_s)$; similarly, a $\pi_{s,k}$-horizontal q-form η is expressed by means of $dt, dx^a, \cdots, dx^a_k$ only (with the components dependent on all the $t, x^a, \cdots x^a_s$).

Let $\eta \in \bigwedge^q(J^s\pi_0)$. There is a unique horizonal form $h\eta \in \bigwedge^q(J^{s+1}\pi_0)$ such that for every section γ of π_0,

$$J^s\gamma^*\eta = J^{s+1}\gamma^* h\eta.$$

The mapping $h : \bigwedge^q(J^s\pi_0) \to \bigwedge^q(J^{s+1}\pi_0)$ is a homomorphism of the exterior algebras and is called *horizontalisation operator*. For $q = 0, 1$ this definition gives

$$hf = f \circ \pi_{s+1,s}$$

$$hdt = dt, \quad hdx^a = \dot{x}^a dt, \quad hd\dot{x}^a = \ddot{x}^a dt, \quad \cdots, \quad hdx^a_s = x^a_{s+1}dt,$$

$$h\,df = \frac{df}{dt}\,dt\,, \quad \frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x^a}\,\dot{x}^a + \frac{\partial f}{\partial \dot{x}^a}\,\ddot{x}^a + \cdots + \frac{\partial f}{\partial x^a_s}\,x^a_{s+1}.$$

For the sake of brevity of notations we shall also write

$$\frac{\bar{d}f}{dt} = \frac{df}{dt} - \frac{\partial f}{\partial x^a_s}\, x^a_{s+1} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x^a}\, \dot{x}^a + \frac{\partial f}{\partial \dot{x}^a}\, \ddot{x}^a + \cdots + \frac{\partial f}{\partial x^a_{s-1}}\, x^a_s,$$

and call it the "cut" total derivative.

Let now $s \geq 1$. A form $\eta \in \bigwedge^q(J^s\pi_0)$ is called *contact* if

$$J^s\gamma^*\,\eta = 0$$

for every section γ of π_0. Obviously, η is contact iff $h\eta = 0$. On our fibred manifolds every q-form for $q \geq 2$ is contact. A function f is contact iff $f = 0$. Contact forms form a closed ideal in the exterior algebra on $J^s\pi_0$ locally generated by the 1-forms

$$\omega^a = dx^a - \dot{x}^a dt, \quad \dot{\omega}^a = d\dot{x}^a - \ddot{x}^a dt, \cdots, \quad \omega^a_{s-1} = dx^a_{s-1} - x^a_s dt, \tag{2}$$

and their exterior derivatives.

Every vector field on $J^r\pi_0$ that is a prolongation of a vector field from $\mathbb{R} \times M$ (i.e., of the form J^rX) is a *symmetry of the contact ideal* on $J^r\pi_0$, i.e. transfers contact forms into contact forms. A symmetry of the contact ideal, however, need not be a prolongation of a projectable vector field on $\mathbb{R} \times M$ [67, 68].

Let $q \geq 1$, and let $\eta \in \bigwedge^q(J^s\pi_0)$ be a *contact* form. We say that η is *one-contact* if for every π_s-vertical vector field X on $J^s\pi_0$ the $(q-1)$-form $i_X\eta$ is π_s-horizontal; we say that η is *k-contact*, $2 \leq k \leq q$, if $i_X\eta$ is $(k-1)$-contact. The following theorem due to Krupka [62] describes the structure of differential forms on fibred manifolds.

Theorem 2.1 *(Krupka 1983) Every q-form η on $J^s\pi_0$, $s \geq 0$, admits the unique decomposition*

$$\pi^*_{s+1,s}\,\eta = h\eta + p_1\eta + \cdots + p_q\eta$$

where $p_i\eta$ is a i-contact q-form on $J^{s+1}\pi_0$, $1 \leq i \leq q$.

The form $p_i\eta$ is called the *i-contact part* of η. In view of the above theorem we shall consider operators $p_i : \bigwedge^q(J^s\pi_0) \to \bigwedge^q(J^{s+1}\pi_0)$, $1 \leq i \leq q$, assigning to every form its i-contact part. Since we consider the base manifold one-dimensional, we have $p_i\eta = 0$ for $i < q-1$. A contact q-form is called *strongly contact* if $\pi^*_{s+1,s}\,\eta = p_q\eta$.

Contact 1-forms on $J^s\pi_0$ annihilate a distribution of constant corank ms, called *contact* or *Cartan distribution* and denoted by $\mathcal{C}_{\pi_s}$. This distribution is not completely integrable. It is equivalently spanned by $m+1$ vector fields

$$\frac{\partial}{\partial t} + \sum_{j=0}^{s-1} x^b_{j+1}\frac{\partial}{\partial x^b_j}, \quad \frac{\partial}{\partial x^a_s}.$$

The contact distribution is important in distinguishing holonomic sections among all sections of π_s: *A section δ of π_s is holonomic* (i.e. $\delta = J^s\gamma$ for a section γ of π_0) *if and only if δ is an integral section of the contact distribution on $J^s\pi_0$.*

π_s-horizontal subdistributions of the contact distribution $\mathcal{C}_{\pi_s}$ are called *semispray distributions*. Accordingly, vector fields spanning semispray distributions, i.e., vector fields belonging to the contact distribution that are everywhere non-vertical are called *semisprays*. In local coordinates every semispray takes the form

$$g\Big(\frac{\partial}{\partial t}+\sum_{j=0}^{s-1}x^a_{j+1}\frac{\partial}{\partial x^a_j}+f^a\frac{\partial}{\partial x^a_s}\Big),$$

where g and f^a are functions of $(t,x^i,\dots,x^i_s)$, $g\neq 0$ at each point of $J^s\pi_0$. In this way *holonomic sections are characterised as integral sections of semisprays.*

A semispray distribution on $J^s\pi_0$ defines a splitting of the tangent bundle $TJ^s\pi_0$ into a direct sum of the horizontal and π_s-vertical subbundles. In this way, certain (nonlinear) connections are related to semispray distributions.

A general setting for many constructions connected with differential equations in jet bundles requires the concepts of a jet field and jet connection (i.e. a nonlinear connection on a fibred manifold), introduced in 1950 by Ehresmann [31, 33] and further developed by Mangiarotti, Modugno, Saunders, Vondra and others [93, 102, 129, 130, 153]. In our exposition, the concept of a *semispray connection* will play an important role.

A *second-order semispray connection* on the fibred manifold π_0 is a (local) section w of the projection $\pi_{2,1}:\mathbb{R}\times T^2M\to\mathbb{R}\times TM$. In coordinates the definition $\pi_{2,1}\circ w=\mathrm{id}_{\mathbb{R}\times TM}$ of a semispray connection w takes the form

$$t\circ w=t,\quad x^a\circ w=x^a,\quad \dot x^a\circ w=\dot x^a,\quad \ddot x^a\circ w=f^a.$$

The functions $f^a(t,x^i,\dot x^i)$ are called *components* of the connection w. Note that under coordinate transformations they transform like the coordinates $\ddot x^a$.

A (local) section γ of π_0 is called a *geodesic* (or a *path*) of a semispray connection w if it satisfies the equation

$$w\circ J^1\gamma=J^2\gamma.$$

In coordinates this turns out to be a system of $m=\dim M$ *second order ordinary differential equations*

$$\ddot x^a=f^a(t,x^i,\dot x^i),\quad 1\le i\le m,$$

for (the components of) sections of π_0, that is, of curves $\mathbb{R}\to M$.

As mentioned above, a second order semispray connection w is related with a semispray distribution on $\mathbb{R}\times TM$; the distribution is spanned by the (global) semispray

$$\Gamma=\frac{\partial}{\partial t}+\dot x^a\frac{\partial}{\partial x^a}+f^a\frac{\partial}{\partial\dot x^a},$$

or annihilated by $2m$ (local) one-forms

$$\omega^a=dx^a-\dot x^a dt,\quad \dot\omega^a_\Gamma=d\dot x^a-f^a dt.$$

Note that geodesics of the connection w are integral curves of the semispray Γ.

Remark 2.2 Contact 1-forms (2) can be completed to a basis of linear forms that is well adapted to the fibred structure. In what follows, we shall often use for a local expression of forms on $E=\mathbb{R}\times TM$ the adapted basis $(dt,\omega^a,d\dot x^a)$ instead of the canonical basis $(dt,dx^a,d\dot x^a)$, and similarly, for forms on $\mathbb{R}\times T^2M$, the adapted basis $(dt,\omega^a,\dot\omega^a,d\ddot x^a)$ instead of $(dt,dx^a,d\dot x^a,d\ddot x^a)$.

2.3 Dynamical forms

Systems of ordinary differential equations on fibred manifolds can be represented by so-called *dynamical forms* and their associated exterior differential systems (EDS), generated by a distribution on the evolution space. The integration problem and the structure of solutions, however, are in general complicated: the distribution is annihilated by a system of smooth 1-forms and may be of nonconstant rank, i.e. spanned by *non-continuous vector fields*. Moreover, if it happens to have a constant rank, it often is not completely integrable. Hence Frobenius or Sussmann-Viflyantsev theorems [137, 147, 148] cannot be used to obtain integrability conditions, and standard integration methods cannot be applied. On the other hand, a natural classification of SODEs on the basis of properties of their associated EDS appears. This is a tool to study the structure of solutions, as well as to obtain integration techniques for singular SODEs.

As above, we consider a fibred manifold $\pi_0 : \mathbb{R} \times M \to \mathbb{R}$, endowed with local coordinates adapted to the product structure, and the jet prolongations of π_0. We shall follow techniques and results by Krupková (see a series of papers [71, 72, 74, 83, 87] and the book [82] for more details and also for higher than second-order equations).

Let ε be a 2-form on $\mathbb{R} \times T^2M$. ε is called *dynamical form* if it is 1-contact and horizontal with respect to the projection onto $\mathbb{R} \times M$ ($\pi_{2,0}$-horizontal).

The above conditions mean that in local coordinates (with respect to the canonical and adapted basis, respectively) a second-order dynamical form reads

$$\varepsilon = E_a dx^a \wedge dt = E_a \omega^a \wedge dt,$$

where the components E_a are allowed to depend upon $t, x^i, \dot{x}^i, \ddot{x}^i$.

A section γ of π_0 is called *path* of a dynamical form $\varepsilon \in \bigwedge^2(\mathbb{R} \times T^2M)$ if

$$\varepsilon \circ J^2\gamma = 0. \tag{3}$$

Equation (3) is referred to as *equation of paths of* ε; in coordinates it becomes the following *system of* m *second-order ODEs for* m *components* $\gamma^i = x^i \circ \gamma$ *of sections of* $\pi_0 : \mathbb{R} \times M \to \mathbb{R}$:

$$E_a(t, x^i, \dot{x}^i, \ddot{x}^i) \circ J^2\gamma = 0, \quad 1 \leq a \leq m. \tag{4}$$

To simplify the notation, from now on we will write

$$E_a(t, x^i, \dot{x}^i, \ddot{x}^i) = 0. \tag{5}$$

It is worth noting that the above equations represent a very general class of ODEs: in particular, they *need not admit a normal (vector field) form*

$$\ddot{x}^a = f^a(t, x^i, \dot{x}^i). \tag{6}$$

The study of geometric structures related with equations of paths of dynamical forms will be our next goal. The key to the analysis of these equations is a relation between dynamical forms and exterior differential systems.

Let $W \subset \mathbb{R} \times T^2M$ be an open set and φ a 2-*contact* form on W. For a dynamical form $\varepsilon \in \bigwedge^2(\mathbb{R} \times T^2M)$ put

$$\alpha = \varepsilon|_W + \varphi,$$

and consider the exterior differential system on W, generated by the system Δ^0_φ of 1-forms

$$i_X\alpha, \quad \text{where } X \text{ runs over } \pi_2\text{-vertical vector fields on } W. \tag{7}$$

The characteristic distribution of this exterior differential system (i.e. the annihilator of Δ^0_φ) is called a *dynamical distribution* of ε and denoted by Δ_φ. The following proposition holds:

Proposition 2.3 *Let ε be a dynamical form on $\mathbb{R} \times T^2M$. Then for any 2-contact form $\varphi \in \bigwedge^2(W)$, paths of ε in W coincide with holonomic integral sections of the exterior differential system defined by Δ^0_φ, i.e. with holonomic integral sections of the dynamical distribution Δ_φ. This means that equations*

$$\varepsilon \circ J^2\gamma = 0$$

on W are equivalent with the equations

$$J^2\gamma^* i_X(\varepsilon + \varphi) = 0, \quad \textit{for every } \pi_2\textit{-vertical vector field } X \textit{ on } W.$$

In view of the above proposition we can pose the *Cauchy initial problem* as a problem of finding all holonomic integral sections of an associated EDS, passing through a given point $x \in \mathbb{R} \times T^2M$.

In the sequel we shall exclusively deal with equations that admit representation by *first-order* EDS.

We call a dynamical form $\varepsilon \in \bigwedge^2(\mathbb{R} \times T^2M)$ *pertinent* with respect to $\mathbb{R} \times TM$, if around every point in $J^2(\mathbb{R} \times T^2M)$ there is an open set W and a 2-contact 2-form φ defined on W such that the 2-form $\alpha = \varepsilon|_W + \varphi$ is $\pi_{2,1}$-projectable. We denote by $[\alpha]$ the family of all such "local extensions" of ε and call it the *Lepage class* of ε.

Pertinent dynamical forms can be easily described in local coordinates:

Proposition 2.4 *A dynamical form $\varepsilon \in \bigwedge^2(\mathbb{R}\times T^2M)$ is pertinent with respect to $\mathbb{R}\times TM$ if and only if its components E_a are affine functions in the variables $\ddot{x}^i$:*

$$E_a = A_a(t, x, \dot{x}) + B_{ab}(t, x, \dot{x})\ddot{x}^b. \tag{8}$$

Then the elements of the Lepage class $[\alpha]$ of ε take the form

$$\begin{aligned}\alpha &= E_a\omega^a \wedge dt + F_{ab}\omega^a \wedge \omega^b + B_{ab}\omega^a \wedge \dot{\omega}^b \\ &= A_a\omega^a \wedge dt + F_{ab}\omega^a \wedge \omega^b + B_{ab}\omega^a \wedge d\dot{x}^b,\end{aligned}$$

where $F_{ab}(t, x, \dot{x})$ are arbitrary functions, skew-symmetric in the indices a, b.

Note that

$$\begin{aligned}\varepsilon &= p_1\alpha, \\ \varphi &= p_2\alpha = F_{ab}\omega^a \wedge \omega^b + B_{ab}\omega^a \wedge \dot{\omega}^b,\end{aligned}$$

and that $\alpha \in \bigwedge^2(U)$ belongs to the Lepage class $[\alpha]$ of a pertinent dynamical form ε if and only if

$$\alpha = \omega^a \wedge (A_a dt + B_{ab}d\dot{x}^b) + F,$$

where F is a 2-contact 2-form on $U \subset \mathbb{R} \times TM$.

Remark 2.5 The Lepage class $[\alpha]$ of ε contains distinguished representatives

$$\alpha_\varepsilon = A_a\omega^a \wedge dt + \frac{1}{4}\Big(\frac{\partial A_a}{\partial \dot{x}^b} - \frac{\partial A_b}{\partial \dot{x}^a}\Big)\omega^a \wedge \omega^b + B_{ab}\omega^a \wedge d\dot{x}^b, \tag{9}$$

completely determined by the dynamical form ε, each defined on the domain of the coordinates $(t, x^i, \dot{x}^i)$. In general, these 2-forms are only local: they do not give rise to a global 2-form on the evolution space. However, as we shall see later, for an important class of equations, the so-called *semi-variational equations*, the above local representatives "unify" into a *global* 2*-form on the evolution space* $\mathbb{R} \times TM$.

For the dynamical distribution Δ_φ (defined on $U = \pi_{2,1}(W) \subset \mathbb{R} \times TM$) we now get

$$\begin{aligned}\Delta_\varphi &= \text{annih}\{i_X\alpha,\ X \text{ runs over } \pi_1\text{-vertical vector fields on } U\}\\ &= \text{annih}\{A_a dt + 2F_{ab}\omega^b + B_{ab}d\dot{x}^b,\ B_{ab}\omega^b\}.\end{aligned}$$

This means that at each point of U, rank $\Delta_\varphi \geq 1$.

Let $\mathcal{C}_{\pi_1}$ be the contact distribution on $\mathbb{R} \times TM$. If Δ_φ is a dynamical distribution of ε defined on U, consider the distribution $\Delta_\varphi \cap \mathcal{C}_{\pi_1}$. We can see that for every $x \in \mathbb{R} \times TM$ the vector space $(\Delta_\varphi \cap \mathcal{C}_{\pi_1})(x)$ is the same for all φ defined in a neighbourhood of x, that is, the local distributions define a *unique global distribution*. It is denoted $\mathcal{D}_\varepsilon$ and called *evolution distribution* of ε. We have

$$\mathcal{D}_\varepsilon = \text{annih}\{A_a dt + B_{ab}d\dot{x}^b,\ \omega^a\}, \tag{10}$$

so that rank $\mathcal{D}_\varepsilon \geq 1$, and the rank may be nonconstant on the evolution space.

Paths of pertinent dynamical forms are solutions of the system of m SODEs

$$B_{ab}(t, x^i, \dot{x}^i)\,\ddot{x}^b + A_a(t, x^i, \dot{x}^i) = 0, \tag{11}$$

and are characterised as follows:

Proposition 2.6 *Let* $\varepsilon \in \bigwedge^2(\mathbb{R} \times T^2M)$ *be a dynamical form, pertinent with respect to* $\mathbb{R} \times TM$.

(1) *Prolongations of paths of* ε *coincide with integral sections of the evolution distribution* $\mathcal{D}_\varepsilon$.

(2) *If* $\alpha = \varepsilon + \varphi$ *is a projectable extension of* ε *defined on an open subset* $U \subset \mathbb{R} \times TM$ *then holonomic integral sections of the dynamical distribution* Δ_φ *coincide with prolongations of paths of* ε *in* U.

Corollary 2.7 *Given a* $(\mathbb{R} \times TM)$*-pertinent dynamical form* ε *and (any) its projectable extension* α *on an open set* $U \subset \mathbb{R} \times TM$*, equations for paths of* ε *(3) have the following equivalent coordinate-free representations:*

(1) $J^1\gamma^* i_X\alpha = 0$ *for every* π_1*-vertical vector field* X *on* U.

(2) $J^1\gamma^* i_X\alpha = 0$ *for every vector field* X *on* U.

(3) *Solutions of (3) in* U *coincide with holonomic integral sections of the characteristic vector fields of the* 2*-form* α*, i.e. vector fields* X *that are solutions of the equation*

$$i_X\alpha = 0.$$

Note that equation $i_X\alpha = 0$ may have solutions X that are not continuous and are not semisprays.

2.4 Geometric classification of SODEs

The *Cauchy initial problem* for equations (11) means finding all integral sections of the distribution $\mathcal{D}_\varepsilon$ passing through a given point in the evolution space. If the Cauchy problem is solved around every point, one can get the family of all maximal one-dimensional submanifolds in $\mathbb{R} \times TM$, that are images of integral sections $\mathcal{D}_\varepsilon$; for brevity we speak about the global *dynamical picture* for the given SODEs.

It is clear that for a concrete system of equations the dynamical picture may be complicated and hard to find. A first step is to classify the evolution distributions.

We introduce the matrices

$$\mathcal{B} = (B_{ab}), \quad (\mathcal{B}|\mathcal{A}) = (B_{ab}\ A_a)$$

where $1 \le a, b, \le m$, and the indices a and b number rows and columns, respectively.

Given an $(\mathbb{R} \times TM)$*-pertinent dynamical form* $\varepsilon \in \bigwedge^2(\mathbb{R} \times T^2M)$*, its evolution distribution* $\mathcal{D}_\varepsilon$ *has the following properties* [72, 74, 77]

Proposition 2.8 *rank* $\mathcal{D}_\varepsilon = 1$ *if and only if the matrix* $\mathcal{B}$ *is regular. In this case,* $\mathcal{D}_\varepsilon$ *is spanned by the following global semispray:*

$$\Gamma = \frac{\partial}{\partial t} + \dot{x}^a \frac{\partial}{\partial x^a} - B^{ab} A_b \frac{\partial}{\partial \dot{x}^a},$$

where $(B^{ab}) = \mathcal{B}^{-1}$*, and the equations for paths of* ε *(11) have an equivalent normal form*

$$\ddot{x}^a = -B^{ab} A_b. \tag{12}$$

$\mathcal{D}_\varepsilon$ *is the horizontal distribution of the global semispray connection* $w : \mathbb{R} \times TM \to \mathbb{R} \times T^2M$ *that is a unique solution of the equation*

$$w^* \varepsilon = 0. \tag{13}$$

Proposition 2.9 *Let* $x \in \mathbb{R} \times TM$ *be a point. The following conditions are equivalent:*

(1) $\mathcal{D}_\varepsilon$ *is weakly horizontal at* x.

(2) *In a neighbourhood of* x *there is a semispray* X *such that* $X(x) \in \mathcal{D}_\varepsilon(x)$.

(3) *rank* $\mathcal{B}(x) =$ *rank* $(\mathcal{B}|\mathcal{A})(x)$.

Proposition 2.10 *Let* $x \in \mathbb{R} \times TM$ *be a point. The following conditions are equivalent:*

(1) *In a neighbourhood of* x *there exist semisprays* $X_1, \dots, X_r$ *such that the vectors* $X_1(x), \dots, X_r(x)$ *are linearly independent and span* $\mathcal{D}_\varepsilon(x)$.

(2) *rank* $\mathcal{B}(x) =$ *rank* $(\mathcal{B}|\mathcal{A})(x) = m + 1 - r$.

Since the rank of $\mathcal{D}_\varepsilon$ need not be constant, the above semisprays need not be continuous.

Note that from the transformation properties of the components E_a of ε it is clear that *the rank of the matrices* $\mathcal{B}$ *and* $(\mathcal{B}|\mathcal{A})$ *is invariant with respect to the change of local coordinates* (on M, as we consider throughout, or of fibred coordinates on the total space for more general fibred manifolds).

Put

$$\mathcal{P} = \{x \in \mathbb{R} \times TM \mid \operatorname{rank} \mathcal{B}(x) = \operatorname{rank} (\mathcal{B}|\mathcal{A})(x)\} \subset \mathbb{R} \times TM$$

$\mathcal{P}$ can be regarded as *the set of admissible initial conditions for the* SODE*s* (11). One should notice that it *need not be a submanifold* of the evolution space. Its complement, i.e. $(\mathbb{R} \times TM) - \mathcal{P}$ then has the meaning of *constraints* on the motion and will be referred to as the set of *primary semispray constraints*.

Remark 2.11 The constraints in $\mathbb{R} \times TM$ defined above are entirely related with the given SODEs, and as such are of "internal" origin. They should be distinguished from "external constraints" given as additional restrictions on the evolution space, e.g. as a fixed submanifold in $\mathbb{R} \times TM$, with no reference to concrete equations (cf. holonomic and non-holonomic constraints in classical mechanics). Equations on manifolds with external constraints will not be studied in this work.

We have the following distinguished classes of dynamical forms [77, 84]:

Definition 2.12 We say that a pertinent dynamical form $\varepsilon \in \bigwedge^2(\mathbb{R} \times T^2M)$, respectively, a system of SODEs (11)

(1) is *regular* if $\operatorname{rank} \mathcal{D}_\varepsilon = 1$,

(2) is *weakly regular* if $\mathcal{D}_\varepsilon$ is weakly horizontal and its rank is locally constant,

(3) has *primary semispray constraints* if $\mathcal{P} \neq \mathbb{R} \times TM$.

Now we are prepared to discuss the integration problem for SODEs (11). To solve the Cauchy problem at an initial point $x \in \mathbb{R} \times TM$ one has to find integral sections of the distribution $\mathcal{D}_\varepsilon$ passing through x. Alternatively, by Corollary 2.7, there is another possibility based on solving the equation $i_X\alpha = 0$. In both cases, however, the integration problem has the following steps:

(1) Find all (local) *vector fields* passing through x that belong to the distribution $\mathcal{D}_\varepsilon$, respectively, satisfy the characteristic equation $i_X\alpha = 0$.

(2) Find integral *sections* of these vector fields.

(3) Exclude nonholonomic solutions (in case of $\mathcal{D}_\varepsilon$ the procedure terminates at the second step).

The concrete application and result of the procedure essentially depend on which class of equations is considered.

Regular equations are the most simple and most frequently studied class of SODEs. As we have seen, they can be represented equivalently either in the *"covariant form"* (11) or in *"contravariant" (normal) form* (12). Solutions are integral sections of a *semispray*: this means that the dynamical picture is a one-dimensional horizontal foliation of the evolution space (at each point the Cauchy problem has a unique maximal solution). In the language of physics, any choice of initial conditions determines a unique global path in the evolution space; this property is called *Newtonian determinism*. For regular equations powerful integration methods are available, based on symmetries of the evolution distribution (for more details see eg. [92, 107]). Regular equations, however, have many specific properties that require individual attention. We shall postpone this study to the next sections.

Weakly regular equations have no primary semispray constraints so that all points in the evolution space are admissible initial conditions. By propositions 2.8 and 2.10 the evolution distribution is locally spanned by r semisprays. If $r > 1$ (the equations are not regular), this distribution is *not completely integrable*. As a result, the structure of solutions is complicated and is certainly not a straightforward generalisation of the one-dimensional foliation corresponding to regular equations (in particular, it is typically non-deterministic).

For integration of weakly regular equations one can utilise dynamical distributions Δ_φ better than the evolution distribution $\mathcal{D}_\varepsilon$. In particular, if $\alpha = \varepsilon + \varphi$ is a projectable extension of ε on $U \subset \mathbb{R} \times TM$ then as we have seen, integral sections of the evolution distribution coincide with holonomic integral sections of the dynamical distribution Δ_φ. In this sense, Δ_φ is an "enveloping" distribution for $\mathcal{D}_\varepsilon$. The point is that the distribution Δ_φ may be much easier than $\mathcal{D}_\varepsilon$ and its integration using known methods may be possible.

We say that α is *semiregular* if Δ_φ is weakly horizontal, of locally constant rank and completely integrable. Since for weakly regular equations $\mathcal{D}_\varepsilon$ is a weakly horizontal subdistribution of Δ_φ in U, the weak horizontality condition is automatically satisfied for any φ.

Let α be a semiregular projectable extension of ε on U. Then, as proved in [77, 84], the dynamical distribution Δ_φ is the *characteristic distribution of the* 2*-form* α and $\mathcal{D}_\varepsilon$ is its subdistribution spanned by semisprays. Thus, solutions of weakly regular equations in U are integral *sections* of semisprays X that are solutions of the equation

$$i_X \alpha = 0,$$

i.e. integral *curves* of semisprays X that satisfy

$$i_X \alpha = 0, \quad T\pi_1.X = \frac{\partial}{\partial t}.$$

Note that for weakly regular equations the equation $i_X \alpha = 0$ with semiregular α always has solutions that are not semisprays.

Since the distribution Δ_φ is completely integrable, its maximal integral manifolds define a *foliation* of the evolution space (we assume that ε is not regular, hence the dimension of the foliation is > 1). Every integral section of Δ_φ is an embedding of an open interval into a leaf of this foliation, and conversely, every section of π_1 lying on a leaf is an integral section of Δ_φ. One can use some of the known integration methods to find the leaves explicitly (see e.g. [92, 107] for Lie methods based on symmetries of distributions or [75, 113] for a method based on a generalised Liouville theorem); solutions of the given SODEs in U are then holonomic sections of the leaves of the foliation.

Equations with primary constraints are the most complicated class of SODEs to study. Obtaining the dynamical picture requires application of the so-called *geometric constraint algorithm* [77]. Without going into details this means studying solutions of the given SODEs in the subset $\mathcal{P}$: one has to investigate the evolution distribution – or, better a dynamical distribution Δ_φ – *restricted* to certain subsets of $\mathcal{P}$ that are *submanifolds* in $\mathbb{R} \times TM$. For a detailed exposition and examples of applications of this algorithm we refer to [82] and [84]. A typical dynamical picture that is obtained is non-deterministic:

there are points in the evolution space where there is no solution, as well as points (initial conditions) admitting a bunch of solutions passing through this point.

The name of the procedure refers to the "constraint algorithm" developed by Dirac as a heuristic technique to obtain Hamilton equations for singular Lagrangians in mechanics [29], and later explained within presymplectic geometry [38, 39, 40] and generalised to singular higher-order, time-dependent Lagrangians [27]. In [77, 84] the constraint algorithm has been revisited and further generalised to become a method of solving singular ODEs (integration of distributions of nonconstant rank that are annihilated by smooth 1-forms).

Remark 2.13 If $\alpha = \varepsilon + \varphi$ is a (local) projectable extension of ε we can consider, instead of the dynamical distribution Δ_φ, the *characteristic distribution* of the 2-form α, that is the distribution

$$\begin{aligned}\chi_\varphi &= \text{span}\{X \in \mathfrak{X}(U) \mid i_X\alpha = 0\} = \text{annih}\{i_X\alpha \mid X \in \mathfrak{X}(U)\}\\ &= \text{annih}\{A_a dt + 2F_{ab}\omega^b + B_{ab}d\dot{x}^b,\ B_{ab}\omega^b, A_b\omega^b\}.\end{aligned}$$

χ_φ is a subdistribution of Δ_φ, its rank ≥ 1, and

$$\text{rank}\,\alpha = \text{corank}\,\chi_\varphi. \tag{14}$$

It can be proved [74] (see also [77, 82, 84]) that the distributions Δ_φ and χ_φ have the same integral *sections*, and that they *coincide in all points where they are weakly horizontal*. Hence, at the points where they are distinct there is no solution of the given SODEs (such points are non-admissible initial conditions).

3 Variational structures in the theory of SODEs

In this section we shall be interested in SODEs arising as Euler-Lagrange equations of a variational functional. We recall the way in which the global first order variational formula for Lagrangian systems on fibred manifolds is obtained, and related important differential forms determined by a Lagrangian: the *Cartan form* and the *Euler-Lagrange form*. The Euler-Lagrange form appears as a global dynamical form representing Euler-Lagrange equations. The Euler-Lagrange mapping, assigning to a Lagrangian its Euler-Lagrange form then becomes one of the morphisms in an exact sequence of sheaves over a fibred manifold, called the *variational sequence*. Equipped with this material one can then study local and global problems in the calculus of variations, as well as explore various variational structures in the study of general differential equations. Our exposition of the main concepts of the calculus of variations in jet bundles is very brief and we refer the interested reader to the articles in this book by Krupka [66] and Vitolo [152], devoted to these topics, and Grigore [45] for discussion of applications in physics. Our main aim is then the exposition of the so-called inverse problem of the calculus of variations in covariant formulation, i.e. the question on the local and global existence of a Lagrangian for dynamical forms. Thus the reader can find in this section necessary and sufficient conditons for variationality of equations for paths of a dynamical form (5), their geometric meaning, and theorems on structure of variational equations. These theorems then lead to a generalisation of Riemannian and Finsler metric structures on manifolds, and have interesting physical consequences in classification of variational forces on manifolds with generalised metrics.

As above, we consider a fibred manifold $\pi_0 : \mathbb{R} \times M \to \mathbb{R}$, and its jet prolongations, and denote by (t, x^a), $1 \leq a \leq m$, local coordinates on $\mathbb{R} \times M$ adapted to the product structure.

3.1 The first variation formula

A *Lagrangian of order* r on π_0, where $r \geq 1$, is defined to be a *horizontal* 1-form λ on $J^r\pi_0$. This means that

$$\lambda = Ldt,$$

where L is a function on $J^r\pi_0$.

In what follows, we shall mostly consider first-order Lagrangians; in this case L is a function of the variables $(t, x^a, \dot{x}^a)$.

Let $\Omega = [a, b] \subset \mathbb{R}$, $a < b$, be an interval, and denote by $\mathcal{S}_\Omega(\pi_0)$ the set of sections of π_0 the domain of which is a neighbourhood of Ω. The function

$$\mathcal{S}_\Omega(\pi_0) \ni \gamma \to \int_\Omega J^1\gamma^*\lambda \in \mathbb{R} \tag{15}$$

is called the *action function* of the Lagrangian λ over Ω.

Consider now a section $\gamma \in \mathcal{S}_\Omega(\pi_0)$. If X is a *projectable* vector field on $\mathbb{R} \times M$ with projection X_0, and $\{\phi_u\}$, resp. $\{\phi_{0u}\}$ are the corresponding local one-parameter groups, we get a one-parameter family of sections, $\gamma_u = \phi_u\gamma\phi_{0u}^{-1}$, defined in a neighborhood of $\phi_{0u}(\Omega)$, and called *deformation of the section* γ *induced by* X. The arising function

$$\mathcal{S}_\Omega(\pi_0) \ni \gamma \to \left(\frac{d}{du}\int_{\phi_{0u}(\Omega)} J^1\gamma_u^*\lambda\right)_{u=0} = \int_\Omega J^1\gamma^*\,\mathcal{L}_{J^1X}\lambda \in \mathbb{R} \tag{16}$$

is called the *first variation of the action function* of λ over Ω, *induced by* X. Note that it is simply the action of the Lagrangian $\mathcal{L}_{J^1X}\lambda$ over Ω.

The above integral has to be decomposed *in an invariant* way into a sum of two terms: the first one representing equations for extremal sections (i.e., *equations for paths of a dynamical form*) and a boundary term. Taking into account proposition 2.3, 2.6 and remark 2.13 we can see that the decomposition of the Lie derivative of λ to $i_{J^1X}d\lambda + di_{J^1X}\lambda$ is not appropriate, since $p_1d\lambda$ is not a dynamical form. We note, however, that *the action does not change if to* λ *a contact form is added.* Moreover, since prolongations of projectable vector fields are symmetries of the contact ideal, such a change of the action has no effect to the variation of the action. Finally, the requirement on the first term means that the contact addition η to λ must be such that $p_1d(\lambda + \eta)$ is a *dynamical form.* The following theorem solves the problem of existence (and uniqueness) of η:

Theorem 3.1 *(Krupka 1973) Given a first-order Lagrangian* λ *there is a unique* 1*-form* θ_λ *on* $J^1\pi_0$ *such that*

(1) $\lambda = h\theta_\lambda$,

(2) $p_1d\theta_\lambda$ *is a dynamical form.*

θ_λ is called the *Cartan form*, or the *Lepage equivalent of* λ [58]. In local coordinates, with respect to the adapted and canonical basis, respectively, θ_λ reads

$$\theta_\lambda = Ldt + \frac{\partial L}{\partial \dot{x}^a}\omega^a = \Big(L - \frac{\partial L}{\partial \dot{x}^a}\dot{x}^a\Big)dt + \frac{\partial L}{\partial \dot{x}^a}dx^a. \tag{17}$$

We put

$$H = -L + \frac{\partial L}{\partial \dot{x}^a}\dot{x}^a, \quad p_a = \frac{\partial L}{\partial \dot{x}^a},$$

and call these functions the *Hamiltonian* and *momenta* related with the (first order) Lagrangian λ.

The dynamical form

$$\varepsilon_\lambda = p_1 d\theta_\lambda \tag{18}$$

is then called the *Euler-Lagrange form* of λ [58]. We have $\varepsilon_\lambda = E_a(L)\,\omega^a \wedge dt$, where the components, called *Euler-Lagrange expressions*, take the form

$$E_a(L) = \frac{\partial L}{\partial x^a} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}^a}. \tag{19}$$

We can see that the 2-form $d\theta_\lambda$ is a projectable extension of the Euler-Lagrange form ε_λ, i.e., it belongs to the Lepage class $[\alpha]$ of ε_λ, and it is a *global* and *closed* representative of this class.

With help of the Cartan form the first variation of the action of λ is invariantly decomposed as follows [58]:

$$\int_\Omega J^1\gamma^*\, \mathcal{L}_{J^1X}\lambda = \int_\Omega J^1\gamma^* i_{J^1X} d\theta_\lambda + \int_{\partial\Omega} J^1\gamma^* i_{J^1X}\theta_\lambda. \tag{20}$$

This formula is called the *integral first variation formula*. Its differential expression, i.e.

$$\mathcal{L}_{J^1X}\lambda = h i_{J^1X} d\theta_\lambda + h d i_{J^1X}\theta_\lambda \tag{21}$$

is then referred to as the *infinitesimal first variation formula.*

A section γ of π_0 is called an *extremal of* λ *on* Ω if the first variation of the action of λ on Ω vanishes for every *vertical* vector field X on $\mathbb{R} \times M$ with the support in $\pi_0^{-1}(\Omega)$ (such a vector field X is often called a *fixed-endpoints variation*). γ is called *extremal of* λ if is an extremal on every interval $\Omega = [a, b] \subset \mathbb{R}$, $a < b$.

Equations for extremals of a Lagrangian λ are called *Euler-Lagrange equations*. It can be proved that *they are precisely equations of paths of the Euler-Lagrange form* ε_λ.

3.2 Locally variational forms: the inverse problem in covariant formulation

Given a first-order Lagrangian, its Cartan form is also of order one and the Euler-Lagrange form is at most of order two. In the higher-order situation, however, if λ is of order r then θ_λ is of order at most $2r - 1$ and ε_λ is of order $2r$; it may be, however projectable onto some lower jet prolongation of π_0.

The mapping

$$\mathcal{E} : \lambda \to \varepsilon_\lambda,$$

from the sheaf of Lagrangians of order r to the sheaf of dynamical forms of order $2r$, assigning to a Lagrangian its Euler-Lagrange form, is called *Euler-Lagrange mapping*.

One of the crucial problems in the calculus of variations is to analyse *the kernel and the image of the Euler-Lagrange mapping*.

Lagrangians belonging to the kernel of the Euler-Lagrange mapping are called *null Lagrangians*. By this definition, λ is a null Lagrangian if its Euler-Lagrange form vanishes, i.e., $\varepsilon_\lambda = 0$.

It is clear that for Lagrangians that differ by a null Lagrangian the Euler-Lagrange forms are equal (up to a possible jet projection, since the Lagrangians may have different orders). We call such Lagrangians *equivalent*.

A dynamical form ε is called *locally variational* if it belongs to the image of the Euler-Lagrange mapping. A dynamical form ε on $J^s\pi_0$ is called *globally variational* if there is $r \geq 1$ and a Lagrangian λ on $J^r\pi_0$ such that $\varepsilon = \varepsilon_\lambda$ (possibly up to a jet projection). We stress that a locally variational form need not be globally variational.

In this section we shall investigate conditions of local variationality (existence of local Lagrangians) for global dynamical forms on $J^2\pi_0$. This problem is referred to as the *local inverse problem of the calculus of variations for dynamical forms*, or, the *inverse problem of the calculus of variations for* SODEs *in covariant formulation*.

We will leave aside the global inverse variational problem (conditions for the existence of a global Lagrangian) for now and return to it in the next section.

We have a fundamental theorem that connects locally variational forms (variational SODEs) with *closed 2-forms* (see [135, 71, 83] for the proof).

Theorem 3.2 *Let $\varepsilon \in \bigwedge^2(\mathbb{R} \times T^2M)$ be a dynamical form, $\varepsilon = E_a\omega^a \wedge dt$ its expression in a local adapted chart, and*

$$E_a(t, x^i, \dot{x}^i, \ddot{x}^i) = 0$$

the corresponding system of equations for paths. The following conditions are equivalent:

(1) *ε is locally variational.*

(2) *ε is pertinent with respect to $\mathbb{R} \times TM$ and the Lepage class $[\alpha]$ of ε has a global closed representative.*

(3) *There is a 2-contact 2-form φ on $\mathbb{R} \times T^2M$ such that $\varepsilon + \varphi$ is closed and projectable onto $\mathbb{R} \times TM$.*

(4) *There is a closed 2-form α on $\mathbb{R} \times TM$ such that $\varepsilon = p_1\alpha$.*

(5) *There is a 2-form α on $\mathbb{R} \times TM$ such that $\varepsilon = p_1\alpha$ and $p_2 d\alpha = 0$.*

(6) *The functions E_a, $1 \leq a \leq m$, satisfy the following identities:*

$$\frac{\partial E_a}{\partial \ddot{x}^b} - \frac{\partial E_b}{\partial \ddot{x}^a} = 0,$$

$$\frac{\partial E_a}{\partial \dot{x}^b} + \frac{\partial E_b}{\partial \dot{x}^a} - 2\frac{d}{dt}\frac{\partial E_b}{\partial \ddot{x}^a} = 0, \tag{22}$$
$$\frac{\partial E_a}{\partial x^b} - \frac{\partial E_b}{\partial x^a} + \frac{d}{dt}\frac{\partial E_b}{\partial \dot{x}^a} - \frac{d^2}{dt^2}\frac{\partial E_b}{\partial \ddot{x}^a} = 0.$$

(7) $\lambda = L\, dt$, *where*

$$L = x^a \int_0^1 E_a(t, ux^i, u\dot{x}^i, u\ddot{x}^i)du \tag{23}$$

is a local Lagrangian for ε.

To a locally variational form ε *the form mentioned in (2)-(5) is unique (denoted by* α_ε*). In coordinates,*

$$\alpha_\varepsilon = E_a\omega^a \wedge dt + \frac{1}{4}\Big(\frac{\partial E_a}{\partial \dot{x}^b} - \frac{\partial E_b}{\partial \dot{x}^a}\Big)\omega^a \wedge \omega^b + \frac{1}{2}\Big(\frac{\partial E_a}{\partial \ddot{x}^b} + \frac{\partial E_b}{\partial \ddot{x}^a}\Big)\omega^a \wedge \dot{\omega}^b. \tag{24}$$

Necessary and sufficient conditions of local variationality (22) are called *Helmholtz conditions*, a Lagrangian given by formula (23) is called *Tonti Lagrangian*, and the 2-form α_ε is called *Lepage* 2*-form*, or more precisely, *Lepage equivalent of* ε. Since every locally variational form is pertinent, we can write $E_a = A_a + B_{ab}\ddot{x}^b$ and express α_ε by means of A_a and B_{ab}. Then we obtain for α_ε precisely the formula (9) in remark 2.5.

The result that locally variational forms are in one-to-one correspondence with closed 2-forms was proved in [21] for regular SODEs and in [55, 71, 135] for general higher order ODEs (see also [72] for the regular case). An analogous result holds for first-order PDEs, as proved in [46]. In [62] and [85] a theorem of this kind was proved for higher order PDEs (in this case, the correspondence between locally variational forms and closed $(n+1)$-forms is no longer one-to-one). We shall meet versions of this theorem for regular SODEs in normal form separately in section 6.

The Helmholtz conditions may be written in different equivalent forms. At the first sight we notice the following one:

$$\frac{\partial E_a}{\partial \ddot{x}^b} - \frac{\partial E_b}{\partial \ddot{x}^a} = 0,$$
$$\frac{\partial E_a}{\partial \dot{x}^b} + \frac{\partial E_b}{\partial \dot{x}^a} - \frac{d}{dt}\Big(\frac{\partial E_a}{\partial \ddot{x}^b} + \frac{\partial E_b}{\partial \ddot{x}^a}\Big) = 0, \tag{25}$$
$$\frac{\partial E_a}{\partial x^b} - \frac{\partial E_b}{\partial x^a} - \frac{1}{2}\frac{d}{dt}\Big(\frac{\partial E_a}{\partial \dot{x}^b} - \frac{\partial E_b}{\partial \dot{x}^a}\Big) = 0.$$

If we rewrite the Helmholtz conditions in terms of the first-order functions A_a, B_{ab}, we get the following identities

$$B_{ab} = B_{ba},$$
$$\frac{\partial B_{ab}}{\partial \dot{x}^c} = \frac{\partial B_{ac}}{\partial \dot{x}^b},$$
$$\frac{\partial A_a}{\partial \dot{x}^b} + \frac{\partial A_b}{\partial \dot{x}^a} - 2\frac{\bar{d}B_{ab}}{dt} = 0, \tag{26}$$
$$\frac{\partial A_a}{\partial x^b} - \frac{\partial A_b}{\partial x^a} - \frac{1}{2}\frac{\bar{d}}{dt}\Big(\frac{\partial A_a}{\partial \dot{x}^b} - \frac{\partial A_b}{\partial \dot{x}^a}\Big) = 0$$

(where $\bar{d}/dt$ stands for the "cut" total derivative), and

$$\frac{\partial B_{ac}}{\partial x^b} - \frac{\partial B_{bc}}{\partial x^a} - \frac{1}{2}\frac{\partial}{\partial \dot{x}^c}\Big(\frac{\partial A_a}{\partial \dot{x}^b} - \frac{\partial A_b}{\partial \dot{x}^a}\Big) = 0. \tag{27}$$

It is known, however, that this last identity is a consequence of the former ones (see e.g. [118, 73]). Thus Helmholtz conditions (26) are equivalent with (22).

The uniqueness of the Lepage equivalent α_ε gives relation to Cartan forms of individual Lagrangians [71]:

Proposition 3.3 *Let ε be a locally variational form on $\mathbb{R}\times T^2M$, α_ε its Lepage equivalent. Then for every (possibly local) Lagrangian of order r for ε, the 2-form $d\theta_\lambda$ is projectable onto an open subset $U \subset \mathbb{R} \times TM$ and*

$$\alpha_\varepsilon|_U = d\theta_\lambda.$$

Since this also means that $\varepsilon_\lambda = 0 \Leftrightarrow \alpha_\varepsilon = 0 \Leftrightarrow d\theta_\lambda = 0 \Leftrightarrow \theta_\lambda =$ (locally) $df \Leftrightarrow \lambda = hdf$, i.e. $L = df/dt$, we immediately learn the *local structure of null Lagrangians*. The same result will be obtained once more in the next section by completely different methods.

The relation between locally variational forms and closed 2-forms expressed by proposition 3.3 is a key to the solution of the problem of *construction* of a Lagrangian to a locally variational form: indeed, if ε is locally variational then $\varepsilon = p_1\alpha$ where $d\alpha = 0$, so that a Lagrangian is obtained, by applying *Poincaré Lemma*, as the horizontal part of the one form $\rho = B\alpha$ (B is the standard homotopy operator). The contact structure, however, admits the introduction of another homotopy operator, denoted by A, that is more convenient, since it is adapted to the canonical decomposition of forms into contact components [62]. A has the following properties: given $\eta \in \bigwedge^q_{J^1\pi_0}(J^2\pi_0)$, $q \geq 2$, it holds

$$\eta = Ad\eta + dA\eta$$
$$Ap_k\eta = p_{k-1}A\eta, \quad k = q-1, q.$$

Applying A to the Lepage equivalent α_ε of a locally variational form ε, we get $A\alpha_\varepsilon = A\varepsilon + A\varphi$, where

$$\lambda = A\varepsilon = \Big(x^a \int_0^1 (E_a \circ \chi)du\Big)dt$$

is a horizontal form, hence a Lagrangian for ε; here the mapping $\chi : [0,1] \times W \to W$ is defined by $\chi(u,t,x^i,\dot{x}^i,\ddot{x}^i) = (t, ux^i, u\dot{x}^i, u\ddot{x}^i)$, where $W \subset \mathbb{R} \times T^2M$ is an appropriate open set. This is the way how the formula (23) for the Tonti Lagrangian appears.

Notice that the Tonti Lagrangian is a *second order* Lagrangian for ε. It has been proved, however, in [146] that *every higher order Lagrangian for ε is locally equivalent with a first order Lagrangian*.

A similar assertion holds also for higher-order locally variational forms in mechanics (i.e. for higher-order variational ordinary differential equations). However, in field theory (partial differential equations) reduction criteria are yet not known.

3.3 The Krupka variational sequence

In his papers [63, 64], Krupka introduced an exact sequence of sheaves, *the variational sequence*, in which the Euler-Lagrange mapping appears as a sequence morphism. By means of this sequence information about the local and global structure of the Euler-Lagrange mapping is obtained. As a surprising result, new, non-classical objects important for the calculus of variations and the theory of differential equations appear, and their local and global properties are described. We will mention only a few concepts in the simplest situation of first-order jet bundles over $\mathbb{R}$, especially the *Helmholtz mapping* and the *Helmholtz form* that are essential for the study of the inverse variational problem, and for investigation of SODEs on manifolds with variational metrics.

Compared with the variational bicomplex theory of Anderson, Dedecker, Takens, Tulczyjev, Vinogradov, and others [4, 5, 6, 24, 138, 149, 150, 151], the variational sequence is well-adapted to study *finite-order* problems. For more about variational sequences and variational bicomplexes the reader is referred to the above mentioned papers as well as to the survey paper by Vitolo in this book [152], to the Notes at the end of Chapter 5 of Olver's book [107] and the first chapter of Anderson and Thompson [7].

As above, consider a fibred manifold $\pi_0 : \mathbb{R} \times M \to \mathbb{R}$ and its jet prolongations. Let Ω_q be the sheaf q-forms on $J^1\pi_0$, $\Omega_{0,c} = \{0\}$, and $\Omega_{q,c}$ the sheaf of strongly contact q-forms on $J^1\pi_0$. Set $\Theta_q = \Omega_{q,c} + d\Omega_{q-1,c}$ where $d\Omega_{q-1,c}$ is the image sheaf of $\Omega_{q-1,c}$ by the exterior derivative d. It can be shown that one gets an exact sequence of soft sheaves

$$0 \to \Theta_1 \to \Theta_2 \to \Theta_3 \to \cdots ,$$

where the morphisms are the exterior derivative, i.e., a subsequence of the De Rham sequence

$$0 \to \mathbb{R} \to \Omega_0 \to \Omega_1 \to \Omega_2 \to \Omega_3 \to \cdots .$$

The quotient sequence

$$0 \to \mathbb{R} \to \Omega_0 \to \Omega_1/\Theta_1 \to \Omega_2/\Theta_2 \to \Omega_3/\Theta_3 \to \cdots$$

which is also exact, is called the first-order *variational sequence* on π_0 [63]. Note that elements of Ω_q/Θ_q are not forms, but *classes* of (local) first-order q-forms. We denote by $[\rho]_{\mathrm{v}}$ an element of Ω_q/Θ_q, that is the (variational) class of $\rho \in \Omega_q$. The quotient mappings are denoted by

$$\mathcal{E}_q : \Omega_q/\Theta_q \to \Omega_{q+1}/\Theta_{q+1}.$$

As proved by Krupka in 1990, the variational sequence is an acyclic resolution of the constant sheaf $\mathbb{R}$ over the total space of the fibred manifold π_0, i.e. $\mathbb{R} \times M$ in our situation. Hence, due to the abstract De Rham theorem, we get *the identification of the cohomology groups of the cochain complex of global sections of the variational sequence with the De Rham cohomology groups $H^q M$ of the manifold M.*

The quotient sheaves Ω_q/Θ_q are determined up to natural isomorphisms of Abelian groups. Thus the classes in Ω_q/Θ_q admit various equivalent characterisations. A simple analysis shows that the sections of the quotient sheaf Ω_1/Θ_1 can be identified with some *horizontal forms* $\lambda = Ldt$, i.e., with some *Lagrangians*; more precisely, with those (generally second-order) Lagrangians that arise by horizontalisation from first-order 1-forms. Elements of Ω_2/Θ_2 can be identified with some *dynamical forms* $\varepsilon = E_a\omega^a \wedge dt$ (that arise from first-order 2-forms by applying the operator p_1 and the factorisation by Θ_2). We say that Lagrangians, respectively, dynamical forms are *source forms* [138] for the quotient sheaf Ω_1/Θ_1, respectively, Ω_2/Θ_2. The quotient mapping

$$\mathcal{E}_1 : \Omega_1/\Theta_1 \to \Omega_2/\Theta_2$$

in this representation of the sheaves coincides with the *Euler-Lagrange mapping*, that is, on source forms, $\mathcal{E}_1(\lambda) = \varepsilon_\lambda$.

In general, source forms for the quotient sheaves Ω_q/Θ_q are $(q-1)$-contact q-forms arising by applying to q-forms the so-called *interior Euler–Lagrange operator* $\mathcal{I}$ [4, 5, 70].

The mapping

$$\mathcal{E}_2 : \Omega_2/\Theta_2 \to \Omega_3/\Theta_3$$

is called the *Helmholtz mapping*. In the source forms representation, the image of a dynamical form is a source 3-form,

$$\mathcal{E}_2(\varepsilon) = H_\varepsilon, \tag{28}$$

called the *Helmholtz form* of ε [63]. In coordinates where $\varepsilon = E_a\omega^a \wedge dt$ we get (see [65])

$$\begin{aligned} H_\varepsilon &= \frac{1}{2}\Big(\frac{\partial E_a}{\partial x^b} - \frac{\partial E_b}{\partial x^a} - \frac{1}{2}\frac{d}{dt}\Big(\frac{\partial E_a}{\partial \dot{x}^b} - \frac{\partial E_b}{\partial \dot{x}^a}\Big)\Big)\omega^b \wedge \omega^a \wedge dt \\ &+ \frac{1}{2}\Big(\frac{\partial E_a}{\partial \dot{x}^b} + \frac{\partial E_b}{\partial \dot{x}^a} - \frac{d}{dt}\Big(\frac{\partial E_a}{\partial \ddot{x}^b} + \frac{\partial E_b}{\partial \ddot{x}^a}\Big)\Big)\dot{\omega}^b \wedge \omega^a \wedge dt \\ &+ \frac{1}{2}\Big(\frac{\partial E_a}{\partial \ddot{x}^b} - \frac{\partial E_b}{\partial \ddot{x}^a}\Big)\ddot{\omega}^b \wedge \omega^a \wedge dt. \end{aligned}$$

The condition $\mathcal{E}_2(\varepsilon) = H_\varepsilon = 0$ for the elements of the quotient sheaves reads $\mathcal{E}_2([\eta]_{\mathrm{v}}) = 0$, and by exactness of the variational sequence means that there exists a class $[\rho]_{\mathrm{v}} \in \Omega_1/\Theta_1$ such that $[\eta]_{\mathrm{v}} = [d\rho]_{\mathrm{v}}$. The source form $\lambda = h\rho$ for $[\rho]_{\mathrm{v}}$ is then a (local) Lagrangian for ε, i.e. on the domain of λ, $\varepsilon = \varepsilon_\lambda$. Indeed, the conditions for the components of the Helmholtz form to vanish are the *Helmholtz conditions* for local variationality of dynamical forms. If in addition, $H^2M = \{0\}$, ρ may be chosen *globally* defined on $\mathbb{R} \times TM$, hence $\lambda = h\rho$ is a global Lagrangian for ε.

Summarizing we can see that the following assertions are true:

Theorem 3.4 *(Krupka 1990) A dynamical form ε is locally variational if and only if $H_\varepsilon = 0$. If $H^2M = \{0\}$ then every locally variational dynamical form is globally variational.*

Similarly, the condition $\mathcal{E}_1(\lambda) = \varepsilon_\lambda = 0$ (that is λ is locally a null Lagrangian) in terms of elements of the quotient sheaves reads $\mathcal{E}_1([\rho]_{\mathrm{v}}) = 0$. By exactness of the variational sequence there exists $f \in \Omega_0$ such that $[\rho]_{\mathrm{v}} = [df]_{\mathrm{v}}$ (note that for functions, $[f]_{\mathrm{v}} = f$), that is, $\lambda = hdf$. If in addition, $H^1M = \{0\}$, f may be chosen *globally* defined. If λ is a first-order Lagrangian then f must be a function on $\mathbb{R} \times M$ (df/dt depends upon $(t, x^i, \dot{x}^i)$ iff $\partial f/\partial \dot{x}^a = 0$, $1 \leq a \leq m$).

From $\lambda = hdf$ we see that df is the Cartan form of the Lagrangian hdf, i.e. $df = \theta_\lambda$. Hence, from the Poincaré Lemma, $f = B\theta_\lambda$, where B is the standard homotopy operator.

As a result, a theorem on null Lagrangians appears [63]:

Theorem 3.5 *A Lagrangian λ is locally null (i.e. $\varepsilon_\lambda = 0$) if and only if in a neighbourhood U of every point $x \in \mathbb{R} \times M$ there is a function f such that $\lambda = hdf$ over U.*

If $H^1 M = 0$ then every Lagrangian that is locally null is globally null, and takes the form $\lambda = hdf$ where f is a function on $\mathbb{R} \times M$.

It holds $f = B\theta_\lambda$, up to an additive constant, where B is the standard Poincaré homotopy operator.

The condition $\lambda = hdf$ in coordinates reads $L = df/dt$. If regarded as an *equation for f* it is called the *total derivative equation* (total divergence equation for n independent variables, $n > 1$).

Corollary 3.6 *The total derivative equation*

$$\frac{df}{dt} = F(t, x^i, \dot{x}^i)$$

is integrable if and only if F satisfies the following conditions:

$$\frac{\partial^2 F}{\partial \dot{x}^a\, \partial \dot{x}^b} = 0, \qquad \frac{\partial F}{\partial x^a} - \frac{d}{dt}\frac{\partial F}{\partial \dot{x}^a} = 0.$$

If we denote $F = g + h_a \dot{x}^a$, the solution is (up to an additive constant)

$$\begin{aligned} f &= B\Big(F\, dt + \frac{\partial F}{\partial \dot{x}^a}\omega^a\Big) = B(g\, dt + h_a dx^a) \\ &= t\int_0^1 g(\tau t, \tau x^i)\, d\tau + x^a \int_0^1 h_a(\tau t, \tau x^i) d\tau. \end{aligned}$$

We have seen that the classes in the variational sequence can be represented by source forms. A different but also important representation is realised by so-called *Lepage forms*. We already met Lepage 1-forms in the context of the first variation formula and Lepage 2-forms as global closed extensions of locally variational forms. The concept is generalised as follows: [70] A q-form η, $q \geq 1$, is called *Lepage q-form* if $p_q d\eta$ is a source form for Ω_q/Θ_q. If η is a Lepage q-form such that $\sigma = p_{q-1}\eta$ is a source form, we say that η is a *Lepage equivalent* of the source form σ.

Apparently, given a Lagrangian λ (source 1-form) its Lepage equivalent is the Cartan form θ_λ. For a locally variational dynamical form ε, the Lepage equivalent is the closed form α_ε (24), since $p_1\alpha = \varepsilon$ and $p_2 d\alpha$ is the zero source form. Lepage equivalents of general dynamical forms are characterised as follows:

Theorem 3.7 *Every dynamical form ε on $\mathbb{R} \times T^2 M$ has a unique global second-order Lepage equivalent α_ε.*

In coordinates,

$$\alpha_E = E_a \omega^a \wedge dt + \frac{1}{4}\Big(\frac{\partial E_a}{\partial \dot{x}^b} - \frac{\partial E_b}{\partial \dot{x}^a}\Big)\omega^a \wedge \omega^b + \frac{1}{2}\Big(\frac{\partial E_a}{\partial \ddot{x}^b} + \frac{\partial E_b}{\partial \ddot{x}^a}\Big)\omega^a \wedge \dot{\omega}^b.$$

If ε is pertinent, we get the formula

$$\begin{aligned} \alpha_\varepsilon = (A_a + B_{ab}\ddot{x}^b)\omega^a \wedge dt &+ \frac{1}{4}\Big(\frac{\partial A_a}{\partial \dot{x}^b} - \frac{\partial A_b}{\partial \dot{x}^a} + \Big(\frac{\partial B_{ac}}{\partial \dot{x}^b} - \frac{\partial B_{bc}}{\partial \dot{x}^a}\Big)\ddot{x}^c\Big)\omega^a \wedge \omega^b \\ &+ \frac{1}{2}\Big(B_{ab} + B_{ba}\Big)\omega^a \wedge \dot{\omega}^b. \end{aligned}$$

We call ε *semi-variational* if the Lepage equivalent α_ε is projectable onto $\mathbb{R} \times TM$. This is the case if and only if

$$B_{ab} = B_{ba}, \quad \frac{\partial B_{ac}}{\partial \dot{x}^b} = \frac{\partial B_{bc}}{\partial \dot{x}^a}, \tag{29}$$

hence α_ε takes the form (9) in remark 2.5.

3.4 The structure of variational and semi-variational SODEs

We shall show that *every second-order locally variational form gives rise to a geometrical structure on the manifold M, generalising the metric structure of Riemannian and Finsler geometry*. The *variational morphisms* and *variational metric structures* which appear can be studied separately: this concerns a generalised "kinetic energy" Lagrangian, the associated (nonlinear) metric connection, curvature, geodesics, and related topics, such as, for example, metrisability of semispray connections and relations with variationality. On the other hand, one can study various objects on manifolds endowed with a variational morphism (generalised metric). This gives us structure results for second order equations and particularly Lagrangians, a different formulation of the inverse variational problem as a question on variational forces, etc. The material in this section was developed in a series of papers [73, 76, 86], and the book [82] by Krupková, opening an area for geometry and physics that waits to be explored.

The starting point is the idea that Helmholtz conditions (26) can be viewed as equations for unknown functions A_a, and as such solved, providing insight into the structure of variational equations and their Lagrangians.

Theorem 3.8 *(Krupková 1987) Let $U \subset \mathbb{R}$ be an open interval, and $W \subset R^m$ an open ball in the centre in the origin. Let ε be a dynamical form on $J^2(U \times W)$, $\varepsilon = E_a \omega^a \wedge dt$. Consider the mapping*

$$\bar{\chi} : [0,1] \times J^1(U \times W) \to J^1(U \times W)$$

defined by

$$\bar{\chi}(v, t, x^a, \dot{x}^a) = (t, x^a, v\dot{x}^a).$$

The following conditions are equivalent:

(1) *ε is variational.*

(2) *It holds*

$$E_a = A_a + B_{ab}\ddot{x}^b$$

where B_{ab} are functions on $J^1(U \times V)$ satisfying

$$B_{ab} = B_{ba}, \quad \frac{\partial B_{ac}}{\partial \dot{x}^b} = \frac{\partial B_{bc}}{\partial \dot{x}^a},$$

and

$$A_a = \Gamma_{abc}\dot{x}^b\dot{x}^c + \Gamma_{ab0}\dot{x}^b + \alpha_{ab}\dot{x}^b + \beta_a,$$

where

$$\Gamma_{abc} = \frac{1}{2}\int_0^1 \Big(\frac{\partial B_{ab}}{\partial x^c} + \frac{\partial B_{ac}}{\partial x^b} - 2\frac{\partial B_{bc}}{\partial x^a}\Big) \circ \bar{\chi}\, dv + \int_0^1 \frac{\partial B_{bc}}{\partial x^a} \circ \bar{\chi}\, v\, dv\,, \tag{30}$$

$$\Gamma_{ab0} = \int_0^1 \frac{\partial B_{ab}}{\partial t} \circ \bar{\chi}\, dv,$$

and α_{ab}, β_a *are arbitrary functions satisfying the conditions*

$$\alpha_{ab} = -\alpha_{ba}, \quad \frac{\partial \alpha_{ab}}{\partial x^c} + \frac{\partial \alpha_{ca}}{\partial x^b} + \frac{\partial \alpha_{bc}}{\partial x^a} = 0, \quad \frac{\partial \beta_a}{\partial x^b} - \frac{\partial \beta_b}{\partial x^a} = \frac{\partial \alpha_{ab}}{\partial t}. \tag{31}$$

Corollary 3.9 *Given a locally variational form on* $\mathbb{R} \times T^2M$ *then its components have the following structure:*

$$E_a = B_{ab}\ddot{x}^b + \Gamma_{abc}\dot{x}^b\dot{x}^c + \Gamma_{ab0}\dot{x}^b + \alpha_{ab}\dot{x}^b + \beta_a,$$

where

$$B_{ab} = B_{ba}, \quad \frac{\partial B_{ac}}{\partial \dot{x}^b} = \frac{\partial B_{bc}}{\partial \dot{x}^a}, \tag{32}$$

the functions Γ_{abc} *and* Γ_{ab0} *are uniquely determined by the* B_{ij}*'s by (30), and* α_{ab}, β_a *are functions of* (t, x^i)*, satisfying the identities (31).*

We can see that (31) are the Helmholtz conditions for the first order dynamical form $\phi = \phi_a\omega^a \wedge dt$, where

$$\phi_a = \alpha_{ab}\dot{x}^b + \beta^a,$$

i.e. they mean that ϕ is locally variational. Equivalently, the 2-form

$$\alpha_\phi = \beta_a dx^a \wedge dt + \frac{1}{2}\alpha_{ab}dx^a \wedge dx^b = \phi_a\omega^a \wedge dt + \frac{1}{2}\alpha_{ab}\omega^a \wedge \omega^b$$

on $U \times W$ is closed (cf. theorem 3.2); α_ϕ is the closed extension of ϕ. This means however that by the Poincaré Lemma, locally $\alpha_\phi = d\nu$, where $\nu = A\alpha_\phi = f_b dx^b + g dt$ with

$$f_b = x^a \int_0^1 (\alpha_{ab} \circ \chi)\, u\, du, \quad g = x^a \int_0^1 \beta_a \circ \chi\, du.$$

Hence, $h\nu = A\phi = V dt$, where

$$V = x^a \int_0^1 (\alpha_{ab}\dot{x}^b + \beta_a) \circ \chi\, du = f_b\dot{x}^b + g \tag{33}$$

is the Tonti Lagrangian for ϕ.

We can see that every locally variational form ε canonically splits into two parts: a second order dynamical form

$$\varepsilon_{\mathcal{B}} = \varepsilon - \phi,$$

uniquely determined by the matrix $\mathcal{B} = (B_{ab})$, and a first order (locally variational) dynamical form ϕ. By the following theorem this splitting is invariant with respect to change of local coordinates.

Theorem 3.10 *(Krupková 1994) The form $\varepsilon_{\mathcal{B}}$ is globally variational. The global Lagrangian for $\varepsilon_{\mathcal{B}}$ is defined on $\mathbb{R} \times TM$ and takes the form $\lambda_{\mathcal{B}} = -T dt$, where*

$$T = \dot{x}^a \dot{x}^b \int_0^1 \Big(\int_0^1 B_{ab} \circ \bar{\chi}\, dv \Big) \circ \bar{\chi}\, v\, dv. \tag{34}$$

Since every local Lagrangian for ε is

$$\lambda = A\varepsilon = A\varepsilon_{\mathcal{B}} + A\phi \;\sim\; \lambda_{\mathcal{B}} + A\phi,$$

we get [73]

Theorem 3.11 *Let $\varepsilon \in \bigwedge^2(\mathbb{R} \times T^2M)$ be a locally variational form. Every local first order Lagrangian for ε reads*

$$L = -T + V,$$

where T is a function on $\mathbb{R} \times TM$ uniquely determined by $\mathcal{B}$ by (34), and V is a (local) function affine in velocities (given by (33)).

If ε is a pertinent dynamical form on $\mathbb{R} \times T^2M$ then the transformation rules for the functions B_{ab} show that B_{ab} are components of a fibred morphism

$$\mathcal{B} : \mathbb{R} \times TM \to T_2^0 M$$

over the identity of M, of the fibred manifold $\mathbb{R} \times TM \to M$ to the bundle $T_2^0M \to M$ of type $(0, 2)$ tensors on M. Conversely, to every such a morphism one associates a (non-unique) pertinent dynamical form ε. Obviously, the morphism is regular iff ε is regular.

We say that a fibred morphism $\mathcal{B} : \mathbb{R} \times TM \to T_2^0M$ is *variational* if the family of associated dynamical forms contains a locally variational form. The Helmholtz conditions show that $\mathcal{B}$ is variational if and only if it satisfies the integrability conditions (32). Taking into account the above results, we can see that a variational morphism gives rise to a canonical locally variational dynamical form $\varepsilon_{\mathcal{B}}$ on $\mathbb{R} \times T^2M$, coming from the global Lagrangian $\lambda_{\mathcal{B}}$.

Equipped with the above observations, we can study SODEs on a *manifold endowed with a morphism*.

Given a manifold $(M, \mathcal{B})$, we say that a pertinent dynamical form ε on $\mathbb{R} \times T^2M$ is *$\mathcal{B}$-related* if

$$\Big(\frac{\partial E_a}{\partial \ddot{x}^b} \Big) = \mathcal{B}. \tag{35}$$

In this way we obtain on $(M, \mathcal{B})$ an *equivalence class* of all $\mathcal{B}$-related dynamical forms, containing a distinguished representative – the canonical dynamical form $\varepsilon_{\mathcal{B}}$. For every ε we then have the difference

$$\phi = \varepsilon - \varepsilon_{\mathcal{B}},$$

which is a dynamical form on $\mathbb{R} \times TM$; it is called a *(covariant) force* related with ε on the manifold $(M, \mathcal{B})$.

We can see that the concept of a $\mathcal{B}$-related dynamical form concerns only *semi-variational forms*, i.e. such that their Lepage equivalents α_ε are projectable onto the evolution space. In this sense the following results exclusively concern *semi-variational equations* [76, 86].

Proposition 3.12 *Let M be a manifold, $\mathcal{B} : \mathbb{R} \times TM \to T^0_2 M$ a variational fibred morphism over id_M. Let $\varepsilon \in \bigwedge^2(\mathbb{R} \times T^2 M)$ be a $\mathcal{B}$-related dynamical form. Then equations of paths of ε, $\varepsilon \circ J^2\gamma = 0$, split canonically, and take the form*

$$-\varepsilon_{\mathcal{B}} \circ J^2\gamma = \phi \circ J^1\gamma.$$

In coordinates,

$$\frac{\partial T}{\partial x^a} - \frac{d}{dt}\frac{\partial T}{\partial \dot{x}^a} = \phi_a,$$

or, explicitly,

$$B_{ab}\ddot{x}^b + \Gamma_{abc}\dot{x}^b\dot{x}^c + \Gamma_{ab0}\dot{x}^b = -\phi_a,$$

where T and Γ_{abc}, Γ_{ab0} are uniquely determined by the morphism $\mathcal{B}$.

Proposition 3.13 *Let ϕ be a force on $(M, \mathcal{B})$. The following conditions are equivalent:*

(1) *ϕ is $(\mathbb{R} \times M)$-pertinent.*

(2) *The Helmholtz form H_ϕ is horizontal with respect to the projection onto $\mathbb{R} \times M$.*

(3) *The components of ϕ satisfy*

$$\frac{\partial \phi_a}{\partial \dot{x}^b} + \frac{\partial \phi_b}{\partial \dot{x}^a} = 0.$$

(4) *The components of ϕ satisfy*

$$\phi_a = \alpha_{ab}\dot{x}^b + \beta_a, \quad \textit{where} \quad \alpha_{ab} = -\alpha_{ba}.$$

In terms of ε we get that *ϕ is $(\mathbb{R} \times M)$-pertinent if and only if ε satisfies the third set of the Helmholtz conditions* (26) (i.e., since ε is $\mathcal{B}$-related by assumption, it satisfies all but the last set of the Helmholtz conditions (26)).

A force ϕ on $(M, \mathcal{B})$ that is locally variational (as a first order dynamical form) is called a *variational force*. As we have seen, *a force is variational if and only if it is $(\mathbb{R} \times M)$-pertinent (i.e. affine in velocities) and satisfies the closure conditions (=first order Helmholtz conditions) (31).*

An important particular case appears when the considered SODEs are *regular*. Recall that in this case the matrix (35) is regular, and prolongations of paths of the dynamical form ε are integral curves of a semispray that is uniquely determined by ε, hence, equations for paths of ε have an equivalent *contravariant form*.

A regular fibred morphism $g : \mathbb{R} \times TM \to T^0_2 M$ is called a *time, position and velocity dependent metric*, or simply a *generalised metric* on M. If moreover, the morphism is variational, we speak about a *variational metric*.

On a manifold M with a variational metric g the canonical global Lagrangian $\lambda_g = -T dt$ (34) is *regular*. With reference to physics, T is called the *kinetic energy* of g, and the Lagrangian system determined by λ_g is called the *free particle* for the manifold (M, g). The corresponding SODEs then take the form

$$\frac{\partial T}{\partial x^a} - \frac{d}{dt}\frac{\partial T}{\partial \dot{x}^a} = 0, \tag{36}$$

or, explicitly,

$$\ddot{x}^a + \Gamma^a_{bc}\dot{x}^b\dot{x}^c + \Gamma^a_{b0}\dot{x}^b = 0, \tag{37}$$

where

$$\Gamma^a_{bc} = g^{ai}\Gamma_{ibc} = \frac{1}{2}\int_0^1 \Big(\frac{\partial g_{ib}}{\partial x^c} + \frac{\partial g_{ic}}{\partial x^b} - 2\frac{\partial g_{bc}}{\partial x^i}\Big) \circ \bar{\chi}\, dv + \int_0^1 \frac{\partial g_{bc}}{\partial x^i} \circ \bar{\chi}\, v\, dv\,,$$
$$\Gamma^a_{b0} = g^{ai}\Gamma_{ib0} = \int_0^1 \frac{\partial g_{ib}}{\partial t} \circ \bar{\chi}\, dv,$$

and they are *equations for geodesics* of the associated semispray connection w (defined by $w^*\varepsilon_g = 0$), respectively, of a semispray Γ spanning the dynamical distribution of ε_g.

Moreover, from theorem 3.8 we get the structure of *variational forces* on a manifold M with a generalised metric g [76]: namely, conditions (31) mean that *every admissible variational force is a Lorentz-type force*. This question was first solved for the case of the Euclidean metric (when g = the unit matrix) in [34] and [106]; in the latter paper the result was obtained also for a velocity dependent metric appearing in the 3-dimensional equations of motion in the special relativity theory.

If the morphism g is projectable onto M (meaning that g is identified with a usual *metric on* M), the (global) function T (34) turns to be the usual *kinetic energy*,

$$T = \frac{1}{2} g_{ab}\dot{x}^a\dot{x}^b,$$

and the connection coefficients read $\Gamma_{ab0} = 0$ and

$$\Gamma_{abc} = \frac{1}{2}\Big(\frac{\partial g_{ab}}{\partial x^c} + \frac{\partial g_{ac}}{\partial x^b} - \frac{\partial g_{bc}}{\partial x^a}\Big).$$

This means that Γ^a_{bc} are the Christoffel symbols of the metric g, and equations of the "free particle" (36) are *equations for geodesics of the Levi-Civita connection*. Given a force ϕ on the (pseudo)-Riemannian manifold (M, g), the corresponding equations take the familiar form of equations of motion in nonconservative mechanics.

If g is a *Finsler metric* on M, i.e. a fibred morphism from the slit tangent bundle $TM - 0 \to M$ to the bundle of metrics on M over the identity of M, satisfying the integrability and the homogeneity conditions

$$\frac{\partial g_{ab}}{\partial \dot{x}^c} = \frac{\partial g_{ac}}{\partial \dot{x}^b}, \qquad \frac{\partial g_{ab}}{\partial \dot{x}^c}\dot{x}^c = 0,$$

then we obtain *equations for geodesics of the Cartan connection* [69, 76]. Indeed, in (37), $\Gamma_{ab0} = 0$, and after a short computation, $\Gamma_{abc}\dot{x}^b\dot{x}^c = 2g_{ab}G^b$, where G^b are the *geodesic coefficients* on the Finsler manifold (M, g) (see e.g. [9] for introduction to Finsler geometry).

4 Symmetries and first integrals

Joining the exterior differential systems approach to variational equations presented in section 2 together with various differential forms related with the equations due to the existence of a Lagrangian, results in specific integration methods for variational SODEs. They all are based on a relation between symmetries of variational structures and first integrals of the equations. The most cited result here is the famous Emmy Noether's theorem, published in 1918, stating that to every symmetry of a Lagrangian there corresponds a conservation law. Since that time various generalisations and modifications of this theorem have been discovered, making the theory of symmetries one of the most interesting and important parts of the variational calculus. In the Liouville theorem this theory becomes a powerful integration method, more effective than the alternative integration method of Lie. The classical Liouville method, however, applies exclusively to regular Lagrangian systems, contrary to the Lie method that is an integration method for any completely integrable distribution. In accounting for singular (degenerate) Lagrangian systems, it turns out that integration based on symmetries is no longer so straightforward. In this section we present an extension of the classical theory of symmetries and related integration methods to non-regular Lagrangian systems, and also to differential equations that need not be variational, but are representable as equations for the characteristics of a closed 2-form of a constant rank.

4.1 Exterior differential systems related with variational SODEs: regular Lagrangians revisited

As we have seen, second-order variational equations have a distinguished projectable extension α_ε that is *global* and *closed*. The dynamical distribution of α_ε is denoted by Δ_ε and called the *Euler-Lagrange distribution*: note, it is defined on $\mathbb{R} \times TM$. Thus, for variational equations we have three generally different *global* representations by exterior differential systems on the evolution space:

(1) the evolution distribution $\mathcal{D}_\varepsilon$, whose integral sections coincide with prolongations of extremals,

(2) the characteristic distribution $\chi_{\alpha_\varepsilon}$ of the closed 2-form α_ε, whose *holonomic* integral sections coincide with prolongations of extremals.

(3) the Euler-Lagrange distribution Δ_ε whose *holonomic* integral sections coincide with prolongations of extremals (and are the same as the holonomic integral sections of $\chi_{\alpha_\varepsilon}$).

Taking into account that for every Lagrangian λ, α_ε is on the domain of λ equal to the Cartan 2-form $d\theta_\lambda$, we can see that Euler-Lagrange equations are

– equations for integral sections of the evolution distribution $\mathcal{D}_\varepsilon$,

– equations for holonomic integral sections of the characteristic distribution of α_ε, thus, for any Lagrangian λ of ε taking the form

$$J^1\gamma^* i_X d\theta_\lambda = 0 \quad \forall \text{ vector field } X \text{ on } \mathbb{R} \times TM,$$

– equations for holonomic integral sections of the dynamical distribution Δ_ε, thus, for any Lagrangian λ of ε taking the form

$$J^1\gamma^* i_X d\theta_\lambda = 0 \quad \forall\, \pi_1\text{-vertical vector field } X \text{ on } \mathbb{R} \times TM.$$

Equations for integral sections of the Euler-Lagrange distribution Δ_ε, i.e. equations for sections δ of π_1 satisfying

$$\delta^* i_X \alpha_\varepsilon = 0 \quad \text{for every } \pi_1\text{-vertical vector field } X \text{ on } \mathbb{R} \times TM \tag{38}$$

are called *Hamilton equations* [37, 71]; solutions of Hamilton equations are then called *Hamilton extremals*. With this terminology, equations for *holonomic Hamilton extremals* are the Euler-Lagrange equations.

Apparently, every prolongation of an extremal is a Hamilton extremal. The set of extremals, however, need not be in bijective correspondence with the set of Hamilton extremals, meaning that Hamilton equations in general are *not equivalent* with the Euler-Lagrange equations. We can see, however, that the following holds:

Proposition 4.1 *If ε is regular then Hamilton equations and Euler-Lagrange equations are equivalent, meaning that every Hamilton extremal is a prolongation of an extremal of ε.*

Let us turn to *regular* locally variational forms in more detail. Combining properties of regular equations with variationality we get

Proposition 4.2 *(Krupková 1987, 1994) Let ε be a regular locally variational form on $\mathbb{R} \times T^2M$. Denote*

$$\varepsilon = E_a \omega^a \wedge dt, \quad E_a = B_{ab}\ddot{x}^b + A_a.$$

The following conditions are equivalent:

(1) *ε is regular.*

(2) *Every local projectable extension α of ε has maximal rank ($= 2m$).*

(3) *The Euler-Lagrange distribution Δ_ε is a semispray connection.*

(4) $\mathcal{D}_\varepsilon = \Delta_\varepsilon = \chi_{\alpha_\varepsilon}$.

(5) *Equation $i_\Gamma \alpha_\varepsilon = 0$ has a unique global solution satisfying the scaling condition $\Gamma(dt) = 1$; it reads*

$$\Gamma = \frac{\partial}{\partial t} + \dot{x}^a \frac{\partial}{\partial x^a} - B^{ab} A_b \frac{\partial}{\partial \dot{x}^a}. \tag{39}$$

The equivalence of (1) and (5) in proposition 4.2 was originally obtained by Goldschmidt and Sternberg [37, 136]. They called the vector field Γ (39) *Euler-Lagrange field*. With help of (19) its components can be easily expressed by means of a Lagrangian for ε.

Equipped with the geometric understanding of regularity of equations we can extend it to Lagrangians as follows: A Lagrangian λ is called *regular* if its Euler-Lagrange form is regular [71].

Now, given a Lagrangian $\lambda = L dt$ for ε we can rewrite the regularity condition in terms of the Lagrange function L. If λ is a *first order Lagrangian*, we get the well-known condition

$$\det \left(\frac{\partial^2 L}{\partial \dot{x}^a \partial \dot{x}^b} \right) \neq 0.$$

If, however λ is a *second order Lagrangian* for the (second order) locally variational form ε, the regularity condition takes the form

$$\det\left(\frac{\partial^2 L}{\partial \dot{x}^a \partial \dot{x}^b} - \frac{\partial^2 L}{\partial \ddot{x}^a \partial x^b} - \frac{\partial^2 L}{\partial x^a \partial \ddot{x}^b} - \frac{d}{dt}\frac{\partial^2 L}{\partial \dot{x}^a \partial \ddot{x}^b}\right) \neq 0.$$

For regular Lagrangians we have the following *theorem on a local canonical form of the closed* 2*-form* α_ε [71]:

Theorem 4.3 *Let ε be a regular locally variational form on $\mathbb{R} \times T^2M$. Then in a neighbourhood of every point in the evolution space $\mathbb{R} \times TM$ there is a chart with coordinates (t, x^a, p_a), and a function H, such that*

$$\alpha_\varepsilon = -dH \wedge dt + dp_a \wedge dx^a.$$

The functions H and p_a defined above are called a *Hamiltonian* and *momenta* of the locally variational form ε.

The Hamiltonian and momenta of ε are non-unique: obviously for any local function $\varphi(t, x^i)$, the family

$$H' = H + \frac{\partial \varphi}{\partial t}, \quad p'_a = p_a - \frac{\partial \varphi}{\partial x^a}$$

is another set of Hamiltonian and momenta for ε. Hamiltonians and momenta of a locally variational form are in one-to-one correspondence with the *first order* (generally, if ε is of higher order, with the *minimal order*) Lagrangians.

The coordinates (t, x^a, p_a) are called *Legendre coordinates*. Expressing the coordinate-free Hamilton equations (38) in Legendre coordinates we get the familiar Hamilton equations of classical mechanics,

$$\frac{dx^a}{dt} = \frac{\partial H}{\partial p_a}, \qquad \frac{dp_a}{dt} = -\frac{\partial H}{\partial x^a},$$

that are equivalent with the Euler-Lagrange equations.

Going back to proposition 2.8, we realised, that to each pertinent *regular* dynamical form ε on $\mathbb{R} \times T^2M$ there exists a *unique* semispray connection w having the same paths as ε. Conversely, however, given a semispray connection w, there is a *family* of pertinent regular dynamical forms, having the same paths as w: they are characterised by the condition $w^*\varepsilon = 0$, i.e. their components are given by

$$E_a = g_{ab}(\ddot{x}^b - f^b),$$

where $g = (g_{ab})$ is an arbitrary *regular* matrix. Consequently, to a given semispray

$$\Gamma = \frac{\partial}{\partial t} + \dot{x}^a \frac{\partial}{\partial x^a} + f^a \frac{\partial}{\partial \dot{x}^a}$$

there corresponds a family of local projectable extensions of the ε's, i.e. a family of 2-forms of rank $2m$ such that Γ spans the characteristic distribution of α; they read

$$\alpha = g_{ab}\omega^a \wedge (d\dot{x}^b - f^b dt) + F,$$

where g is any regular matrix, and F is an arbitrary 2-contact 2-form.

A regular matrix g above is called a *multiplier* for Γ. It has an intrinsic meaning as a *regular fibred morphism* between the fibred manifolds $R \times TM \to M$ and $T^0_2 M \to M$ over the identity of M, where the latter is the bundle of all tensors of type $(0,2)$ over M (g is a "time, position and velocity dependent metric" on M, considered in Sec. 3.4).

Now, taking into account the relation between semisprays and dynamical forms we can extend to semisprays the inverse variational problem:

A semispray Γ is called *variational* [72] if the related family of dynamical forms contains a locally variational form, i.e. if in a neighbourhood of every point in the evolution space there exists a regular matrix g such that $\varepsilon = g_{ab}(\ddot{x}^b - f^b)\omega^a \wedge dt$ is variational. For a variational semispray Γ we thus have a *regular* first order Lagrangian λ such that $\mathcal{D}_{\varepsilon_\lambda} = \text{span}\{\Gamma\}$.

Any *regular* matrix g relating a semispray Γ with a locally variational form is called a *variational multiplier* for Γ. Since the related dynamical form satisfies the Helmholtz conditions, we can see that every variational multiplier is *symmetric* and is a solution of Helmholtz conditions that become *partial differential equations for g*. Existence and multiplicity of variational multipliers for a given semispray (SODEs in normal form) is a crucial problem; we call it the *contravariant inverse problem of the calculus of variations*. Although this problem for a *general* system of m SODEs in normal form is yet unsolved, many interesting particular achievements have been reached. We devote Sections 5 and 6 to the inverse problem for semisprays and to various geometric structures that are helpful in dealing with this question.

Due to the global existence of the Euler-Lagrange distribution Δ_ε for variational equations, that, as we have seen above describes *Hamilton extremals*, we have along with regular and weakly regular dynamical forms (cf. Sec. 2.4) another distinguished class:

A locally variational form ε is called *semiregular* if its Euler-Lagrange distribution is weakly horizontal and of a constant rank.

Since in this case Δ_ε *coincides with the characteristic distribution of the* 2*-form* α_ε [74] (see remark 2.13 in Sec. 2.4), and since α_ε has constant rank and is closed, we get that *the Euler-Lagrange distribution of a semiregular locally variational form is completely integrable*.

Given a *semiregular* locally variational form ε on $\mathbb{R} \times T^2 M$, a $(2m+1-p)$-dimensional immersed submanifold $\mathcal{Q}$ of the evolution space $\mathbb{R} \times TM$ is called *Lagrangian submanifold* of ε if $\mathcal{Q}$ is foliated by the leaves of Δ_ε.

Integration methods for the characteristic distribution of a closed 2-form, and in particular, for a semiregular locally variational form will be subject of the rest of this section.

4.2 Symmetries of Lagrangian structures

Given a locally variational form (variational SODEs) one can study invariance transformations and symmetries of the corresponding *differential forms* and *distributions*. We shall be interested in relations between various kinds of symmetries, and in relations of symmetries and conservation laws. Applications to exact integration of the Euler-Lagrange equations will then be subject of the next section.

Geometric foundations of the *theory of invariant variational problems* were laid by Trautman [142, 143] and Krupka [60]. Here we focus on ordinary differential equations (see also the book [82] by Krupková). For a general setting the reader can consult the paper by Krupka in this book [66].

First, recall the concept of invariance transformation and of symmetry of a differential form and of a distribution on a manifold $\mathcal{N}$ (see e.g. [90, 92]). The reader should be aware of the varied terminology appearing in the literature: invariance transformations are often called *finite symmetries* and symmetries are then called *infinitesimal symmetries*.

A local diffeomorphism $\phi : \mathcal{N} \to \mathcal{N}$ is called *invariance transformation* of a q-form η if

$$\phi^* \eta = \eta \,.$$

Let X be a vector field on $\mathcal{N}$, and $\{\phi_u\}$ its local one-parameter group of transformations. X is called a *symmetry* of η, if for every u, ϕ_u is an invariance transformation of η. Equivalently, the symmetry condition takes the form

$$\mathcal{L}_X \eta = 0.$$

Given a distribution $\mathcal{D}$ of a constant rank on $\mathcal{N}$, ϕ is called *invariance transformation* of $\mathcal{D}$ if for all $x \in \mathcal{N}$,

$$T\phi(\mathcal{D}_x) \subset \mathcal{D}_{\phi(x)}.$$

The vector field X is called a *symmetry* of $\mathcal{D}$ if for all u, ϕ_u is an invariance transformation of $\mathcal{D}$. The set of all symmetries of a distribution is a Lie algebra with respect to the Lie bracket of vector fields, characterised as follows [92]:

Theorem 4.4 *Let $\mathcal{D}$ be a distribution of a constant rank. The following three conditions are equivalent:*

(1) *X is a symmetry of $\mathcal{D}$,*

(2) *for every vector field Z belonging to $\mathcal{D}$, the Lie bracket $[X, Z]$ belongs to $\mathcal{D}$,*

(3) *for every one-form η belonging to the annihilator $\mathcal{D}^0$ of $\mathcal{D}$, the Lie derivative $\mathcal{L}_X \eta$ belongs to $\mathcal{D}^0$.*

We can see that if X (with the local one-parameter group $\{\phi_u\}$) is a symmetry of $\mathcal{D}$, and $\mathcal{Q}$ is an integral manifold of $\mathcal{D}$ then $\phi_u(\mathcal{Q})$ is also an integral manifold of $\mathcal{D}$. In other words, the local flow of a symmetry transfers integral mappings into integral mappings, and, consequently integral manifolds into integral manifolds.

Let $\mathcal{D}$ be a distribution on $\mathcal{N}$. A function f, defined on an open subset U of $\mathcal{N}$ is called a *first integral* of $\mathcal{D}$ if $df \in \mathcal{D}^0$, i.e., if $i_X df = 0$ for every vector field $X \in \mathcal{D}$. We can see that if f is a first integral of $\mathcal{D}$ on U and $\iota : \mathcal{Q} \to U \subset \mathcal{N}$ is an integral manifold of $\mathcal{D}$ then $d(f \circ \iota) = 0$, i.e., the function f is a constant along $\mathcal{Q}$.

First integrals $f_1,\ f_2$ of $\mathcal{D}$ defined on U are called *independent* if the forms df_1, df_2 are linearly independent at each point of U. Let $f_1, \dots, f_p$ be independent first integrals of $\mathcal{D}$, defined on U. The set $\{f_1, \dots, f_p\}$ is called *a complete set of independent first integrals of $\mathcal{D}$ at a point $x \in U$* (resp. *on U*) if

$$\mathcal{D}^0(x) = \mathrm{span}\{df_1(x), \dots, df_p(x)\}$$

(resp. if the above condition holds at each point of U).

If $\{f_1, \ldots, f_p\}$ is a complete set of independent first integrals of $\mathcal{D}$ then the submanifolds of U characterised by the equations $f_1 = c_1, \ldots, f_p = c_p$, where $c_1, \ldots, c_p$ are constants, are integral manifolds of $\mathcal{D}$ of maximal dimension, and one can find local coordinates (y^i) on $\mathcal{N}$ such that $y^1 = f_1, \ldots, y^p = f_p$.

If the distribution $\mathcal{D}$ has constant rank k and is completely integrable then in a neighbourhood of every point in $\mathcal{N}$ the existence of a complete set of independent first integrals is guaranteed by Frobenius Theorem. Indeed, around every point in $\mathcal{N}$ there is an adapted chart (U, φ), $\varphi = (y^i)$, to the foliation defined by $\mathcal{D}$, such that

$$\mathcal{D} = \text{span}\Big\{\frac{\partial}{\partial y^1}, \ldots, \frac{\partial}{\partial y^k}\Big\} = \text{annih}\{dy^{k+1}, \ldots, dy^n\}$$

on U. Apparently, $\{y^{k+1}, \ldots, y^n\}$ is a complete set of independent first integrals of $\mathcal{D}$ on U. The reader is referred to the paper [131] which describes an integration strategy for systems of ordinary differential equations in the spirit of the foregoing discussion.

Let us return to a fibred manifold $\pi_0 : \mathbb{R} \times M \to \mathbb{R}$ and its jet prolongations. Let λ be a Lagrangian on $\mathbb{R} \times TM$, ε_λ its Euler-Lagrange form. An isomorphism ϕ of π_0 is called *invariance transformation of* λ if

$$J^1\phi^*\lambda = 0,$$

and a *generalised invariance transformation of* λ if

$$J^1\phi^*\varepsilon_\lambda = 0.$$

The condition

$$\mathcal{L}_{J^1X}\lambda = 0$$

for a π_0-projectable vector field X on $\mathbb{R} \times M$ be a symmetry of the Lagrangian λ is called *Noether equation*. In coordinates it takes the form

$$J^1X(L) + L\frac{dX^0}{dt} = 0. \tag{40}$$

Similarly, the condition

$$\mathcal{L}_{J^2X}\varepsilon_\lambda = 0$$

for a π_0-projectable vector field X on $\mathbb{R} \times M$ be a symmetry of the Euler-Lagrange form of λ is called *Noether-Bessel-Hagen equation*. In coordinates,

$$J^2X(E_a) + E_b\frac{\partial X^b}{\partial x^a} + E_a\frac{dX^0}{dt} = 0, \tag{41}$$

where for the higher components of J^2X formula (1) has to be used.

The Noether equation (respectively, Noether-Bessel-Hagen equation) can be used to find *all symmetries* of a given Lagrangian (respectively, of a given dynamical form), or, if a group of transformations of $\mathbb{R} \times TM$ is given, to find all invariant Lagrangians (respectively, dynamical forms).

To find all Euler-Lagrange expressions (locally variational forms) possessing prescribed symmetries one has to combine the Noether-Bessel-Hagen equation with Helmholtz conditions (cf. [134]).

For the sake of brevity, symmetries on the evolution space that are of the form J^1X, where X is a vector field on the extended configuration space $\mathbb{R} \times M$, are called *point symmetries*.

As shown in [60], joining the symmetry requirements with the first variation formula (20), we immediately obtain the famous Noether theorem and its generalisation as follows:

Theorem 4.5 ***Noether Theorem.*** *[105, 60] Let λ be Lagrangian on $\mathbb{R} \times TM$, θ_λ the Cartan form of λ. Assume that for a π_0-projectable vector field X on $\mathbb{R} \times M$, J^1X is a symmetry of λ. Then for every extremal γ of λ,*

$$i_{J^1X}\theta_\lambda \circ J^1\gamma = const. \tag{42}$$

Equation (42) saying that the function $i_{J^1X}\theta_\lambda$ is constant along extremals, is called a *conservation law*.

Theorem 4.6 ***Generalised Noether Theorem.*** *[60] Let ε be a locally variational form on $\mathbb{R} \times T^2M$, let a π_0-projectable vector field X on $\mathbb{R} \times M$ be a point symmetry of ε. If λ is a (local) Lagrangian for ε on an open set $W \subset \mathbb{R} \times TM$, and ρ is the unique closed 1-form such that $\mathcal{L}_{J^1X}\lambda = h\rho$, and if γ is an extremal of ε defined on $\pi_1(W) \subset \mathbb{R}$, then*

$$J^1\gamma^*(di_{J^1X}\theta_\lambda - \rho) = 0.$$

This means that a point symmetry of a locally variational form ε gives rise to *conservation laws*

$$(i_{J^1X}\theta_\lambda - g) \circ J^1\gamma = const,$$

where g is a function (defined on an appropriate open set) such that $dg = \rho$. Thus, knowing a point symmetry of ε, the Generalised Noether Theorem provides us, for *every* Lagrangian, with a *function constant along extremals*.

Given a locally variational form ε on $\mathbb{R} \times T^2M$, we can consider symmetries of the dynamical form itself, but also symmetries of its Lagrangians, Cartan forms and Cartan 2-forms (that unify, as we know, into a global closed 2-form – the Lepage equivalent of ε). We now clarify relations between point symmetries of these differential forms.

Proposition 4.7 *Let λ be a Lagrangian on $\mathbb{R} \times TM$, ε_λ its Euler-Lagrange form. Let X be a π_0-projectable vector field on $\mathbb{R} \times M$. Then*

$$\mathcal{L}_{J^2X}\varepsilon_\lambda = \varepsilon_{\mathcal{L}_{J^1X}\lambda}.$$

Corollary 4.8 (1) *Given a point symmetry X of ε_λ there exists a unique closed one-form ρ such that*

$$\mathcal{L}_{J^1X}\lambda = h\rho.$$

(2) *Every point symmetry of a Lagrangian is a point symmetry of its Euler-Lagrange form.*

Moreover, we have [80]:

Theorem 4.9 *Let X be a π_0-projectable vector field on $\mathbb{R} \times M$.*

(1) *X is a point symmetry of λ if and only if it is a point symmetry of the Cartan form θ_λ.*

(2) *X is a point symmetry of the Euler-Lagrange form ε_λ if and only if it is a point symmetry of the Lepage equivalent $\alpha_{\varepsilon_\lambda} = d\theta_\lambda$.*

(3) *Every point symmetry X of the Cartan form θ_λ is a point symmetry of the Cartan 2-form $d\theta_\lambda = \alpha_{\varepsilon_\lambda}$.*

(4) *If rank of $\alpha_\varepsilon = d\theta_\lambda$ is constant then every point symmetry X of α_ε is a point symmetry of the characteristic distribution of α_ε. If X is vertical then it is also a point symmetry of the Euler-Lagrange distribution Δ_ε.*

(5) *If ε_λ is semiregular then every point symmetry of ε_λ (hence of $\alpha_{\varepsilon_\lambda}$) is a point symmetry of the Euler-Lagrange distribution Δ_ε.*

Summarizing briefly, the set of point symmetries of a locally variational form ε is the same as the set of point symmetries of the closed 2-form α_ε. It contains the set of all point symmetries of any Lagrangian λ of ε (that coincides with the set of point symmetries of the Cartan form θ_λ), and for semiregular ε is contained in the set of point symmetries of the Euler-Lagrange distribution Δ_ε.

The Noether Theorems provide us with "constants of the motion" related with symmetries, i.e. with functions that remain constant along prolonged extremals, each corresponding to a symmetry of a Lagrangian. For *regular* Lagrangian systems, when there is no difference between the evolution distribution $\mathcal{D}_\varepsilon$, Euler-Lagrange distribution Δ_ε, and the characteristic distribution of the Lepage 2-form α_ε, hence prolongations of extremals coincide with integral sections (maximal integral manifolds) of (any of) the distributions, there is no difference between constants of the motion and first integrals of $\mathcal{D}_\varepsilon$. For non-regular Lagrangian systems, however, one may find functions constant along (prolonged) extremals that are not first integrals of the characteristic distribution of α_ε [82]. So there arises a question whether also in these more complicated cases the Noether Theorems can be used for an explicit integration of the characteristic and/or the Euler-Lagrange distribution. It was proved in [81] that the answer is affirmative, and *Noether Theorems hold also in the following stronger formulation*:

Theorem 4.10 (1) *Let λ be Lagrangian on $\mathbb{R} \times TM$, θ_λ the Cartan form of λ. Assume that a π_0-projectable vector field X on $\mathbb{R} \times M$ is a point symmetry of λ. Then the function $i_{J^1X}\theta_\lambda$ is a first integral of the characteristic distribution $\chi_{\alpha_\varepsilon}$ of the 2-form α_ε.*

(2) *Let ε be a locally variational form on $\mathbb{R} \times T^2M$, let a π_0-projectable vector field X on $\mathbb{R} \times M$ be a point symmetry of ε. If λ is a (local) Lagrangian for ε on an open set $W \subset \mathbb{R} \times TM$, and $\rho = dg$ is the unique closed 1-form such that $\mathcal{L}_{J^1X}\lambda = h\rho$, then $i_{J^1X}\theta_\lambda - g$ is a first integral of the characteristic distribution $\chi_{\alpha_\varepsilon}$ of the 2-form α_ε.*

So far, we have studied symmetries determined by vector fields *on the extended configuration space* $\mathbb{R} \times M$. Now we shall consider *general vector fields on the evolution* *space* $\mathbb{R} \times TM$. Given a Lagrangian system a significant role is played by *symmetries of the corresponding Lepage forms*, i.e. Cartan 1-forms θ_λ of individual Lagrangians and the closed 2-form α_ε (as we have seen, for every Lagrangian λ for ε, on the domain of λ, $\alpha_{\varepsilon_\lambda}$ identifies with the Cartan 2-form $d\theta_\lambda$).

We have the following relations between different kinds of symmetries associated with a Lagrangian system [80]:

Proposition 4.11 *Let X be a vector field on the evolution space $\mathbb{R} \times TM$.*

(1) *Every symmetry of the Cartan form θ_λ is a symmetry of the Cartan 2-form $d\theta_\lambda = \alpha_{\varepsilon_\lambda}$.*

(2) *Every symmetry of $d\theta_\lambda = \alpha_{\varepsilon_\lambda}$ is a symmetry of the characteristic distribution $\chi_{\alpha_{\varepsilon_\lambda}}$ of the closed 2-form $\alpha_{\varepsilon_\lambda}$.*

Symmetries of Lepage forms provide us with *first integrals of the characteristic distribution*:

If $\mathcal{L}_X \alpha_{\varepsilon_\lambda} = 0$ then $d i_X \alpha_{\varepsilon_\lambda} = 0$ so that locally

$$i_X \alpha_{\varepsilon_\lambda} = df,$$

i.e. by definition of the characteristic distribution, $i_X \alpha_{\varepsilon_\lambda}$ belongs to the annihilator of $\chi_{\alpha_{\varepsilon_\lambda}}$. This means that f is a first integral of the distribution $\chi_{\alpha_{\varepsilon_\lambda}}$; it can be obtained by integration using the Poincaré Lemma.

Similarly, if $\mathcal{L}_X \theta_\lambda = 0$ then $i_X d\theta_\lambda = i_X \alpha_{\varepsilon_\lambda} = -d i_X \theta_\lambda$, so that the function

$$f = i_X \theta_\lambda$$

is a first integral of the characteristic distribution $\chi_{\alpha_{\varepsilon_\lambda}}$.

Recall that if the locally variational form ε_λ is *semiregular*, or even *regular* then the characteristic distribution $\chi_{\alpha_{\varepsilon_\lambda}}$ coincides with the Euler-Lagrange distribution $\Delta_{\varepsilon_\lambda}$, i.e. symmetries of Lepage one and two-forms provide us with first integrals of the Euler-Lagrange distribution $\Delta_{\varepsilon_\lambda}$. Moreover, since in this case $\Delta_{\varepsilon_\lambda}$ is completely integrable, they can be used to find extremals.

A method of integration of the Euler-Lagrange distribution based on symmetries of Lepage forms and related first integrals is subject of the next section.

It should be pointed out that there is a rich bibliography dealing with various questions concerning symmetries and first integrals in Lagrangian mechanics, some of the references are listed in the Bibliography. The review paper by Sarlet and Cantrijn [122] has been heavily cited in the literature as has the classification of symmetries in Lagrangian dynamics due to Prince [108, 109]. One can find also other interesting results, for example, relating the theory of symmetries with the inverse problem of the calculus of variations [7, 117, 119, 120, 126], or clarifying the place of symmetries of the Helmholtz form in the theory of SODEs [67, 68]. There is a comprehensive application of the generalised Noether theorem to the geodesic equations of a (pseudo) Riemannian metric to be found in [114, 115].

4.3 The Liouville Theorem for closed 2-forms and integration methods for semiregular variational equations

In this section we present a generalisation of the Liouville theorem of the classical calculus of variations [91] to *characteristic distributions of (arbitrary) closed two-forms of a constant rank* [74, 75, 113], and in particular, to *semiregular variational equations* (see also [82] for exposition and related integration methods for higher order SODEs, and [54] for variational equations); the generalisation is based on the Darboux Theorem for closed 2-forms and on the understanding of complete integrals as distributions [8].

Consider a manifold $\mathcal{N}$ of dimension n. Let α be a closed two-form of a constant rank on $\mathcal{N}$. Denote by χ_α the characteristic distribution of α. Thus, the rank of α is an even number, and $\operatorname{rank}\alpha = \operatorname{corank}\chi_\alpha$. We set

$$\operatorname{rank}\alpha = 2p, \quad \operatorname{rank}\chi_\alpha = r,$$

so that $2p + r = n$.

Theorem 4.12 ***Darboux Theorem***. *Given a closed two-form α as above, then at each point $x \in \mathcal{N}$ there is a local chart with coordinates*

$$(a^1, \dots, a^p, b_1, \dots, b_p, y^{2p+1}, \dots, y^n)$$

such that

$$\alpha = \sum_{K=1}^{p} da^K \wedge db_K.$$

Charts characterised by Darboux Theorem are called *Darboux charts* of α. In a Darboux chart one has for the characteristic distribution

$$\chi_\alpha = \operatorname{span}\{\partial/\partial y^\sigma,\ 2p+1 \le \sigma \le n\} = \operatorname{annih}\{da^K,\ db_K,\ 1 \le K \le p\},$$

so that the Darboux functions $a^1, \dots, a^p, b_1, \dots, b_p$ form *a complete set of independent first integrals of the characteristic distribution* χ_α. This, however, means that *the characteristic distribution χ_α of α is completely integrable*.

Let us turn to symmetries. If X is a vector field on $\mathcal{N}$, and α is closed, then $\mathcal{L}_X\alpha = di_X\alpha$. This means that if X is a symmetry of α then (locally)

$$i_X\alpha = df,$$

where *f is a first integral of the characteristic distribution χ_α*. In particular, every vector field belonging to χ_α is a symmetry of α; the corresponding first integrals are trivial (constant functions).

If ρ is a one-form (possibly defined on an open subset $U \subset \mathcal{N}$) such that $\alpha = d\rho$ on U, we can see that $\mathcal{L}_X\rho = 0$ means that $i_X\alpha = df$, where $f = -i_X\rho$. Thus, *every symmetry X of ρ is a symmetry of α, and $i_X\rho$ is a first integral of the characteristic distribution χ_α*.

Note that a first integral corresponding to a symmetry of α is unique up to a constant function. Conversely, to a given first integral the corresponding symmetry is not unique (all the symmetries form a class modulo the characteristic distribution).

Let $X_1,\ X_2$ be two symmetries of α. Then

$$i_{[X_1,X_2]}\alpha = \mathcal{L}_{X_1} i_{X_2}\alpha - i_{X_2}\mathcal{L}_{X_1}\alpha = d i_{X_1} i_{X_2}\alpha,$$

and we have the following assertion which can be viewed as a generalisation of the classical *Poisson Theorem*.

Proposition 4.13 *The set of all symmetries of a closed two-form α is a Lie algebra.*

If $f_1,\ f_2$ are first integrals of the characteristic distribution χ_α, and $X_1,\ X_2$ are symmetries of α corresponding to f_1 and f_2, respectively, then

$$\{f_1, f_2\} := i_{X_1} i_{X_2}\alpha = i_{X_1} df_2 = -i_{X_2} df_1$$

is a first integral of χ_α, corresponding to the symmetry $[X_1, X_2]$ of α.

The first integral $\{f_1, f_2\}$ is called the *Poisson bracket* of the first integrals $f_1,\ f_2$.

In particular, for every symmetry X of α and every Z belonging to the characteristic distribution we have $i_{[X,Z]}\alpha = d i_X i_Z \alpha = 0$, i.e., the symmetry $[X, Z]$ of α belongs to the characteristic distribution, giving rise to trivial first integrals. Consequently, using the second condition of theorem 4.4, we get the relation between symmetries of a closed two-form of a constant rank and of its characteristic distribution as follows: *Every symmetry of α is a symmetry of the characteristic distribution χ_α.*

The problem of finding solutions of the characteristic distribution χ_α can be considered as a problem of finding a covering of the manifold $\mathcal{N}$ by Darboux charts. The classical idea of how to proceed is the following: use p symmetries of α to find p independent first integrals, the remaining p first integrals then can be computed by means of differentiation and integration from α ("by quadratures" in the classical terminology). However, not any family of p independent first integrals of α is appropriate for using the Liouville integration formulas: one has to start with the so-called "first integrals in involution". Therefore to find a proper generalisation of the theorem, we will be interested not only in the procedure alone, but first of all in a geometric characterisation of admissible families of first integrals. Consider a closed two-form α of a constant rank $= 2p$ on $\mathcal{N}$. Let $\mathcal{I}$ be a distribution defined on an open subset U of $\mathcal{N}$. $\mathcal{I}$ is called a *complete integral* of α on U [74] if

(1) $\mathcal{I}$ is completely integrable and corank $\mathcal{I} = p$ on U, and ,

(2) α belongs to the differential ideal generated by $\mathcal{I}$.

Condition (2) in the above definition means that $\alpha = 0$ on the maximal integral manifolds of $\mathcal{I}$.

The *existence* of local complete integrals is guaranteed by Darboux theorem: if $(U, \varphi), \varphi = (a^K, b_K, y^\sigma), 1 \le K \le p,\ 1 \le \sigma \le r$, is a Darboux chart of α, then putting

$$\mathcal{I} = \text{annih}\{da^K,\ 1 \le K \le p\}$$

we get a complete integral of α defined on U.

The geometric meaning of complete integrals is as follows:

Proposition 4.14 *If $\mathcal{I}$ be a complete integral of α, defined on an open set U, then on U, the characteristic distribution χ_α is a subdistribution of $\mathcal{I}$. Consequently, every leaf of the foliation of the distribution $\mathcal{I}$ is foliated by the leaves of the characteristic distribution χ_α.*

In order to distinguish systems of first integrals of the characteristic distribution that define complete integrals of α, we have the following definition: Let a^K, $1 \le K \le p$, be independent first integrals of the characteristic distribution χ_α. We say that the integrals a^K are *in involution* if the distribution $\mathcal{I} = \text{annih}\{da^K,\, 1 \le K \le p\}$ is a complete integral of α.

The following theorem answers the question of under which conditions a family of first integrals of χ_α (respectively, a family of symmetries of α) defines a complete integral [75].

Theorem 4.15 *Let a^K, $1 \le K \le p$, be independent first integrals of the characteristic distribution of α, and let X_K, $1 \le K \le p$, be some corresponding symmetries of α, i.e., such that $i_{X_K}\alpha = da^K$, for all K.*

Alternatively, let X_K, $1 \le K \le p$, be nontrivial symmetries of α (i.e., such that $i_{X_K}\alpha \ne 0$ for all K, that is, $X \notin \chi_\alpha$), linearly independent at each point of their domain of definition, and let a^K, $1 \le K \le p$, be some corresponding first integrals of χ_α.

Then the following conditions are equivalent:

(1) *The distribution*

$$\mathcal{I} = \textit{annih}\{i_{X_K}\alpha,\, 1 \le K \le p\} = \textit{annih}\{da^K,\, 1 \le K \le p\}$$

is a complete integral of α.

(2)

$$\{a^K, a^L\} = i_{X_K} i_{X_L}\alpha = i_{X_K}\, da^L = 0, \quad \forall\, 1 \le K,\, L \le p.$$

(3)

$$\textit{annih}\{da^K,\, 1 \le K \le p\} = \textit{span}\{X_K,\, Z_\sigma,\, 1 \le K \le p,\, 1 \le \sigma \le r\},$$

where the vector fields $Z_1, \dots, Z_r$ span χ_α.

Both the definition of a complete integral and condition (3) above express the geometric content of the classical concept of "a system of first integrals in involution" that is defined by vanishing of all the Poisson brackets of the first integrals, i.e. by condition (2) of the theorem.

The following theorem is a generalisation of the *Liouville theorem* of classical mechanics to closed two-forms of a constant rank, and it is of fundamental importance for integration of distributions [75]:

Theorem 4.16 *Let α be a closed two-form of constant rank $2p$ on $\mathcal{N}$, let $\mathcal{I}$ be a complete integral of α. Then at each point of the domain of $\mathcal{I}$ there exists a chart (U, φ), $\varphi = (a^K, b_K, y^\sigma)$, $1 \le K \le p$, $1 \le \sigma \le \dim \mathcal{N} - 2p$, such that*

(1) $\mathcal{I} = \textit{annih}\{da^K,\, 1 \le K \le p\}$,

(2) $\alpha = da^K \wedge db_K$,

(3) *the set of functions $\{a^K, b_K,\, 1 \le K \le p\}$ is a complete set of independent first integrals of the characteristic distribution χ_α of α.*

Note that condition (1) means that $(U,\,\varphi)$ is an adapted chart to the distribution $\mathcal{I}$, and (2) means that it is a Darboux chart of α.

One can easily get *explicit formulas for the first integrals* b_K, $1 \leq K \leq p$, as follows: Given a complete integral of χ_α (= a set of independent first integrals of χ_α satisfying $\{a^K, a^L\} = 0$), the functions a^K can be completed to local coordinates, (y^J, a^K), on $\mathcal{N}$. Now, since locally $\alpha = d\rho$, we have in the adapted coordinates

$$\rho = \rho_J\, dy^J + \hat{\rho}_K\, da^K,$$

where

$$\frac{\partial \rho_J}{\partial y^I} - \frac{\partial \rho_I}{\partial y^J} = 0\,,$$

since by assumption $d\rho = 0$ on the leaves of $\mathcal{I}$. The latter are the integrability conditions for the existence of a local function $S(y^J, a^K)$ such that

$$\rho_J = \frac{\partial S}{\partial y^J}. \tag{43}$$

Now, we have

$$\rho = \frac{\partial S}{\partial y^J}\, dy^J + \hat{\rho}_K\, da^K = dS + \Big(\hat{\rho}_K - \frac{\partial S}{\partial a^K}\Big) da^K. \tag{44}$$

Putting

$$b_K = \hat{\rho}_K - \frac{\partial S}{\partial a^K} \tag{45}$$

gives us the desired set of p independent first integrals of the distribution χ_α. The functions $\hat{\rho}_K$ and S can be easily found using Poincaré Lemma. Indeed, since α belongs to the ideal generated by $\mathcal{I}^0 = \text{span}\{da^1, \ldots, da^p\}$, we have

$$\alpha = g_{JK}\, dy^J \wedge da^K + h_{LK}\, da^L \wedge da^K,$$

and $\rho = B\alpha$, where B is the standard homotopy operator. This, however, means that

$$\rho_J = -a^L \int_0^1 (g_{JL} \circ \psi)\, u\, du$$
$$\hat{\rho}_K = y^J \int_0^1 (g_{JK} \circ \psi)\, u\, du + 2a^L \int_0^1 (h_{LK} \circ \psi)\, u\, du,$$

where the mapping ψ is defined by $\psi(u, y^J, a^K) = (uy^J, ua^K)$. Knowing ρ, we obtain a suitable S from (43) easily again by application of the Poincaré Lemma, this time, however, with use of a homotopy operator $A^{\mathcal{I}}$ adapted to the foliation defined by $\mathcal{I}$:

$$S = A^{\mathcal{I}}\rho = y^J \int_0^1 (\rho_J \circ \mu)\, dv,$$

where

$$\mu(v, y^J, a^K) = (vy^J, a^K)\,.$$

Note that S is unique up to a function $h(a^K)$ (i.e. up to first integrals of the distribution $\mathcal{I}$). Substituting all this into (45) an explicit formula for required p first integrals b_K appears, where the unknown first integrals are obtained by means of integration and differentiation ("by quadratures").

The above geometric version of the Liouville Theorem has the following reformulation that is close to the classical one:

Theorem 4.17 ***Generalised Liouville Theorem.*** *Let α be a closed two-form of constant rank $2p$ on $\mathcal{N}$. Let a^K, $1 \leq K \leq p$, be independent first integrals in involution of the characteristic distribution χ_α of α, defined on an open subset U of $\mathcal{N}$. Then the system of functions $\{a^K, b_K,\, 1 \leq K \leq p\}$, where b_K are defined by (45), is a complete set of independent first integrals of χ_α.*

Another reformulation of the same result could be called *Coordinate-free Liouville Theorem* [74]:

Corollary 4.18 *Let $\mathcal{I}_1$ be a complete integral of a closed two-form α of a constant rank on $\mathcal{N}$. Then to every point x of the domain of $\mathcal{I}_1$ there exists a neighborhood U and a complete integral $\mathcal{I}_2$ of α on U such that $\mathcal{I}_1 \cap \mathcal{I}_2$ is the characteristic distribution χ_α of α on U.*

Finally, the next reformulation of the Liouville Theorem is apparent from (44) and generalises the classical *Jacobi Theorem* [75].

Corollary 4.19 *Let $\mathcal{I}$ be a complete integral of a closed two-form α of a constant rank on $\mathcal{N}$. Then to every point x of the domain of $\mathcal{I}$ there exists a neighborhood U and a one-form $\bar{\rho}$ on U such that $\alpha = d\bar{\rho}$ and $\bar{\rho}$ belongs to $\mathcal{I}^0$ (i.e., $\bar{\rho}$ vanishes on the maximal integral manifolds of $\mathcal{I}$). Otherwise speaking, given a (local) one-form ρ such that $\alpha = d\rho$, there exists a function S such that*

$$\rho - dS \in \mathcal{I}^0. \tag{46}$$

In a chart adapted to the complete integral $\mathcal{I}$ (i.e. such that $\mathcal{I}^0 = \mathrm{span}\{da^K,\, 1 \leq K \leq p\}$), condition (46) reads

$$\bar{\rho} = \rho - dS = -b_K\, da^K,$$

hence

$$b_K = -i_{\partial/\partial a^K}\, \bar{\rho},$$

that obviously are independent first integrals of χ_α.

We have two important *applications of the geometric Liouville theory*:

1. Lagrangian systems. [74] This is the case when the manifold $\mathcal{N}$ is the evolution space $\mathbb{R}\times TM$ of a fibred manifold $\pi_0 : \mathbb{R}\times M \to \mathbb{R}$, and $\alpha = \alpha_\varepsilon$ is the Lepage equivalent of a *semiregular* locally variational form ε on $\mathbb{R}\times T^2M$ (recall that $d\alpha_\varepsilon = 0$, and $\alpha_\varepsilon = d\theta_\lambda$ for every, possibly local and higher-order, Lagrangian λ of ε). Then, as we know, rank α_ε is constant $= 2p$ and the characteristic distribution $\chi_{\alpha_\varepsilon}$ coincides with the Euler-Lagrange distribution Δ_ε. The Liouville theorem then provides an integration method for obtaining the foliation of the evolution space determined by the Euler- Lagrange distribution Δ_ε. If rank α_ε is maximal $(= 2m)$, i.e. if ε is regular, the foliation exactly corresponds to prolonged extremals of ε. If rank $\alpha_\varepsilon = 2p < 2m$, prolongations of extremals coincide with holonomic sections of the leaves of the foliation, while (general) sections of the leaves are *Hamilton extremals* of ε (solutions of Hamilton equations).

Note that by proposition 4.14, the leaves of the foliation defined by a complete integral $\mathcal{I}$ of α_ε are *Lagrangian submanifolds* (we also speak about *Lagrangian foliation* in this context).

If ε is *regular* and $\mathcal{I} = \text{annih}\{da^i, 1 \leq i \leq m\}$ is a complete integral of α_ε such that

$$\det\Big(\frac{\partial a^i}{\partial \dot{x}^j}\Big) \neq 0, \tag{47}$$

we can consider in the Liouville theorem adapted charts (t, x^i, a^i). Since ρ is the Cartan form θ_λ of a Lagrangian λ of ε, generalised Jacobi theorem (Corollary 4.19) then gives $\theta_\lambda - dS = -Hdt + p_i dx^i - dS = -b_i da^i$, so that

$$\frac{\partial S}{\partial t} = H, \quad \frac{\partial S}{\partial x^i} = p_i, \quad \frac{\partial S}{\partial a^i} = b_i.$$

Summarising, we get the famous *classical Jacobi Theorem* that *every solution* $S(t, x^i, a^i)$ *of the partial differential equation*

$$\frac{\partial S}{\partial t} = H\Big(t, x^i, \frac{\partial S}{\partial x^i}\Big) \tag{48}$$

satisfying the condition

$$\det\Big(\frac{\partial p_i}{\partial a^j}\Big) = \det\Big(\frac{\partial^2 S}{\partial x^i \partial a^j}\Big) \neq 0$$

(that arises from the requirement (47) and means that $(t, x, a) \to (t, x, \dot{x}) \to (t, x, p)$ are coordinate transformations on the evolution space), *provides first integrals of the Euler-Lagrange equations, given by the formula* $b_i = \partial S/\partial x^i$.

Equation (48) is called the Hamilton-Jacobi equation*, and we can see that it appears as a consequence of the geometrical Liouville theorem.* It is shown in [113, 111] that these ideas apply *mutatis mutandis* to contact manifolds where the Cartan form is replaced by the exterior derivative of the contact 1-form of the Reeb field.

2. Completely integrable distributions. [75] Let $\mathcal{D}$ be a *completely integrable* distribution of a constant rank on a manifold $\mathcal{N}$. It is clear that *if corank* $\mathcal{D}$ *is an even number,* $2p$*, then around every point in* $\mathcal{N}$ *there is a closed* 2*-form* α *of rank* $2p$ *such that* $\mathcal{D}$ *is the characteristic distribution of* α. Indeed, by Darboux Theorem, a local 2-form α with the above properties appears as eg.

$$\alpha = df_1 \wedge df_{p+1} + df_2 \wedge df_{p+2} + \cdots df_p \wedge df_{2p},$$

where $f_1, f_2, \ldots, f_{2p}$ are independent first integrals of the distribution $\mathcal{D}$; their existence is guaranteed by Frobenius Theorem.

To apply the Liouville integration method for getting maximal integral manifolds of $\mathcal{D}$ one needs to know in a neighbourhood of every point at least one appropriate 2-form α. This, however may be a problem (except, as we have seen above, when the given equations come from a Lagrangian!). On the other hand, it is clear, that one can try to find such a form directly, by solving the corresponding equations and conditions:

(1) $d\alpha = 0$,

(2) $\operatorname{rank} \alpha = 2p$,

(3) $i_{Z_k}\alpha = 0$, $1 \leq k \leq \operatorname{rank} \mathcal{D}$, where Z_k are independent generators of $\mathcal{D}$,

that as we know, are locally solvable. Effectively this means finding a (local) 1-form ρ, $\rho = g_i dy^i$, such that $\alpha = d\rho$, which produces equations for the functions g_i. Examples of this process for non-variational SODEs can be found in [113, 110].

5 Geometry of regular SODEs on $\mathbb{R} \times$ TM

In what follows we will be analysing a system of second order differential equations in normal or contravariant form,

$$\ddot{x}^a = f^a(t, x, \dot{x}) \tag{49}$$

on a manifold M with local coordinates (x^a) and with associated bundles $\pi : \mathbb{R}\times M \to M$, $\pi_0 : \mathbb{R}\times M \to \mathbb{R}$ and $\pi_{1,0} : E \to \mathbb{R}\times M$. As before the evolution space $E := \mathbb{R}\times TM$ has local adapted coordinates (t, x^a, u^a) or $(t, x^a, \dot{x}^a)$. The geodesic equations are, of course, a special example and one might expect the analysis to be modelled on that (autonomous) situation. However, we take the position that even the autonomous case is best described on extended configuration space $\mathbb{R}\times M$ and evolution space $\mathbb{R} \times TM$ and that one should put aside the historical fact that autonomous systems were discussed on M and its tangent and cotangent bundles. We warn the reader that in the remainder of this article the configuration space M has dimension n (rather than m): this is to maintain consistency with the extensive literature in this part of the topic.

5.1 A nonlinear connection and the Jacobi endomorphism

We follow [15, 21, 52, 53] and give the basic evolution space geometry of regular SODEs as it stood around 1985 (with some enhancements).

From (49) we construct on E a second-order differential equation field:

$$\Gamma = \frac{\partial}{\partial t} + u^a \frac{\partial}{\partial x^a} + f^a \frac{\partial}{\partial u^a} \tag{50}$$

whose integral curves are the 1-jets of the solution curves of the given equations.

The vertical and contact structures of the bundle $\pi_{1,0} : E \to \mathbb{R}\times M$ are combined in S, an intrinsic $(1, 1)$ tensor field on E and known as the *vertical endomorphism*. In coordinates:

$$S = V_a \otimes \omega^a. \tag{51}$$

where $V_a := \frac{\partial}{\partial u^a}$ are the vertical basis fields with respect to the bundle projection $\pi_{1,0}$, and $\omega^a := dx^a - u^a dt$ are the local contact forms as before. From the first order deformation $\mathcal{L}_\Gamma S$ a *nonlinear connection* is constructed as follows: $\mathcal{L}_\Gamma S$ has eigenvalues $0, 1, -1$ with corresponding eigenspaces spanned locally by Γ, the n *vertical* fields V_a and n *horizontal* fields

$$H_a := \frac{\partial}{\partial x^a} - \Gamma^b_a \frac{\partial}{\partial u^b}\,, \qquad \text{where } \Gamma^b_a := -\frac{1}{2}\frac{\partial f^b}{\partial u^a}, \tag{52}$$

respectively. Note that span$\{\Gamma, H_a\}$ is the horizontal space of this connection. (See [21] for an intrinsic development of the foregoing.)

The vector fields $\{\Gamma, H_a, V_a\}$ form a local basis on E, with dual basis $\{dt, \omega^a, \psi^a\}$ where

$$\psi^a := du^a - f^a dt + \Gamma^a_b \omega^b.$$

(Compare this basis with the one adapted to the second-order semispray connection given in section 2.) The Γ^b_a form the components of the nonlinear connection thus induced by Γ. The resulting direct sum decomposition of $T(E)$ is $I_E = P_\Gamma + P_H + P_V$ where I_E is the identity type $(1,1)$-tensor field on E and P_Γ, P_H and P_V are the three projection operators given in coordinates by

$$P_\Gamma = \Gamma \otimes dt\ , \quad P_H = H_a \otimes \omega^a\ , \quad P_V = V_a \otimes \psi^a. \tag{53}$$

The components of the Jacobi endomorphism (sometimes called the Douglas tensor), $\Phi := P_V \circ \mathcal{L}_\Gamma P_H$, a type (1,1) tensor field on E, can be calculated from

$$[\Gamma, H_a] = \Gamma^b_a H_b + \Phi^b_a V_b, \tag{54}$$

giving

$$\Phi = \Phi^b_a\, V_b \otimes \omega^a = \left(B^b_a - \Gamma^b_c \Gamma^c_a - \Gamma(\Gamma^b_a)\right) V_b \otimes \omega^a, \tag{55}$$

where $B^b_a := -\dfrac{\partial f^b}{\partial x^a}$. Other useful results:

$$[\Gamma, V_a] = -H_a + \Gamma^b_a V_b,\ \ [H_a, H_b] = R^d_{ab} V_d,\ \ [H_a, V_b] = V_b(\Gamma^c_a) V_c = [H_b, V_a]; \tag{56}$$

the second of these is effectively the definition of the curvature, R, of the nonlinear connection Γ^b_a. In coordinates

$$R^d_{ab} := \frac{1}{2}\left(\frac{\partial^2 f^d}{\partial x^a \partial u^b} - \frac{\partial^2 f^d}{\partial x^b \partial u^a} + \frac{1}{2}\left(\frac{\partial f^c}{\partial u^a}\frac{\partial^2 f^d}{\partial u^c \partial u^b} - \frac{\partial f^c}{\partial u^b}\frac{\partial^2 f^d}{\partial u^c \partial u^a}\right)\right).$$

The following identity is important:

$$V_a(\Phi^c_b) - V_b(\Phi^c_a) = 3R^c_{ab}. \tag{57}$$

In [21] vertical and horizontal lifts to E of vector fields on $\mathbb{R}\times M$ are intrinsically defined; here it suffices to give their coordinate descriptions. Given $X \in \mathfrak{X}(\mathbb{R}\times M)$ with coordinate representation $X = X^0 \frac{\partial}{\partial t} + X^a \frac{\partial}{\partial x^a}$ then

$$X^{\mathsf{v}} = (X^a - u^a X^0) V_a \qquad \text{and} \qquad X^{\mathsf{h}} = (X^a - u^a X^0) H_a.$$

This means, for example, that for any vertical vector $\mu \in T_q(E)$ there exists a unique vector $\eta \in T_{\pi_{1,0}(q)}(\mathbb{R}\times M)$ with $dt(\eta) = 0$ such that $\eta^{\mathsf{v}} = \mu$.

Using $\Phi := P_V \circ \mathcal{L}_\Gamma P_H$ and $I_E = P_\Gamma + P_H + P_V$ it is a simple matter to show that

$$P_V \circ \mathcal{L}_\Gamma P_V = -\Phi. \tag{58}$$

5.2 SODEs along the tangent bundle projection

In a series of papers [95, 96, 97, 128] Martínez, Cariñena, Sarlet *et al* developed the calculus of derivations along the tangent bundle projection with particular application to SODEs. While exterior calculus per se is not available in the scenario, there are some computational and conceptual advantages in this scheme which we will attempt to utilise.

So now we introduce vector fields and forms *along the projection* $\pi_{1,0} : E \to \mathbb{R} \times M$. We follow [130, 128]. Vector fields along $\pi_{1,0}$ are sections of the pull back bundle ${\pi_{1,0}}^*(T(\mathbb{R}\times M))$ over E. $\mathfrak{X}(\pi_{1,0})$ denotes the $C^\infty(E)$-module of such vector fields. Similarly, $\bigwedge(\pi_{1,0})$ denotes the graded algebra of scalar-valued forms along $\pi_{1,0}$ and $V(\pi_{1,0})$ denotes the $\bigwedge(\pi_{1,0})$-module of vector-valued forms along $\pi_{1,0}$. *Basic vector fields and 1-forms* along $\pi_{1,0}$ are elements of $\mathfrak{X}(\mathbb{R} \times M)$ and $\mathfrak{X}^*(\mathbb{R} \times M)$ respectively identified with vector fields and forms along $\pi_{1,0}$ by composition with $\pi_{1,0}$. Using this device tensor fields along the projection can be expressed as tensor products of basic vector fields and 1-forms with coefficients in $C^\infty(E)$. The canonical vector field along $\pi_{1,0}$ is

$$\mathbf{T} = \frac{\partial}{\partial t} + u^a \frac{\partial}{\partial x^a},$$

and the natural bases for $\mathfrak{X}(\pi_{1,0})$ and $\mathfrak{X}^*(\pi_{1,0})$ are then $\{\mathbf{T}, \frac{\partial}{\partial x^a}\}$ and $\{dt, \omega^a\}$. The set of equivalence classes of vector fields along $\pi_{1,0}$ modulo $\mathbf{T}$ is denoted $\overline{\mathfrak{X}(\pi_{1,0})}$ so that $\overline{X} \in \overline{\mathfrak{X}(\pi_{1,0})}$ satisfies $dt(\overline{X}) = 0$. Then the obvious bijection between $\overline{\mathfrak{X}(\pi_{1,0})}$ and $V(E)$ provides a *vertical lift* from $\mathfrak{X}(\pi_{1,0})$ to $V(E)$, given in coordinates by:

$$X^V = \overline{X}^a \frac{\partial}{\partial u^a} = (X^a - u^a X^0)\frac{\partial}{\partial u^a}$$

where $X = X^0 \frac{\partial}{\partial t} + X^a \frac{\partial}{\partial x^a}$.

On the matter of horizontal lifts we part company with [128], following [53], and say that the *horizontal lift* X^H of $X \in \mathfrak{X}(\pi_{1,0})$ is given by $X^H = \overline{X}^a H_a$. (There are many reasons for this: for example, it is consistent with the horizontal lift of [21] and it respects the eigenvector structure of $\mathcal{L}_\Gamma S$, for this reason it is also known as the *strong* horizontal lift, see [28].) Finally, we can *lift along* Γ by $X^\Gamma := dt(X)\Gamma$ for any $X \in \mathfrak{X}(\pi_{1,0})$ (so that $\mathbf{T}^\Gamma = \Gamma$). Then any vector field $W \in \mathfrak{X}(E)$ can be decomposed as

$$W = (W_\Gamma)^\Gamma + (W_H)^H + (W_V)^V$$

for unique $W_\Gamma \in \text{span}\{\mathbf{T}\}$, $W_H \in \mathfrak{X}(\pi_{1,0})$ with $W_H(t) = W(t)$ and $W_V \in \overline{\mathfrak{X}(\pi_{1,0})}$. This decomposition is the main aim of the lifting exercise. In coordinates,

$$\begin{aligned}
W_\Gamma &= dt(W)\mathbf{T} \, , \\
W_H &= dt(W)\frac{\partial}{\partial t} + dx^a(W)\frac{\partial}{\partial x^a} = dt(W)\mathbf{T} + \omega^a(W)\frac{\partial}{\partial x^a} \, , \\
W_V &= \psi^a(W)\frac{\partial}{\partial x^a} \, .
\end{aligned}$$

The *dynamical covariant derivative* ∇ and the Jacobi endomorphism, Φ, are then defined as objects along the projection through the following commutation relations on E:

$$[\Gamma, X^V] = -X^H + (\nabla X)^V \quad \text{and} \quad [\Gamma, X^H] = (\nabla X)^H + \Phi(X)^V. \tag{59}$$

In coordinates $\Phi = \Phi^a_b \frac{\partial}{\partial x^a} \otimes \omega^b$ (we make no notational distinction between the Jacobi endomorphism in this context and in that of the previous section). We extend ∇ to act on forms by setting $\nabla(F) := \Gamma(F)$ for $F \in \bigwedge^0(\pi_{1,0})$; then it can be shown that $\nabla(\langle X, \alpha\rangle) = \langle \nabla X, \alpha\rangle + \langle X, \nabla\alpha\rangle$ and so ∇ can be extended to tensor fields along $\pi_{1,0}$ in the usual way. $\nabla \mathbf{T} = 0$ and, in coordinates,

$$\nabla\omega^a = -\Gamma^a_b\omega^b, \quad \nabla dt = 0, \quad \nabla\frac{\partial}{\partial x^a} = \Gamma^b_a\frac{\partial}{\partial x^b}.$$

5.3 The Massa and Pagani connection

Massa and Pagani [98], Byrnes [11, 12], Crampin *et al.* [16] and Mestdag and Sarlet [125, 101, 100] have separately proposed various linear connections on E induced by a SODE Γ. They all use the dynamical covariant derivative ∇ to determine derivatives *along* Γ, but differ in the derivatives *of* Γ. This is essentially equivalent to different choices of the torsion of the connection.

Massa and Pagani introduce a linear connection on E by imposing some natural requirements. If we denote their connection by $\hat{\nabla}$, these are that the covariant differentials $\hat{\nabla}dt$, $\hat{\nabla}S$, and $\hat{\nabla}\Gamma$ are all zero and that the vertical sub-bundle is flat.

Crampin *et al.* [16] firstly define a covariant derivative along $\pi_{1,0}$ and then induce one on E by lifting; the Massa and Pagani connection on E can be produced in the same way (see [53] for the details of what follows).

For each $Y \in \mathfrak{X}(E)$, $U \in \mathfrak{X}(\pi_{1,0})$ and $f \in C^\infty(E)$,

$$\hat{D}_Y U := [P_H(Y), U^V]_V + [P_\Gamma(Y), U^V]_V + [P_V(Y), U^H]_H + Y(U(t))\mathbf{T},$$
$$\hat{D}_Y(f) := Y(f)$$

is a covariant derivative along the projection. Note that $\hat{D}_\Gamma = \nabla$ and it is useful to introduce the further notations

$$\hat{D}^H_Y := \hat{D}_{Y^H} \quad \text{and} \quad \hat{D}^V_Y := \hat{D}_{Y^V}.$$

The components of $\hat{D}$ are as follows:

$$\hat{D}_\Gamma\mathbf{T} = 0 \qquad \hat{D}_{H_a}\mathbf{T} = 0 \qquad \hat{D}_{V_a}\mathbf{T} = 0$$
$$\hat{D}_\Gamma\frac{\partial}{\partial x^a} = \Gamma^b_a\frac{\partial}{\partial x^b} \qquad \hat{D}_{H_b}\frac{\partial}{\partial x^a} = \frac{\partial\Gamma^c_a}{\partial u^b}\frac{\partial}{\partial x^c} \qquad \hat{D}_{V_b}\frac{\partial}{\partial x^a} = 0\,.$$

We now use $\hat{D}$ to recover the linear connection $\hat{\nabla}$ of Massa and Pagani on E in the manner of [16].

$$\hat{\nabla}_Y X := (\hat{D}_Y X_\Gamma)^\Gamma + (\hat{D}_Y X_H)^H + (\hat{D}_Y X_V)^V,$$
$$\hat{\nabla}_Y(f) := Y(f)$$

for all $Y, X \in \mathfrak{X}(E)$ and $f \in C^\infty(E)$ is a linear covariant derivative.

(That this is the Massa and Pagani connection can be verified by calculating the covariant differentials of S, dt and Γ along with $\hat{\nabla}_{V_a}X$ or directly from the components below:)

$$\hat{\nabla}_\Gamma \Gamma = 0, \qquad \hat{\nabla}_\Gamma H_a = \Gamma^b_a H_b, \qquad \hat{\nabla}_\Gamma V_a = \Gamma^b_a V_b,$$
$$\hat{\nabla}_{H_a} \Gamma = 0, \qquad \hat{\nabla}_{H_a} H_b = \frac{\partial \Gamma^c_a}{\partial u^b} H_c, \qquad \hat{\nabla}_{H_a} V_b = \frac{\partial \Gamma^c_a}{\partial u^b} V_c,$$
$$\hat{\nabla}_{V_a} \Gamma = 0, \qquad \hat{\nabla}_{V_a} H_b = 0, \qquad \hat{\nabla}_{V_a} V_b = 0.$$

A key feature of $\hat{\nabla}$ for us is that $\hat{\nabla}_X \Gamma = 0$ for all $X \in \mathfrak{X}(E)$. It's also worth noting the following important facts:

Let $X, Y \in \mathfrak{X}(\pi_{1,0})$. Then

$$\hat{\nabla}_\Gamma X^V = (\nabla X)^V, \qquad \hat{\nabla}_\Gamma X^H = (\nabla X)^H,$$
$$\hat{\nabla}_{Y^H} X^H = (\hat{D}_{Y^H} X)^H, \qquad \hat{\nabla}_{Y^H} X^V = (\hat{D}_{Y^H} X)^V,$$
$$\hat{\nabla}_{Y^V} X^H = (\hat{D}_{Y^V} X)^H, \qquad \hat{\nabla}_{Y^V} X^V = (\hat{D}_{Y^V} X)^V,$$
$$\hat{\nabla}_{Y^H} X^\Gamma = Y^H(dt(X))\Gamma, \qquad \hat{\nabla}_{Y^V} X^\Gamma = Y^V(dt(X))\Gamma.$$

Now with every linear connection there is an associated torsion and shape map (see [52, 53]). The torsion is defined by

$$\hat{T}(X, Y) := \hat{\nabla}_X Y - \hat{\nabla}_Y X - [X, Y].$$

It will be useful later to have the following identity (valid for any connection) for a two-form Ω

$$\begin{aligned} d\Omega(X, Y, Z) = &\hat{\nabla}_X \Omega(Y, Z) + \hat{\nabla}_Y \Omega(Z, X) + \hat{\nabla}_Z \Omega(X, Y) \\ &+ \Omega(\hat{T}(X, Y), Z) + \Omega(\hat{T}(Y, Z), X) + \Omega(\hat{T}(Z, X), Y). \end{aligned} \tag{60}$$

The shape map A_Z associated with a vector field $Z \in \mathfrak{X}(E)$ is an endomorphism of tangent spaces of E constructed from Lie and parallel transport along the flow, $\{\zeta_t\}$, of Z.

If $\tau_t : T_x E \to T_{\zeta_t(x)} E$ is the parallel transport map along $\{\zeta_t\}$, then

$$A_Z(\xi_x) := \left.\frac{d}{dt}\right|_{t=0} (\tau_t^{-1} \circ \zeta_{t*})(\xi_x)$$

measures the deformation of tangent spaces by the flow. It is a simple matter to show that

$$A_Z = \hat{\nabla} Z + \hat{T}_Z = \hat{\nabla}_Z - \mathcal{L}_Z,$$

where $\hat{T}_Z(X) := \hat{T}(Z, X)$ and $(\hat{\nabla} Z)(X) := \hat{\nabla}_X Z$.

In our case the torsion contains *all* the significant geometry of the SODE:

$$\hat{T}(\Gamma, V_a) = H_a, \qquad \hat{T}(\Gamma, H_a) = -\Phi^b_a V_b,$$
$$\hat{T}(V_a, V_b) = 0, \qquad \hat{T}(V_a, H_b) = 0,$$
$$\hat{T}(H_a, H_b) = -R^c_{ab} V_c.$$

Since Γ is auto-parallel we have $A_\Gamma(X) = \hat{T}(\Gamma, X)$ and

$$A_\Gamma = -P_V \circ \mathcal{L}_\Gamma P_H - P_H \circ \mathcal{L}_\Gamma P_V = -\Phi - P_H \circ \mathcal{L}_\Gamma P_V,$$

in coordinates

$$A_\Gamma = -\Phi^a_b V_a \otimes \omega^b + H_a \otimes \psi^a,$$

which gives geometric insight into the Jacobi endomorphism.

The Riemann curvature of $\hat{\nabla}$ is given in the usual way by

$$\mathfrak{R}(X,Y)Z := \hat{\nabla}_X \hat{\nabla}_Y Z - \hat{\nabla}_Y \hat{\nabla}_X Z - \hat{\nabla}_{[X,Y]} Z,$$

with components (see [1] for a full development including Bianchi identities, etcetera)

$$\begin{aligned}
&\mathfrak{R}(X,Y)\Gamma = 0, \qquad \mathfrak{R}(V_a, V_b)X = 0, \qquad \mathfrak{R}(\Gamma, V_a)X = 0,\\
&\mathfrak{R}(\Gamma, H_a)V_b = (-R^c_{ab} - V_b(\Phi^c_a))V_c, \qquad \mathfrak{R}(\Gamma, H_a)H_b = (-R^c_{ab} - V_b(\Phi^c_a))H_c,\\
&\mathfrak{R}(V_a, H_b)V_c = V_a(\Gamma^d_{bc})V_d, \qquad \mathfrak{R}(V_a, H_b)H_c = V_a(\Gamma^d_{bc})H_d,\\
&\mathfrak{R}(H_a, H_b)V_c = -V_c(R^d_{ab})V_d, \qquad \mathfrak{R}(H_a, H_b)H_c = -V_c(R^d_{ab})H_d,
\end{aligned}$$

where $\Gamma^a_{bc} := V_b(\Gamma^a_c)$.

Equipped with the Massa and Pagani connection a very rich geometry of SODEs waits to be explored. Such a geometry holds the promise of extensions of the qualitative techniques of the geometry of geodesics to a much broader class of differential equations.

6 The inverse problem for semisprays

In this section we will describe a particular thread of mathematical development in the inverse problem of the calculus of variations. While it has not yet provided a solution to, for example, the Douglas problem for $n = 3$, it has provided deep insight into both the quantitative and qualitative structure of regular second order ordinary differential equations, uncovering the remarkable universal features shared by them all, features previously thought to belong only to special classes. This line of research continues to be productive as the prospect of a solution for $n = 3$ gets closer.

6.1 History and setting of the problem

The inverse problem for second order equations in normal form has a rather different history and current state to the problem in covariant form as discussed in section 3.2 and elsewhere in section 3. This is essentially because in the covariant case we ask if the system *as it stands* is variational and in the contravariant case we have to search for a variational covariant form. So the inverse problem for semisprays involves deciding whether the solutions of a given system of second-order ordinary differential equations (49), namely

$$\ddot{x}^a = f^a(t, x, \dot{x}),$$

are the solutions of a set of Euler-Lagrange equations

$$\frac{\partial^2 L}{\partial \dot{x}^a \partial \dot{x}^b}\ddot{x}^b + \frac{\partial^2 L}{\partial x^b \partial \dot{x}^a}\dot{x}^b + \frac{\partial^2 L}{\partial t \partial \dot{x}^a} = \frac{\partial L}{\partial x^a}, \quad a, b = 1, \ldots, n \tag{61}$$

for some Lagrangian function $L(t, x^b, \dot{x}^b)$.

Because the Euler-Lagrange equations are not generally in normal form, the problem is to find a so-called (non-degenerate) multiplier matrix $g_{ab}(t, x^c, \dot{x}^c)$ such that

$$g_{ab}(\ddot{x}^b - f^b) \equiv \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^a}\right) - \frac{\partial L}{\partial \dot{x}^a}.$$

As in previous sections we use the notation g_{ab} to stress that the multipliers we consider are regular (non-degenerate).

The most commonly used set of necessary and sufficient conditions for the existence of the g_{ab} are the so-called *Helmholtz conditions* due to Douglas [30] and put in the following form by Sarlet [118]:

$$g_{ab} = g_{ba}, \quad \Gamma(g_{ab}) = g_{ac}\Gamma^c_b + g_{bc}\Gamma^c_a, \quad g_{ac}\Phi^c_b = g_{bc}\Phi^c_a, \quad \frac{\partial g_{ab}}{\partial u^c} = \frac{\partial g_{ac}}{\partial u^b}, \qquad (62)$$

where we have replaced $\dot{x}$ by u and utilised all our notations to date.

These algebraic-differential conditions ultimately require the application of a theory of integrability in order to determine the existence and uniqueness of their solutions. To date the integrability theories that have been used are associated with the names of Riquier-Janet, Cartan-Kähler and Spencer. Of these we will outline only the use of the Cartan-Kähler theorem in its exterior differential systems manifestation (in section 6.3).

Before proceeding with the mathematical description and analysis, we provide the reader with some historical perspective into this local inverse problem for second order ordinary differential equations.

Helmholtz [47] first discussed whether systems of second order ordinary differential equations are Euler-Lagrange for a first-order Lagrangian (that is, one depending on velocities but not accelerations) *in the form presented* (the covariant inverse problem), and found necessary conditions for this to be true. Mayer [99] later proved that the conditions are also sufficient.

However, in 1886, a year earlier than Helmholtz published his celebrated result, in a paper that unfortunately remained unknown for years, Sonin [133] found out that *one* SODE

$$\ddot{x} - f = 0 \qquad (63)$$

can always be put into the form of an Euler-Lagrange equation by multiplying $\ddot{x} - f$ *by a suitable function* $g \neq 0$. He also characterized the multiplicity of the solution, i.e. provided a description of all Lagrangians for (63). Now, Sonin's result can be proved easily using the Helmholtz conditions, that for one equation (63) reduce to a single partial differential equation for the unknown function $g(t, x, \dot{x})$:

$$\frac{\partial g}{\partial t} + \frac{\partial g}{\partial x}\dot{x} + f\frac{\partial g}{\partial \dot{x}} + g\frac{\partial f}{\partial \dot{x}} = 0.$$

Since $g \neq 0$, this equation takes the form

$$\frac{\partial \ln g}{\partial t} + \frac{\partial \ln g}{\partial x}\dot{x} + \frac{\partial \ln g}{\partial \dot{x}}f + \frac{\partial f}{\partial \dot{x}} = 0,$$

that is well-known be solvable; its general solution depends upon a single arbitrary function of any two specific solutions of the corresponding homogeneous equation. Consequently, the most general Lagrangian for (63) depends upon one arbitrary function of two parameters.

Later Hirsch [50] formulated independently, and in a more general setting, the *multiplier problem*, that is, the question of the existence of multiplier functions which convert a system of second order ordinary differential equations in normal form into Euler-Lagrange equations. Surprisingly it turned out that *a solution to the multiplier problem need not exist if there is more than one equation*. Hirsch gave certain self-adjointness conditions for the problem but they are not effective in classifying second order equations according to the existence and uniqueness of the corresponding multipliers.

This multiplier problem was completely solved by Douglas in 1941 [30] for two degrees of freedom, that is, a pair of second order equations on the plane. He produced an exhaustive classification of all such equations in normal form. In each case Douglas identified all (if any) Lagrangians producing Euler-Lagrange equations whose normal form is that of the equations in that particular case. His method avoided Hirsch's self-adjointness conditions and he produced his own necessary and sufficient algebraic-differential conditions. His approach was to generate a sequence of integrability conditions, solving these using Riquier-Janet theory. While this approach is singularly effective and forms the basis of current efforts, it has been particularly difficult to see how to cast it into a form suitable for higher dimensions.

Interest from the physics community in the non-uniqueness aspects of the inverse problem provided the next contribution to solving the Helmholtz conditions. Henneaux [48] and Henneaux and Shepley [49] developed an algorithm for solving the Helmholtz conditions for *any given system of second order equations*. In particular, they solved the problem for spherically symmetric problems in dimension 3. In this fundamental case Henneaux and Shepley showed that a two-parameter family of Lagrangians produce the same equations of motion. Startlingly these Lagrangians produced inequivalent quantum mechanical hydrogen atoms. Further mathematical aspects of this case were elaborated by Crampin and Prince [17, 19].)

At around the same time Sarlet [118] showed that the part of the Helmholtz conditions which ensures the correct time evolution of the multiplier matrix could be replaced by a possibly infinite sequence of purely algebraic initial conditions. Along with the work of Henneaux this provided a prototype for geometrising Douglas's Helmholtz conditions.

Over the next 10 years or so Cantrijn, Cariñena, Crampin, Ibort, Marmo, Prince, Sarlet, Saunders and Thompson explored the tangent bundle geometry of second order ordinary differential equations in general and the Euler-Lagrange equations in particular. The inverse problem provided central inspiration for their examination of the integrability theorems of classical mechanics, multi-Lagrangian systems, geodesic first integrals and equations with symmetry. Using the geometrical approach to second order equations of Klein and Grifone [41, 42, 56, 57], the Helmholtz conditions for non-autonomous second order equations on a manifold were reformulated in terms of the corresponding non-linear connection on its tangent bundle (see section 5.1). This occurred in 1985 after a sequence of papers [15, 21, 118]. The work of Sarlet [121], and collectively Martínez, Cariñena and Sarlet [95, 96, 97] on derivations along the tangent bundle projection (see section 5.2) opened the way to the geometrical reformulation of Douglas's **solution** of the two-degree of freedom case. This was achieved in 1993 by Crampin, Sarlet, Martínez, Byrnes and Prince and is reported in [23]. A number of dimension n classes were subsequently solved ([124, 123, 20]). The reader is directed to the review by Prince [110] for more details of this program up to the turn of the current century.

Separately Anderson and Thompson [7] applied exterior differential systems theory to some special cases of the geometrised problem with considerable success. In order to pursue the EDS approach Aldridge [1] used the Massa and Pagani connection of section 5.3 and recovered all the dimension n results to date along with an overall classification scheme for this general case. It appears that the inverse problem still holds many accessible secrets.

6.2 The general problem: geometric formulations

In this section we outline the progress in the geometric formulation and solution of the inverse problem for semisprays to be found in the work of Aldridge, Cantrijn, Cariñena, Crampin, Martínez, Prince, Sarlet, Thompson and their collaborators. See [1, 3, 14, 15, 19, 20, 21, 23, 118, 124, 127]. In what follows we denote the Cartan form of theorem 3.1 and equation (17) by θ_L rather than θ_λ. A natural starting point is the following proposition from Goldschmidt and Sternberg, see [37, 136] and compare with proposition 4.2:

Proposition 6.1 *If L is a regular Lagrangian (so that the matrix whose entries are $\dfrac{\partial^2 L}{\partial u^a \partial u^b}$ is everywhere nonsingular), then there is a unique vector field Γ, called the Euler-Lagrange field, on E such that*

$$i_\Gamma d\theta_L = 0 \quad \textit{and} \quad dt(\Gamma) = 1.$$

This vector field is a SODE, *and the equations satisfied by its integral curves are the Euler-Lagrange equations for L.*

By careful observation of the properties of the Cartan 2-form, $d\theta_L$, and following the work of Klein [56, 57], Grifone [43, 44] and Crampin [15] in the autonomous case, Crampin, Prince and Thompson [21] give the fundamental geometric version of the Helmholtz conditions:

Theorem 6.2 *(Crampin, Prince & Thompson 1984) Given a* SODE Γ, *necessary and sufficient conditions for the existence of a regular Lagrangian, whose Euler-Lagrange field is Γ, are that there exists $\Omega \in \bigwedge^2(E)$ such that*

(1) Ω *has maximal rank*

(2) $\Omega(V_1, V_2) = 0,\ \forall\, V_1, V_2 \in V(E)$

(3) $i_\Gamma \Omega = 0$

(4) $d\Omega = 0$

The above theorem is a "contravariant" version of Theorem 3.2 adapted to *semisprays* (i.e. where the assumption on regularity of the corresponding dynamical form ε is added). Its benefit is a *direct construction of the related closed* 2*-form* (the Lepage equivalent of ε) *in terms of the vector field* Γ.

The 2-form Ω is modelled on the Cartan 2-form $d\theta_L$ and the necessity of the conditions in the theorem follows from the fact that $d\theta_L = \dfrac{\partial^2 L}{\partial u^a \partial u^b}\psi^a \wedge \omega^b$. The sufficiency of the conditions means that $\Omega = g_{ab}\psi^a \wedge \omega^b$ where g_{ab} satisfies the Helmholtz conditions (62).

This appearance of the Helmholtz conditions is rather surprising since the statement of the theorem seems to entail quite general conditions on the characteristic vector field of a maximal rank 2-form on an odd-dimensional manifold. It is, however, the fact that the manifold is an evolution space and the vector field is a SODE that produces the interesting structure.

The Helmholtz conditions can be made to appear explicitly in the following way. It is straightforward to show that the only non-zero components of Ω are $\Omega(V_a, H_b)$ so that $\Omega = g_{ab}\psi^a \wedge \omega^b$ for some functions g_{ab} on E. The closure conditions $d\Omega(X, Y, Z) = 0$ then produce Helmholtz conditions (the maximal rank condition $\det(g_{ab}) \neq 0$ has to be applied separately).

The simplest way to see how the Helmholtz conditions in Sarlet's form (62) arise from theorem 6.2 is to put $\Omega := g_{ab}\psi^a \wedge \omega^b$ and compute $d\Omega$:

$$\begin{aligned} d\Omega = &(\Gamma(g_{ab}) - g_{cb}\Gamma^c_a - g_{ac}\Gamma^c_b)dt \wedge \psi^a \wedge \omega^b \\ &+ (H_d(g_{ab}) - g_{cb}V_a(\Gamma^c_d))\psi^a \wedge \omega^b \wedge \omega^d \\ &+ V_c(g_{ab})\psi^c \wedge \psi^a \wedge \omega^b \\ &+ g_{ab}\psi^a \wedge \psi^b \wedge dt \\ &+ g_{ca}\Phi^c_b\omega^a \wedge \omega^b \wedge dt \\ &+ g_{ca}H_b(\Gamma^c_d)\omega^a \wedge \omega^b \wedge \omega^d. \end{aligned}$$

The four Helmholtz conditions are

$$\begin{aligned} &d\Omega(\Gamma, V_a, V_b) = 0, \qquad &&d\Omega(\Gamma, V_a, H_b) = 0, \\ &d\Omega(\Gamma, H_a, H_b) = 0, \qquad &&d\Omega(H_a, V_b, V_c) = 0. \end{aligned}$$

The remaining conditions arising from $d\Omega = 0$, namely

$$d\Omega(H_a, H_b, V_c) = 0 \qquad \text{and} \qquad d\Omega(H_a, H_b, H_c) = 0,$$

can be shown to be derivable from the first four (notice that this last condition is void in dimension 2).

The further development of this geometric line of enquiry relied on two-fold inspiration from Douglas's work. Firstly, Douglas generates hierarchies of integrability conditions on the basic Helmholtz conditions by repeated use of the dynamical covariant derivative. Secondly, he utilises the various possible Jordan normal forms of Φ. By combining these two devices he creates his famous exhaustive classification scheme for the two degrees of freedom case. In 1992 Anderson and Thompson [7] used Douglas's hierarchies and the basic differential geometry of SODEs on the evolution space (outlined in section 5.1) to construct an exterior differential systems approach to the inverse problem. Amongst other things they completely solved the case where Φ is a multiple of the identity for arbitrary n. However, they did not pursue other possible Jordan normal forms. In the 1994 paper [23], Crampin, Sarlet, Martínez, Byrnes and Prince give a geometric outline of Douglas's two-fold approach and point the way forward for a study of the higher dimensional cases. This 1994 formulation was based upon the calculus of derivations along the tangent bundle projection, an approach which does not allow direct access to EDS techniques. Nonetheless the tangent bundle projection calculus was subsequently pursued in [20, 124, 123] and a number of important cases for arbitrary n were solved in the Riquier-Janet framework. Finally, Douglas's entire analysis was geometrised in this sense by Sarlet, Thompson and Prince in [127]. We will return to the EDS study of the inverse problem and the role of the Massa and Pagani connection in the next two sections.

Before we turn to a discussion of the integrability of the Helmholtz conditions we will give the formulation of theorem 6.2 in terms of the calculus along the tangent bundle projection described in section 5.2. The observation we made after theorem 6.2 about the simple structure of $d\theta_L$ in the adapted co-frame means that the 2-form Ω on E of theorem 6.2 which we seek is completely determined by a symmetric nondegenerate type (0,2) tensor along π, of the form $g = g_{ab}\omega^a \otimes \omega^b$ (i.e. g vanishes on $\mathbf{T}$). To be precise, Ω is the so-called Kähler lift of g, $\Omega = g^K$, which vanishes on Γ and satisfies

$$g^K(X^V, Y^V) = g^K(X^H, Y^H) = 0, \qquad g^K(X^V, Y^H) = g(X, Y).$$

In this formulation of the conditions (62) then reads

$$\nabla g = 0, \qquad g(\Phi X, Y) = g(X, \Phi Y), \qquad \hat{D}^V_X g(Y, Z) = \hat{D}^V_Y g(X, Z). \tag{64}$$

Loosely speaking the integrability conditions on (64) are generated by repeated differentiation by ∇, and by application of the commutators of $\hat{D}^V$, $\hat{D}^H$ and $\hat{D}^\Gamma = \nabla$.

Repeated differentiation with ∇ produces the two hierarchies

$$g(\nabla^r \Phi(X), Y) = g(\nabla^r \Phi(Y), X) \qquad r = 1, 2, 3, \dots \tag{65}$$

and

$$\sum_{(XYZ)} g(\nabla^r R(X, Y), Z) = 0 \qquad r = 0, 1, 2, \dots, \tag{66}$$

where the sum is over even permutations of X, Y, Z and where R is a type $(1, 2)$ tensor-field along $\pi_{1,0}$, which is a component of the curvature of the connection $\hat{D}$, and is related to Φ by $\hat{D}^V_X \Phi(Y) - \hat{D}^V_Y \Phi(X) = 3R(X, Y)$ (cf equation (57)).
Application of the commutators produces the second order conditions

$$\begin{aligned} &\hat{D}^V_U \hat{D}^H_Z g(Y, X) + g(\theta(Y, X)Z, U) \\ &\qquad = \hat{D}^V_Z \hat{D}^H_U g(Y, X) + g(Z, \theta(Y, X)U), \end{aligned}$$

where θ is a type $(1, 3)$ tensor field along $\pi_{1,0}$ (here written as a type $(1, 1)$ tensor valued twice covariant tensor), which is another part of the curvature of the connection $\hat{D}$, and which is symmetric in all of its arguments as a consequence of the first Bianchi identities. There are two remarks which should be made about these integrability conditions. The first is that there is some question about the claim in [124, 123] that they are complete (that is, sufficient as well as necessary). Aldridge [1] explicitly argues that there are others, and Grifone and Muzsnay [43, 44], using Spencer theory and a different starting point produce what appear to be many more independent conditions. The second is that Thompson [139, 1] has shown that there is a termination theorem for both the above hierarchies. That is, when, for a given r, a condition adds no new algebraic information about g, no further information is available from the hierarchy.

In [23] it is shown the first broad classification of Douglas in [30] is by the linear dependence or independence of Φ and its ∇ derivatives, as follows.

Case I: Φ is a multiple of the identity tensor I.

Case II: $\nabla\Phi$ is a linear combination of Φ and I.

Case III: $\nabla^2\Phi$ is a linear combination of $\nabla\Phi$, Φ and I.

Case IV: $\nabla^2\Phi$, $\nabla\Phi$, Φ and I are linearly independent.

Further subcases arise according to the diagonalisability or Φ and the integrability of the corresponding eigenspaces. Unfortunately, a classification in the n degrees of freedom case does not arise so easily, but a number of important cases do appear, for example when Φ is a multiple of the identity and when Φ is diagonalisable with distinct eigenvalues and each of the corresponding two dimensional (on $\mathbb{R}\times TM$) eigenspaces is Frobenius integrable. These cases are shown to be variational in [124] and [20] respectively.

While this geometric formulation along the projection is satisfying and lends itself to an analysis using the Jordan normal forms of Φ, it simultaneously suffers from its reliance on Riquier-Janet theory and its incompatibility with EDS. Using the Massa & Pagani connection and the shape map, A_Γ, Aldridge [1] produced a corresponding geometric formulation of theorem 6.2 on $\mathbb{R}\times TM$ which is amenable to EDS:

Theorem 6.3 *(Aldridge 2003)*
Given a SODE Γ, *necessary and sufficient conditions for the existence of a Lagrangian, whose Euler-Lagrange field is* Γ, *are that there exists* $\Omega \in \bigwedge^2(E)$:

1. Ω *has maximal rank.*
2. $\Omega(V_1, V_2) = 0$ *for all vertical* $V_1,\ V_2$,
3. $\Omega(A_\Gamma(X), Y) = \Omega(A_\Gamma(Y), X)$,
4. $\hat{\nabla}_\Gamma\Omega = 0$,
5. $(\hat{\nabla}_{Z^V}\Omega)(X^V, Y^H) = (\hat{\nabla}_{X^V}\Omega)(Z^V, Y^H)$.

Aldridge also produces the following attractive result whose proof relies on equation (60).

Proposition 6.4 *The algebraic conditions arising from* $d\Omega = 0$ *are all consequences of*

$$\Omega(\hat{T}(X,Y),Z) + \Omega(\hat{T}(Z,X),Y) + \Omega(\hat{T}(Y,Z),X) = 0. \tag{67}$$

In this picture the two hierarchies of conditions (65),(66) become

$$\Omega((\overset{r}{\hat{\nabla}}_\Gamma A_\Gamma)X^H, Y^H) - \Omega((\overset{r}{\hat{\nabla}}_\Gamma A_\Gamma)Y^H, X^H) = 0,$$
$$\sum_{(XYZ)} \Omega((\overset{r}{\hat{\nabla}}_\Gamma R)(X^H, Y^H), Z^H) = 0, \quad r = 0, 1, \ldots.$$

Following the work of Aldridge, there has been renewed interest in applying EDS theory to the inverse problem, see [3]. In order to do this we return to the evolution space $\mathbb{R}\times TM$ where we can do exterior calculus. The main thrust is to pick up where Anderson and Thompson left off and to use the Jordan normal forms of Φ. As we will see in section 6.4, the EDS process produces the integrability conditions in a satisfactory manner and as a bonus gives us a handle on the maximal rank condition.

The standard EDS reference is the book [10] and for the inverse problem Anderson and Thompson's memoir [7]. We will give a brief synopsis of the method in this context.

6.3 EDS and the inverse problem: outline

The EDS process for finding the two forms of the inverse problem involves three steps: finding a differential ideal, creating an equivalent linear Pfaffian system, and lastly using the Cartan - Kähler theorem to determine the final generality of the solution.

The task is to find all the closed, maximal rank two forms on E of the form

$$g_{ab}\psi^a \wedge \omega^b.$$

So let Σ be the submodule of two forms span$\{\psi^a \wedge \omega^b\}$, and let $\{\Omega^k\}$ be a subset of two forms in Σ. Initially we take $\{\Omega^k : k = 1, \ldots, n^2\}$ to be some basis for Σ. Then the inverse problem becomes that of finding the submodule of closed, maximal rank two forms in Σ, i.e. finding functions r_k such that $d(r_k\Omega^k) = 0$. *Note that $\{\Omega^k\}$ is a working subset of Σ which will shrink as we progress.*

The first EDS step is to find the maximal submodule, Σ', of Σ that generates a differential ideal (that is, an ideal closed under exterior differentiation). We will find (or not) our closed two forms in this ideal.

We use the following recursive process: starting with the submodule $\Sigma^0 := \Sigma$ and a basis $\{\Omega^k\}$, find the submodule $\Sigma^1 \subseteq \Sigma^0$ such that $d\Omega \in \langle\Sigma^0\rangle$ for all $\Omega \in \Sigma^1$. That is, find the functions r_k on $\mathbb{R}\times TM$ such that $d(r_k\Omega^k) \in \langle\Sigma^0\rangle$ and hence $r_k d\Omega^k \in \langle\Sigma^0\rangle$. This is an algebraic problem.

Having found these r_k and hence Σ^1, we check if $\Sigma^1 = \Sigma^0$ and so is already a differential ideal. If not, we iterate the process, finding the submodule $\Sigma^2 \subset \Sigma^1 \subset \Sigma^0$ and so on until at some step, a differential ideal is found or the empty set is reached. If at any point during this process it is not possible to create a maximal rank two form, the inverse problem has no solution. That is, if $\{\Omega^1, ..., \Omega^d\}$ is a basis for Σ^i, then $\wedge^n(\sum_{k=1}^d \Omega^k)$ must be non-zero at each step.

Equipped with Σ', the next step in the EDS process is to express the problem of finding the closed two forms in Σ' as a Pfaffian system. Let the differential ideal $\langle\Sigma'\rangle$ be spanned by the set $\{\Omega^k :\ k = 1, \ldots, d\}$, and calculate

$$d\Omega^k = \xi^k_h \wedge \Omega^h$$

where the ξ^i_j's are now known one forms.

Since $d\Omega = \xi_j \wedge \Omega^j$ for all $\Omega \in \Sigma'$, and we are looking for those Ω's such that $d\Omega = 0$, we find all possible d-tuples of one forms $(\rho^A_k) = (\rho^A_1, ..., \rho^A_d)$ such that $\rho^A_k \wedge \Omega^k = 0$, ($A = 1, \ldots, e$ say). Then if

$$0 = d(r_k\Omega^k) = dr_k \wedge \Omega^k + r_k\xi^k_h \wedge \Omega^h = (dr_k + r_h\xi^h_k) \wedge \Omega^k$$

we must find e functions p_A on $\mathbb{R}\times TM$ with

$$dr_k + r_h\xi^h_k = p_A\rho^A_k. \tag{68}$$

At this point, the problem becomes that of solving (68) for r_k in terms of the unknown function p_A, and secondly, finding the restrictions on the choice of these p_A's. Having found the r_k's we can easily construct the g_{ab}, and hence the Lagrangians, or simply take the Cartan two form(s) $\Omega = r_k\Omega^k$.

The general method for finding the solution for this (r_k, p_A)-problem in EDS is to define an extended manifold $N = E \otimes \mathbb{R}^d \otimes \mathbb{R}^e$ with co-ordinates $\{x^a, r_k, p_A\},\ a \in \{1, ..., 2n\},\ k \in \{1, ..., d\},\ A \in \{1, ..., e\}$ and look for $2n+1$ dimensional manifolds that are sections over E and on which the one forms

$$\sigma_k := dr_k + r_h \xi_k^h + p_A \rho_k^A$$

are zero.

In the remainder of this section, we will give a brief outline of the process of finding the generality of the solutions to this last problem, see [7] or [10] for details.

To find these manifolds, σ_k are considered constraint forms for a distribution on N whose integral submanifolds we want. To find these integral manifolds, we choose a basis of forms on N, $\{\alpha_a, \sigma_k, \pi_A\}$ where $\{\alpha_a\}$ are a pulled back basis for E, $\pi_A = dp_A$, and σ_k as defined above completes the basis.

The condition that we want sections over E is that

$$\alpha_1 \wedge ... \wedge \alpha_{2n+1}$$

be non-zero on the $2n+1$ dimensional integral manifolds given by the constraint forms.

According to [7], to determine the existence and generality of the solutions to (68), we calculate the exterior derivatives $d\sigma_k$ modulo the ideal generated by the forms σ_k.

$$d\sigma_k \equiv \pi_k^1 \wedge \alpha_i + t_k^{ij} \alpha_i \wedge \alpha_j \mod \operatorname{span}\{\sigma_k\} \tag{69}$$

where π_k^i are some linear combination of dp_A. As $d\sigma_k$ expands with no $dp_A \wedge dp_B$ terms, the system is quasi-linear.

Because we want $\alpha_1 \wedge ... \wedge \alpha_{2n+1} \neq 0$ on the integral manifolds, we need to absorb all the $\alpha_i \wedge \alpha_j$ terms into the $\pi_k^i \wedge \alpha_i$ terms. This is done by changing the basis forms π_A to $\bar{\pi}_A = \pi_A - l_A^j \alpha_j$. If any of the $\alpha_i \wedge \alpha_j$ terms can not be absorbed, then asking for $d\sigma_k = 0$ mod span$\{\sigma^l\}$ is incompatible with the independence condition and therefore there is no solution.

Once the $\alpha_i \wedge \alpha_j$ terms have been removed, the next step is to create the tableau Π shown below from which the Cartan characters, $s_1, s_2, ..., s_k$, can be calculated allowing us to apply the Cartan test for involution.

$$\Pi = \begin{array}{c|cccc} & \alpha_1 & \alpha_2 & \dots & \alpha_n \\ \hline \sigma_1 & \pi_1^1 & \pi_1^2 & \dots & \pi_1^n \\ \sigma_2 & \pi_2^1 & \pi_2^2 & \dots & \pi_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ \sigma_d & \pi_d^1 & \pi_d^2 & \dots & \pi_d^n \end{array}$$

The basis $\{\alpha_i\}$ is chosen so that the number, s_1, of independent one-forms in column 1 of Π is maximum, the number, s_2, of independent one forms in column 2 also independent of those in column 1 is maximum with $s_2 \leq s_1$, and so on.

Once the Cartan characters are found, the Cartan test for involution is performed as follows:

Let t denote the number of ways in which the forms π^i_k can be modified by $\bar{\pi}_A = \pi_A - l^j_A \alpha_j$, without changing (69). The differential system (68) is in involution if

$$t = s_1 + 2s_2 + 3s_3 + ... + ks_k.$$

If the Cartan test fails, then it is necessary to prolong the differential system by differentiating the original equations to obtain a new differential system on J^1N, the first jet bundle of local sections of N over M, and then begin the process again. (See [10] for details, but to the best of our knowledge this failure has not been observed in the inverse problem.)

Once a series of Cartan characters is found that passes the Cartan test with last non-zero character s_l, then the general solution to the differential system will depend on s_l arbitrary functions of l variables.

6.4 EDS and the inverse problem: results

We give a brief outline of the results of Aldridge [1] and Aldridge *et al* [3]. Aldridge's starting point is theorem 6.3 which he uses to produce the hierarchies given in section 6.2. Working on $\mathbb{R} \times TM$ he produces further, apparently independent conditions, in contrast to the approach of [124]. Turning to the EDS results, Aldridge shows that each of the differential ideal steps entails a corresponding level in the two hierarchies. For example,

Theorem 6.5 *The first differential ideal step.*

Let $\Sigma^0 := span\{\Omega^k\} = span\{\psi^a \wedge \omega^b\}$*; if* Σ^1 *is the submodule of* Σ^0 *that satisfies the differential ideal condition:* $d\Omega \in \langle \Sigma^0 \rangle \; \forall \Omega \in \Sigma^1$*, then all* $\Omega \in \Sigma^1$ *satisfy the standard algebraic Helmholtz conditions:*

$$\Omega(X^V, Y^H) = \Omega(Y^V, X^H),$$
$$\Omega(A_\Gamma(X^H), Y^H) = \Omega(A_\Gamma(Y^H), X^H),$$
$$\sum_{(XYZ)} \Omega(R(X^H, Y^H), Z^H) = 0.$$

However, Aldridge's additional conditions come in to play at further differential ideal steps, and he falls short of identifying all the algebraic conditions produced by the differential ideal process. It is still not clear whether any further algebraic conditions arise from the later steps in the EDS process.

Using the EDS algorithm Aldridge alone and with his collaborators recover the results of [7, 124, 20] in a far more efficient manner than in the original work, often improving on the strength of the statements. For example, the statement that Σ^0 forms a differential ideal if and only if Φ is a multiple of the identity, is proved in just a few lines in [3], whereas in [7], only the reverse statement is shown. Aldridge gives a number of examples of the application of the EDS process to the cases of Douglas with a clear indication of the higher dimensional approach. In particular, he suggests that the classification be based first on the Jordan normal form of Φ and then on stage at which the differential ideal process terminates. This idea is pursued in [3].

There is a natural extension of this approach to the class of SODEs for which no Lagrangian exists. Suppose that the differential ideal process terminates with a submodule Σ^k whose basis does not permit a non-degenerate two form in Σ^k. For a given n such submodules are denumerable. For example, for $n = 2$ the final differential ideal could be generated by $\{\psi^1 \wedge \omega^1\}$, $\{\psi^2 \wedge \omega^2\}$, $\{\psi^1 \wedge \omega^2 + \psi^1 \wedge \omega^2\}$, and so on. The first two of these clearly contain no non-degenerate two forms and so cannot lead to variationality. Using the differential ideal condition on such differential ideals leads directly back to the classes of SODEs responsible. In this way we can list those classes of equations which are not variational and for which the EDS process terminates at the differential ideal step. At the moment we do not have non-existence examples which are not in this class.

In our view the adoption of EDS approach to the inverse problem represents a return to mainstream mathematics for this long-lived and productive area of enquiry.

6.5 The metrisability of connections

In the context of the inverse problem and differential geometry the following question is a natural one: suppose that we have a linear symmetric connection on M with autoparallel equations

$$\ddot{x}^a + \Gamma^a_{bc}\dot{x}^b\dot{x}^c = 0, \tag{70}$$

does there exist a multiplier $g_{ab}(x^c)$ which makes this system variational, and, of course, is g a metric for which Γ^a_{bc} is the Levi-Civita connection?

So let's assume a solution, g_{ab}, depending only on (x^a), to the Helmholtz conditions for (70). Notice that the last of equations (62) is automatically satisfied and we are assuming that g is symmetric.

Using the constructs of section 5.1 we find that

$$\Gamma^a_b = \Gamma^a_{bc}u^c, \quad \Phi^a_b = R^a_{dcb}u^c u^d$$

where R^a_{dcb} are the components of the Riemann curvature of the connection. The Helmholtz condition $g_{ac}\Phi^c_b = g_{bc}\Phi^c_a$ gives

$$g_{fa}R^a_{(cd)b} = g_{ba}R^a_{(cd)f}, \tag{71}$$

and the condition $\Gamma(g_{ab}) = g_{ac}\Gamma^c_b + g_{bc}\Gamma^c_a$, which along with the symmetry of Γ^a_{bc}, gives

$$\Gamma^a_{bc} = \frac{1}{2}g^{da}\left(\frac{\partial g_{cd}}{\partial x^b} + \frac{\partial g_{bd}}{\partial x^c} - \frac{\partial g_{bc}}{\partial x^d}\right).$$

So if g exists Γ^a_{bc} is its Levi-Civita connection (and (71) is just $R_{abcd} = R_{cdab}$). The existence question is in general trickier, but for $n = 2$ it's straightforward because of the relation between the Riemann and Ricci tensors:

$$R^a_{cdb} = \delta^a_b R_{cd} - \delta^a_d R_{cb},$$

(recall that $R_{ab} := R^c_{abc}$ and $R := R^a_a$). We find that if g exists then

$$R_{ab} = \frac{1}{2}Rg_{ab},$$

meaning that sufficient conditions for the existence of g are that $2\dfrac{R_{ab}}{R}$ should satisfy the Helmholtz conditions (62).

There is a corresponding question in Finsler geometry whose solution is not currently available.

6.6 The Grifone-Muzsnay approach

There is another approach to the inverse problem currently in use and due to Grifone and Muzsnay ([43, 44]). So far it only applies to autonomous equations and so it would take us too far out of our way to give a self-contained account. Such an account can, however, be found in section 3 of Muzsnay and Thompson's paper [104]. The basic idea is to study the multiplier and the obstructions to variationality through the Euler-Lagrange operator P_1. The integrability of this operator is examined using Spencer theory. Given an autonomous SODE and a function $L : TM \to \mathbb{R}$, this operator is defined as follows:
$P_1 : L \mapsto \omega_L \in \bigwedge^1(TM)$ with

$$\omega_L = i_\Gamma \Omega_L + d\mathcal{L}_\Delta L - dL$$

where, in this context, $\Omega_L := d(dL \circ S) = d(\frac{\partial L}{\partial \dot{x}^a} dx^a)$, $\Gamma := \dot{x}^a \frac{\partial}{\partial x^a} + f^a \frac{\partial}{\partial \dot{x}^a}$, $S := \frac{\partial}{\partial \dot{x}^a} \otimes dx^a$, and Δ is the Liouville field $\Delta := \dot{x}^a \frac{\partial}{\partial \dot{x}^a}$.
It is straightforward to show that

$$\omega_L = \left(\Gamma \left(\frac{\partial L}{\partial \dot{x}^a} \right) - \frac{\partial L}{\partial x^a} \right) dx^a,$$

so we must look for L which make $\omega_L \equiv 0$, and to this end the integrability of P_1 must be studied.

While this study has not solved any new classes of inverse problems in the sense of Douglas, there is an intriguing and as yet unresolved conflict with the along-the-projection approach to the integrability conditions, with the Spencer approach apparently producing many more independent conditions.

The Grifone-Muzsnay approach has been used the analysis of a certain inverse problem on Lie groups. In a number of recent papers, Thompson, Muzsnay and others [140, 104, 116] examine the variationality of the autoparallel flow of a certain canonical connection on a Lie group. This inverse problem provides a rich source of examples of both fixed and arbitrary dimension. The zero connection ∇ on a Lie group G is the canonical symmetric connection defined by $\nabla_X Y = \frac{1}{2}[X, Y]$, $\forall X, Y \in \mathfrak{g}$. Using both direct attacks on the Helmholtz conditions and the Spencer theory analysis of the Euler-Lagrange operator the authors produce a considerable variety of results. For example, ∇ is variational for every two-step nilpotent Lie group and for every Lie of dimension up to three. There are detailed studies of the affine and Euclidean groups of the plane. The position of these examples in any general classification of the inverse problem is not well understood.

6.7 Computational challenges and future impact

There are two challenges for symbolic computation inherent in the inverse problem. The first is that of solving the Helmholtz conditions for a specific SODE or class of SODEs including the consideration of global issues. The second is that of making progress with the Douglas type solution in fixed or arbitrary dimension.

The first of these problems is straightforward: the major symbolic computation tools have linear PDE solvers which can handle mixed algebraic and differential conditions like the Helmholtz conditions. The $n = 3$ case presents no significant difficulties for almost any type of smooth equations. For example, `dimsym` [132], the REDUCE differential equation symmetry package, can easily handle the spherically symmetric potential problem, reporting all representatives in this class of potentials and all obstructions to the globality of solutions (see also [2]). No special differential geometry needs to be implemented. Indeed, the ease with which concrete examples are currently handled is rather dispiriting to the theorist.

The second challenge is far more interesting. There is no reason why, given the current state of computer algebra implementations of EDS, that for a given n, a user should not be able to specify the Jordan normal form of Φ and various related information on the integrability of the eigenspaces and have the full EDS algorithm applied automatically. Further, there is a real possibility that the nonexistence classes generated by the differential ideal process (see section 6.4) can be elaborated for arbitrary n. Given the significant manual computational complexity of these tasks, symbolic computation is probably the only way that we will see an exhaustive solution of the inverse problem for $n = 3$.

The developments in the inverse problem described in these pages have yet to have a significant effect outside mathematics and physics. However, the depth of our theoretical understanding and our current computational tools are surely ready to have an impact in areas such as molecular quantum mechanics where a "designer Lagrangian" is the starting point for a molecular model. And in physics itself there is still the long standing open question concerning the inequivalent quantum mechanics arising from multiple Lagrangians for the classical equations of motion. The idea that the symmetries that come along with the multiple Lagrangians may provide a quotient space containing the only "right" Lagrangian is still unexplored. The solution to this problem will almost certainly have an impact on the study of the inverse problem itself.

Acknowledgements. Research supported by grants GAČR 201/06/0922 of the Czech Science Foundation, and MSM 6198959214 of the Ministry of Education, Youth and Sports of the Czech Republic.

Bibliography

[1] J. E. Aldridge: *Aspects of the inverse problem in the calculus of variations* (Ph.D. Thesis, La Trobe University, Australia, 2003)

[2] J. E. Aldridge and G. E. Prince: Computer algebra solution of the inverse problem in the calculus of variations *Comp. Phys. Comm.* **115** (1998) 489–509

[3] J. Aldridge, G. Prince, W. Sarlet and G. Thompson: The EDS approach to the inverse problem in the calculus of variations *J. Math. Phys.* **47** (2006) 103508 (22 pages)

[4] I. Anderson: Aspects of the inverse problem to the calculus of variations *Arch. Math. (Brno)* **24** (1988) 181–202

[5] I. Anderson: *The variational bicomplex* (Utah State University, Technical Report, 1989)

[6] I. Anderson and T. Duchamp: On the existence of global variational principles *Am. J. Math.* **102** (1980) 781–867

[7] I. Anderson and G. Thompson: *The inverse problem of the calculus of variations for ordinary differential equations* (Memoirs Amer. Math. Soc. **98** No. 473, 1992)

[8] S. Benenti: L'intégration de l'equation d'Hamilton–Jacobi par séparation des variables: histoire et résultats récents *In: La Mécanique Analytique de Lagrange et son héritage* (Suppl. al. No 124 (1990) Atti Accad. Sci. Torino, Torino, 1990) 119–144

[9] D. Bao, S.-S. Chern, Z. Shen: *An introduction to Riemann–Finsler geometry* (Berlin: Springer, 2000)

[10] R. L. Bryant, S.-S. Chern, R. B. Gardner, H. L. Goldschmidt and P. A. Griffiths: *Exterior differential systems* (Berlin: Springer, 1991)

[11] G. B. Byrnes: A complete set of Bianchi identities for tensor fields along the tangent bundle projection *J. Phys. A: Math. Gen.* **27** (1993) 6617–6632

[12] G. B. Byrnes: A linear connection for higher-order ordinary differential equations *J. Phys. A: Math. Gen.* **29** (1996) 1685–1694

[13] É. Cartan: *Leçons sur les invariants intégraux* (Paris: Hermann, 1922)

[14] M. Crampin: On the differential geometry of the Euler-Lagrange equations, and the inverse problem of Lagrangian dynamics *J. Phys. A: Math. Gen.* **14** (1981) 2567–2575

[15] M. Crampin: Tangent bundle geometry for Lagrangian dynamics *J. Phys. A: Math. Gen.* **16** (1983) 3755–3772

[16] M. Crampin, E. Martínez and W. Sarlet: Linear connections for systems of second–order ordinary differential equations *Ann. Inst. H. Poincaré Phys. Théor.* **65** (1996) 223–249.

[17] M. Crampin and G. E. Prince: Generalizing gauge freedom for spherically symmetric potentials *J. Phys. A: Math. Gen.* **18** (1985) 2167–2175

[18] M. Crampin and G. E. Prince: Equivalent Lagrangians and dynamical symmetries: some comments *Phys. Lett.* **108A** (1985) 191–194

[19] M. Crampin and G. E. Prince: Alternative Lagrangians for spherically symmetric potentials *J. Math. Phys.* **29** (1988) 1551–1555

[20] M. Crampin, G. E. Prince, W. Sarlet and G. Thompson: The inverse problem of the calculus of variations: separable systems *Acta Appl. Math.* **57** (1999) 239–254

[21] M. Crampin, G. E. Prince and G. Thompson: A geometric version of the Helmholtz conditions in time dependent Lagrangian dynamics *J. Phys. A: Math. Gen.* **17** (1984) 1437–1447

[22] M. Crampin, W. Sarlet and F. Cantrijn: Higher order differential equations and higher order Lagrangian mechanics *Math. Proc. Camb. Phil. Soc.* **99** (1986) 565–587

[23] M. Crampin, W. Sarlet, E. Martínez, G. B. Byrnes and G. E. Prince: Toward a geometrical understanding of Douglas's solution of the inverse problem in the calculus of variations *Inverse Problems* **10** (1994) 245–260

[24] P. Dedecker and W. M. Tulczyjew: Spectral sequences and the inverse problem of the calculus of variations, Proc. Internat. Coll. on Diff. Geom. Methods in Math. Phys., Salamanca 1979 *In:* Lecture Notes in Math. **836** (Berlin: Springer, 1980) 498–503

[25] M. de León and D. M. de Diego: Classification of symmetries for higher order Lagrangian systems II: the non-autonomous case *Extracta Mathematicae* **8** (1994) 111–114

[26] M. de León and D. M. de Diego: Symmetries and constants of the motion for higher order Lagrangian systems *J. Math. Phys.* **36** (1995) 4138–4161

[27] M. de León, J. Marín-Solano and J. C. Marrero: The constraint algorithm in the jet formalism *Diff. Geom. Appl.* **6** (1996) 275–300

[28] M. de León and P. Rodrigues: *Methods of differential geometry in analytical mechanics.* (North-Holland Mathematics Studies, 158. Amsterdam-New York: North-Holland, 1989)

[29] P. A. M. Dirac: Generalised Hamiltonian dynamics *Canad. J. Math.* **II** (1950) 129–148

[30] J. Douglas: Solution of the inverse problem of the calculus of variations *Trans. Am. Math. Soc.* **50** (1941) 71–128

[31] Ch. Ehresmann: Les connections infinitésimales dans un espace fibré différentiable *In: Coll. Topologie, Bruxelles 1950* (Liége, 1951) 29–55

[32] Ch. Ehresmann: Les prolongements d'une variété différentiable: 1. Calcul des Jets, prolongement principal *C. R. Acad. Sci. Paris* **233** (1951) 598–600

[33] Ch. Ehresmann: Les prolongements d'une space fibré différentiable *C. R. Acad. Sci. Paris* **240** (1955) 1755–1757

[34] E. Engels and W. Sarlet: General solution and invariants for a class of Lagrangian equations governed by a velocity-dependent potential energy *J. Phys. A: Math. Gen.* **6** (1973) 818–825

[35] R. Ghanam, G. Thompson and E. J. Miller: Variationality of four-dimensional Lie group connections *J. Lie Theory* **14** (2004) 395–425

[36] H. Gluck: *Global theory of dynamical systems, Lecture Notes in Mathematics 819* (Berlin: Springer, 1980)

[37] H. Goldschmidt and S. Sternberg: The Hamilton-Cartan formalism in the calculus of variations *Ann. Inst. Fourier* **23** (1973) 203–267

[38] M. J. Gotay and J. M. Nester: Presymplectic Lagrangian systems I: the constraint algorithm and the equivalence theorem *Ann. Inst. H. Poincaré, Sect. A* **30** (1979) 129–142

[39] M. J. Gotay and J. M. Nester: Presymplectic Lagrangian systems, II: the second order equation problem *Ann. Inst. H. Poincaré, Sect. A* **32** (1980) 1–13

[40] M. J. Gotay, J. M. Nester and G. Hinds: Presymplectic manifolds and the Dirac–Bergmann theory of constraints *J. Math. Phys.* **19** (1978) 2388–2399

[41] J. Grifone: Structure presque tangente et connexions I *Ann. Inst. Fourier* **22** (1), (1972) 287–334

[42] J. Grifone: Structure presque tangente et connexions II *Ann. Inst. Fourier* **22** (3), (1972) 291–338

[43] J. Grifone and Z. Muzsnay: On the inverse problem of the variational calculus: existence of Lagrangians associated with a spray in the isotropic case *Ann. Inst. Fourier* **49** (1999) 1387–1421

[44] J. Grifone and Z. Muzsnay: *Variational principles for second-order differential equations* (World Scientific, 2000)

[45] D. R. Grigore: Lagrangian formalism on Grassmann manifolds *In: Handbook of Global Analysis* (Amsterdam: Elsevier, 2007) 327–373

[46] A. Haková and O. Krupková: Variational first-order partial differential equations *J. Differential Equations* **191** (2003) 67–89

[47] H. Helmholtz: Uber der physikalische Bedeutung des Princips der kleinsten Wirkung *J. Reine Angew. Math.* **100** (1887) 137–166

[48] M. Henneaux: On the inverse problem of the calculus of variations *J. Phys. A: Math. Gen.* **15** (1982) L93–L96

[49] M. Henneaux and L. C. Shepley: Lagrangians for spherically symmetric potentials *J. Math. Phys.* **23** (1988) 2101–2107

[50] A. Hirsch: Die Existenzbedingungen des verallgemeinerten kinetischen Potentials *Math. Ann.* **50** (1898) 429–441

[51] L. A. Ibort and C. López–Lacasta: *J. Phys. A: Math. Gen.* **23** (1990) 4779–4792

[52] M. Jerie and G. E. Prince: A generalised Raychaudhuri equation for second-order differential equations *J. Geom. Phys.* **34** (2000) 226–241

[53] M. Jerie and G. E. Prince: Jacobi fields and linear connections for arbitrary second order ODE's *J. Geom. Phys.* **43** (2002) 351–370

[54] L. Klapka: Integrals of motion and semiregular Lepagean forms in higher order mechanics *J. Phys. A: Math. Gen.* **16** (1983) 3783–3794

[55] L. Klapka: Euler-Lagrange expressions and closed two-forms in higher order mechanics *In: Geometrical Methods in Physics, Proc. Conf. on Diff. Geom. and Appl. Vol. 2, Nové Město na Moravě, Sept. 1983, D. Krupka, Ed.* (J. E. Purkyně Univ. Brno, Czechoslovakia, 1984) 149–153

[56] J. Klein: Espaces variationels et mécanique *Ann. Inst. Fourier* **12** (1962) 1–124

[57] J. Klein: Structures symplectique ou J-symplectiques homogènes sur l'espace tangent à une variété *Symp. Math.* **14** (1974) 181–192

[58] D. Krupka: *Some geometric aspects of variational problems in fibered manifolds* (Folia Fac. Sci. Nat. Univ. Purk. Brunensis, Physica **14**, Brno, Czechoslovakia, 1973); 65pp; `arXiv:math-ph/0110005`

[59] D. Krupka: A geometric theory of ordinary first order variational problems in fibered manifolds. I. Critical sections *J. Math. Anal. Appl.* **49** (1975) 180–206

[60] D. Krupka: A geometric theory of ordinary first order variational problems in fibered manifolds. II. Invariance *J. Math. Anal. Appl.* **49** (1975) 469–476

[61] D. Krupka: On the local structure of the Euler-Lagrange mapping of the calculus of variations *In: Proc. Conf. on Diff. Geom. Appl., Nové Město na Moravě (Czechoslovakia), 1980,* (Charles Univ. Prague, 1982, 181–188) ; `arXiv:math-ph/0203034`

[62] D. Krupka: Lepagean forms in higher order variational theory *In: Modern Developmens in Analytical Mechanics I: Geometrical Dynamics, Proc. IUTAM-ISIMM Symposium, Torino, Italy, 1982,* (Torino: Accad. Sci. Torino, 1983) 197–238

[63] D. Krupka: Variational sequences on finite order jet spaces *In: Proc. Conf., Brno (Czechoslovakia) 1989, J. Janyška and D. Krupka, Eds.* (Singapore: World Scientific, 1990) 236–254

[64] D. Krupka: Variational sequences in mechanics *Calc. Var.* **5** (1997) 557–583

[65] D. Krupka: Global variational principles: Foundations and current problems *In: Global Analysis and Applied Mathematics, AIP Conference Proceedings 729* (American Institute of Physics, 2004) 3–18

[66] D. Krupka: Global variational theory in fibred spaces *In: Handbook of Global Analysis* (Amsterdam: Elsevier, 2007) 773–836

[67] D. Krupka, O. Krupková, G. Prince, and W. Sarlet: Contact symmetries and variational sequences *In: Proc. 9th Internat. Conf. on Diff. Geom. Appl., Prague, 2004* (Charles University, Prague, 2005) 599-609

[68] D. Krupka, O. Krupková, G. E. Prince and W. Sarlet: Contact symmetries of the Helmholtz form *Diff. Geom. Appl.* (to appear)

[69] D. Krupka and A. Sattarov: The inverse problem of the calculus of variations for Finsler structures *Math. Slovaca* **35** (1985) 217–222

[70] D. Krupka and J. Šeděnková: Variational sequences and Lepagean forms *In: Differential Geometry and its Applications, Proc. Conf., Prague, 2004, J. Bureš, O. Kowalski, D. Krupka, J. Slovák, Eds.* (Charles Univ., Prague, Czech Republic, 2005) 605–615

[71] O. Krupková: Lepagean 2-forms in higher order Hamiltonian mechanics, I. Regularity *Arch. Math. (Brno)* **22** (1986) 97–120

[72] O. Krupková: Lepagean 2-forms in higher order Hamiltonian mechanics, II. Inverse problem *Arch. Math. (Brno)* **23** (1987) 155–170

[73] O. Krupková: A note on the Helmholtz conditions *In: Differential Geometry and Its Applications, Proc. Conf., Brno, August 1986, Communications, D. Krupka and A. Švec, Eds.* (J.E. Purkyně University, Brno, Czechoslovakia, 1987) 181–188

[74] O. Krupková: *Variational analysis on fibered manifolds over one-dimensional bases* (PhD Thesis, Silesian University, Opava et Charles University, Prague, 1992); 67 pp

[75] O. Krupková: Liouville and Jacobi theorems for vector distributions *In: Proc. of the 5th International Conference on Differential Geometry and Its Applications, August 1992, Opava, Czechoslovakia, O. Kowalski and D. Krupka, Eds.* (Silesian University, Opava, 1993) 75–87

[76] O. Krupková: Variational metrics on $R \times TM$ and the geometry of nonconservative mechanics *Math. Slovaca* **44** (1994) 315–335

[77] O. Krupková: A geometric setting for higher-order Dirac-Bergmann theory of constraints *J. Math. Phys.* **35** (1994) 6557–6576

[78] O. Krupková: Higher-order constrained systems on fibered manifolds: An exterior differential systems approach *In: New Developments in Differential Geometry, Proc. Colloq. on Diff. Geom., Debrecen, 1994, L. Tamássy and J. Szenthe, Eds.* (Dordrecht: Kluwer, 1996) 255–278

[79] O. Krupková: A new look at Dirac's theory of constrained systems *In: Gravity, Particles and Space-Time, P. Pronin and G. Sardanashvily, Eds.* (Singapore: World Scientific, 1996) 507–517

[80] O. Krupková: Symmetries and first integrals of time-dependent higher-order constrained systems *J. Geom. Phys.* **18** (1996) 38–58

[81] O. Krupková: Noether theorem and first integrals of constrained Lagrangean systems *Math. Bohemica* **122** (1997) 257–265

[82] O. Krupková: *The geometry of ordinary variational equations, Lecture Notes in Mathematics 1678* (Berlin: Springer, 1997)

[83] O. Krupková: Mechanical systems with non-holonomic constraints *J. Math. Phys.* **38** (1997) 5098–5126

[84] O. Krupková: Differential systems in higher-order mechanics *In: Proceedings of the Seminar on Differential Geometry, D. Krupka, Ed.* (Mathematical Publications 2, Silesian University, Opava, 2000) 87–130

[85] O. Krupková: Hamiltonian field theory revisited: A geometric approach to regularity *In: Steps in Differential Geometry, Proc. Colloq. Diff. Geom., Debrecen, July 2000, L. Kozma, P.T. Nagy and L. Tamássy, Eds.* (Debrecen: Debrecen University, 2001) 187–207

[86] O. Krupková: Variational metric structures *Publ. Math. Debrecen* **62** (2003) 461–495; paper in honour of L. Tamássy

[87] O. Krupková: The geometry of variational equations *In: Global Analysis and Applied Mathematics, AIP Conference Proceedings 729* (American Institute of Physics, 2004) 19–38

[88] O. Krupková and A. Vondra: On some integration methods for connections on fibered manifolds *In: Proc. of the 5th International Conference on Differential Geometry and Its Applications, 1992, Opava, Czechoslovakia, O. Kowalski and D. Krupka, Eds.* (Silesian University, Opava, 1993) 89–101

[89] Th. Lepage: Sur les champs géodésiques du Calcul des Variations *Bull. Acad. Roy. Belg., Cl. des Sciences* **22** (1936) 716–729

[90] P. Libermann and C.-M. Marle: *Symplectic geometry and analytical mechanics* (Dordrecht: Reidel, 1987)

[91] J. Liouville: *Journal de Math.* **20** (1855) 137

[92] V. Lychagin: *Lectures on geometry of differential equations, Part I, Part II* (Roma: Consiglio Nazionale delle Ricerche (G.N.F.M.), 1992)

[93] L. Mangiarotti and M. Modugno: Fibered spaces, jet spaces and connections for field theories *In: Proc. of the Meeting "Geometry and Physics", 1982, Florence* (Bologna: Pitagora, 1982) 135–165

[94] G. Marmo, G. Mendella and W. M. Tulczyjew: Symmetries and constants of the motion for dynamics in implicit form *Ann. Inst. Henri Poincaré, Phys. Theor.* **57** (1992) 147–166

[95] E. Martínez, J. F. Cariñena and W. Sarlet: Derivations of differential forms along the tangent bundle projection *Diff. Geom. and Applic.* **2** (1992) 17–43

[96] E. Martínez, J. F. Cariñena and W. Sarlet: Derivations of differential forms along the tangent bundle projection II *Diff. Geom. and Applic.* **3** (1993) 1–29

[97] E. Martínez, J. F. Cariñena and W. Sarlet: A geometric characterization of separable second-order differential equations *Math. Proc. Camb. Phil. Soc.* **113** (1993) 205–224

[98] E. Massa and E. Pagani: Jet bundle geometry, dynamical connections, and the inverse problem of Lagrangian mechanics *Ann. Inst. Henri Poincaré, Phys. Theor.* **61** (1994) 17–62

[99] A. Mayer: Die Existenzbedingungen eines kinetischen Potentiales *Berich. Verh. Konig. Sach. Gesell. wissen. Leipzig, Math. Phys. Kl.* **84** (1896) 519–529

[100] T. Mestdag: *Berwald-type connections in time-dependent mechanics and dynamics on affine Lie algebroids.* (PhD thesis, Universiteit Gent, Belgium, 2003)

[101] T. Mestdag and W. Sarlet: The Berwald-type connection associated to time-dependent second-order differential equations *Houston J. Math.* **27** (2001) 763–797

[102] M. Modugno: Torsion and Ricci tensor for non-linear connections *Diff. Geom. Appl.* **1** (1991) 177–192

[103] G. Morandi, C. Ferrario, G. Lo Vecchio, G. Marmo and C. Rubano: The inverse problem in the calculus of variations and the geometry of the tangent bundle *Phys. Rep.* **188** (1990) 147–284

[104] Z. Muzsnay and G. Thompson: Inverse problem of the calculus of variations on Lie groups *Differential Geom. Appl.* **23** (2005) 257–281

[105] E. Noether: Invariante Variationsprobleme *Nachr. kgl. Ges. Wiss. Göttingen, Math. Phys. Kl.* (1918) 235–257

[106] J. Novotný: On the inverse variational problem in the classical mechanics *In: Proc. Conf. on Diff. Geom. and Its Appl., Nové Město na Moravě, 1980, O. Kowalski, Ed.* (Universita Karlova, Prague, 1981) 189–195

[107] P. J. Olver: *Applications of Lie Groups to Differential Equations,* 2^{nd} *Edition* (Berlin: Springer, 1993)

[108] G. E. Prince: Toward a classification of dynamical symmetries in classical mechanics *Bull. Aust. Math. Soc.* **27** (1983) 53–71

[109] G. E. Prince: A complete classification of dynamical symmetries in classical mechanics *Bull. Austral. Math. Soc.* **32** (1985) 299–308

[110] G. E. Prince: The inverse problem in the calculus of variations and its ramifications *In: Geometric Approaches to Differential Equations, ed. P. Vassiliou* (Lecture Notes of the Australian Mathematical Society, CUP, 2001)

[111] G. E. Prince, J. Aldridge and G. B. Byrnes: A universal Hamilton-Jacobi equation for second-order ODEs *J. Phys. A: Math. Gen.* **32** (1999) 827–844

[112] G. E. Prince, J. Aldridge, S. E. Godfrey and G. B. Byrnes: The separation of the Hamilton-Jacobi equation for the Kerr metric *J. Aust. Math. Soc. Series B* **41** (1999) 248–259

[113] G. E. Prince, G. B. Byrnes, J. Sherring and S. E. Godfrey: A generalization of the Liouville-Arnol'd theorem *Math. Proc. Camb. Phil. Soc.* **117** (1995) 353–370

[114] G. E. Prince and M. Crampin: Projective differential geometry and geodesic conservation laws in general relativity, I: projective actions *Gen. Rel. Grav.* **16** (1984) 921–942

[115] G. E. Prince and M. Crampin: Projective differential geometry and geodesic conservation laws in general relativity, II: conservation laws *Gen. Rel. Grav.* **16** (1984) 1063–1075

[116] M. Rawashdeh and G. Thompson: The inverse problem for six-dimensional codimension two nilradical Lie algebras *J. math. Phys.* **47** (2006) 112901, 29*pp*

[117] W. Sarlet: Symmetries, first integrals and the inverse problem of Lagrangian mechanics I *J. Phys. A: Math. Gen.* **14** (1981) 2227–2238

[118] W. Sarlet: The Helmholtz conditions revisited. A new approach to the inverse problem of Lagrangian dynamics *J. Phys. A: Math. Gen.* **15** (1982) 1503–1517

[119] W. Sarlet: Symmetries, first integrals and the inverse problem of Lagrangian mechanics II *J. Phys. A: Math. Gen.* **16** (1983) 1383–1396

[120] W. Sarlet: Noether's Theorem and the inverse problem of Lagrangian mechanics *In: Modern Developments in Analytical Mechanics, Proc. IUTAM-ISIMM Symposium, Torino, Italy 1982, S. Benenti, M. Francaviglia and A. Lichnerowicz, Eds.* (Torino: Accad. delle Scienze di Torino, 1983) 737–751

[121] W. Sarlet: Geometrical structures related to second-order equations *In: Differential Geometry and its Applications, Proc. Conf. 1986 Brno, Czechoslovakia, D. Krupka and A. Švec, eds.* (J. E. Purkyně University, Brno, 1987) 279–288

[122] W. Sarlet and F. Cantrijn: Generalizations of Noether's theorem in classical mechanics *SIAM Review* **23** (1981) 467–494

[123] W. Sarlet and M. Crampin: Addendum to:"The integrability conditions in the inverse problem of the calculus of variations for second-order ordinary differential equations" [Acta. Appl. Math. 54 (1998), no. 3, 233–273; by Sarlet, Crampin and E. Martínez *Acta Appl. Math.* **60** (2000) 213–224

[124] W. Sarlet, M. Crampin and E. Martínez: The integrability conditions in the inverse problem of the calculus of variations for second-order ordinary differential equations *Acta Appl. Math.* **54** (1998) 233–273

[125] W. Sarlet and T. Mestdag: Aspects of time-dependent second-order differential equations: Berwald-type connections *In: Steps in differential geometry (Debrecen, 2000)* (Debrecen: Inst. Math. Inform., 2001) 283–293

[126] W. Sarlet, G. E. Prince and M. Crampin: Adjoint symmetries for time dependent second order equations *J. Phys. A: Math. Gen.* **23** (1990) 1335–1347

[127] W. Sarlet, G. Thompson and G. E. Prince: The inverse problem of the calculus of variations: the use of geometrical calculus in Douglas's analysis *Trans. Amer. Math. Soc.* **354** (2002) 2897–2919

[128] W. Sarlet, A. Vandecasteele, F. Cantrijn and E. Martínez: Derivations of forms along a map: the framework for time–dependent second–order equations *Diff. Geom. Applic.* **5** (1995) 171–203

[129] D. J. Saunders: Jet fields, connections and second-order differential equations *J. Phys. A: Math. Gen.* **20** (1987) 3261–3270

[130] D. J. Saunders: *The Geometry of Jet Bundles* (Cambridge University Press, 1989)

[131] J. Sherring and G. E. Prince: Geometric aspects of reduction of order *Trans. Am. Math. Soc.* **334** (1992) 433–453

[132] J. Sherring, A. K. Head and G. E. Prince: DIMSYM and LIE: symmetry determination packages *Mathl. Comput. Modelling* **25** (1997) 153–164

[133] N. Ya. Sonin: On the definition of maximal and minimal properties *Warsaw Univ. Izvestiya* **1–2** (1886) 1–68 (in Russian)

[134] M. Štefanová and O. Štěpánková: Variationality and invariance of systems of partial differential equations *Scripta Fac. Sci. Nat. UJEP Brunensis (Physica)* **5** (1985) 283–288

[135] O. Štěpánková: *The inverse problem of the calculus of variations in mechanics* (Thesis, Charles University, Prague, 1984); 63 pp. (in Czech)

[136] S. Sternberg: *Differential geometry,* 2^{nd} *edition* (New York: Chelsea, 1983)

[137] H. Sussmann: Orbits of families of vector fields and integrability of distributions *Trans. Amer. Math. Soc.* **180** (1973) 171–188

[138] F. Takens: A global version of the inverse problem of the calculus of variations *J. Diff. Geom.* **14** (1979) 543–562

[139] G. Thompson: Private communication (2000)

[140] G. Thompson: Variational connections on Lie groups *Differential Geom. Appl.* **18** (2003) 255–270

[141] E. Tonti: Variational formulation of nonlinear differential equations I, II *Bull. Acad. Roy. Belg. Cl. Sci.* **55** (1969) 137–165, 262–278

[142] A. Trautman: Noether equations and conservation laws *Comm. Math. Phys.* **6** (1967) 248–261

[143] A. Trautman: Invariance of Lagrangian systems *In: General Relativity, Papers in honour of J. L. Synge* (Oxford: Clarendon Press, 1972) 85–99

[144] W. M. Tulczyjew: The Euler-Lagrange resolution, Internat. Coll. on Diff. Geom. Methods in Math. Phys., Aix-en-Provence 1979 *In: Lecture Notes in Math. 836* (Berlin: Springer, 1980) 22–48

[145] M. M. Vainberg: *Variational methods in the theory of nonlinear operators* (Moscow: GITL, 1959); (in Russian)

[146] A. L. Vanderbauwhede: Potential operators and variational principles *Hadronic J.* **2** (1979) 620–641

[147] V. P. Viflyantsev: Frobenius Theorem for distributions of non-constant rank *Uspekhi Mat. Nauk* **197** (1977) 177–178 (in Russian)

[148] V. P. Viflyantsev: On completely integrable distributions of non-constant rank *Mat. Zametki* **26** (1979) 921–929 (in Russian)

[149] A. M. Vinogradov: A spectral sequence associated with a non-linear differential equation, and algebro-geometric foundations of Lagrangian field theory with constraints *Soviet Math. Dokl.* **19** (1978) 144–148

[150] A. M. Vinogradov: The $\mathcal{C}$-spectral sequence, Lagrangian formalism, and conservation laws. I. The linear theory *J. Math. Anal. Appl.* **100** (1984) 1–40

[151] A. M. Vinogradov: The $\mathcal{C}$-spectral sequence, Lagrangian formalism, and conservation laws. II. The nonlinear theory *J. Math. Anal. Appl.* **100** (1984) 41–129

[152] R. Vitolo: Variational sequences *In: Handbook of Global Analysis* (Amsterdam: Elsevier, 2007) 1115–1163

[153] A. Vondra: Semisprays, connections and regular equations in higher order mechanics *In: Differential Geometry and Its Applications, Proc. Conf., Brno, Czechoslovakia, 1989, J. Janyška and D. Krupka Eds.* (Singapore: World Scientific, 1990) 276–287

[154] A. Vondra: *Connections in the geometry of non-autonomous regular higher-order dynamics* (PhD Thesis, Masaryk University, Brno, 1991); 114 pp.

O. KRUPKOVÁ

Department of Algebra and Geometry, Palacký University,
Tomkova 40, 779 00 Olomouc, Czech Republic,
and
Department of Mathematics, LaTrobe University, Victoria 3086, Australia
E-mail: `krupkova@inf.upol.cz`

G. E. PRINCE

Department of Mathematics, LaTrobe University, Victoria 3086, Australia
E-mail: `G.Prince@latrobe.edu.au`

Handbook of Global Analysis
Demeter Krupka and David Saunders

Elements of noncommutative geometry

Giovanni Landi

Contents

1 Introduction

Starting from the early eighties, there has been an increasing interest in noncommutative (and quantum) geometry both in mathematics and in mathematical physics, with motivations going back to quantum mechanics where classical observables such as position and momenta do not commute any longer. The aims are to carry geometrical concepts over to a new class of spaces whose algebras of functions are not commutative in general and to use them in a variety of applications. In particular, it has emerged that such noncommutative spaces retain a rich topology and geometry expressed first of all in K-theory and K-homology, and in a variety of finer aspects of the theory. Developments have occurred in several different fields of both pure mathematics and mathematical physics. In mathematics these include fruitful interactions with analysis, number theory, category theory and representation theory. In mathematical physics, noncommutative geometry has been used for the quantum Hall effect, for applications to the standard model in particle physics and to renormalization in quantum field theory, to models of spacetimes with noncommuting coordinates, to noncommutative gauge theories and string theory.

By now the know territory is so vast and new regions are discovered at such a high speed that the number of relevant papers is overwhelming. It is impossible to even think of covering 'everything'. In this report we attempt to a friendly introduction to some aspects of noncommutative geometry and of its applications and we confine ourself mainly to noncommutative differential geometry as originated from the work of A. Connes.The commutative Gel'fand-Naimark theorem states that thereis a complete equivalence between the category of (locally) compact Hausdorff spaces and(proper and) continuous maps and the category of commutative (not necessarily) unitalC^*-algebras and *-homomorphisms. Any commutativeC^*-algebra can be realized as the C^*-algebra of complex valued functions on a(locally) compact Hausdorff space. A noncommutativeC^*-algebra will then be thought of as the algebra of continuous functions on some'virtual noncommutative space'. The attention will be switched from spaces, which ingeneral do not even exist 'concretely', to algebras of functions 'defined on them'. Moregeneral suitable *-algebras will play the role of smooth functions.Noncommutative spin geometry is based on the notion of a *spectral triple* $(\mathcal{A}, \mathcal{H}, D)$ [16, 17, 18]. Here $\mathcal{A}$ is a noncommutative *-algebra, $\mathcal{H}$ is a Hilbert space on which $\mathcal{A}$ is realized faithfully as an algebra of bounded operators, and D is an operator on $\mathcal{H}$ withsuitable properties and which contains (almost all) the 'geometric' information. In fact, there is also a *real structure* given via an antilinear isometry J on $\mathcal{H}$ and suitable compatibility conditions. With any closed n-dimensional Riemannian spin manifold M there is associated acanonical spectral triple with $\mathcal{A} = C^\infty(M)$, the algebra of complex valuedsmooth functions on M; $\mathcal{H} = L^2(M, S)$, the Hilbert space of square integrablesections of the irreducible spinor bundle over M;and D the Dirac operator associated with the Levi-Civita connection of the metric (and J related to the charge conjugation operator). Interesting examples of noncommutative manifolds are provided by the noncommutative torus [12, 60] and the toric noncommutative manifolds of [21].

The recent constructions of spectral triples, with the consequent analysis of the corresponding spectral geometry, for the manifold of the quantum group $SU_q(2)$ in [29, 30], for its quantum homogeneous spaces (the spheres of Podleś [57]) in [28, 27], and for a quantum Euclidean four-sphere in [26], have provided a number of interesting examples at the frontiers between noncommutative geometry and quantum groups theory. A common feature of these examples is that their geometry is isospectral to the undeformed one and the dimension (and the dimension spectrum) is the same as in the commutative limit. Furthermore, they show that in order to have a real spectral triple one is forced to weaken the original requirements that the real structure should satisfy, a phenomenon first observed in [28] for the equatorial Podleś sphere.

Additional examples of (not necessarily isospectral) noncommutative geometry on quantum spaces have been constructed on quantum two spheres [31, 25, 56, 66] on the quantum group $SU_q(2)$ [10, 19] as well as on quantum flag manifolds [46].

Yang-Mills and gravity theories stem from the notion of connection (gauge orlinear) on vector bundles. The possibility of extending these notions to therealm of noncommutative geometry relies on another classical duality. The Serre-Swan theorem [69, 16] states that there is a complete equivalence between the category of (smooth) vector bundles over a (smooth) compact manifold and bundle maps and the category of projective modules of finite type over commutative algebras and module morphisms. The space $\Gamma(E, M)$ of (smooth) sections of a vector bundle E over a compact manifold M is a projective module of finite type over the algebra $C(M)$ of (smooth) functions over M and any finite projective $C(M)$-module is realized as the module of sections of some bundle over M. With a noncommutative algebra $\mathcal{A}$ as the starting ingredient the (analogueof) vector bundles will be projective modules of finite type over $\mathcal{A}$. One thendevelops a full theory of connections which culminates in the definition of a Yang-Mills action. Needless to say, starting with the canonical triple associated with an ordinary manifold one recovers usual gauge theories.

Noncommutative Yang-Mills theory onnoncommutative tori have been constructed in [24]. On the toric noncommutative spheres of [21] similar gauge theories have been recently constructed in [48, 49] with a crucial use of twisted symmetries.

Useful introduction to several aspects on noncommutative geometry are available in [47, 53, 39, 54, 43, 72]. A recent bird view of the field which shows applications ranging from number theory to physics is [22].

2 Algebras instead of spaces

The well known classical duality between ordinary spaces and (suitable) commutative algebras is expressed by the Gel'fand-Naimark theorem: the algebra of functions on a Hausdorff topological space is the only possible kind of commutative C^*-algebra.

Example 2.1 Let $C_0(M)$ be the C^*-algebra of complex valued continuous functions on a locally compact Hausdorff topological space M which vanish at infinity. The (commutative) product is pointwise multiplication, the * operation is just complex conjugation and the norm is the supremum norm,

$$\|f\|_\infty = \sup_{x\in M} |f(x)| \, .$$

This algebra has no unit. For a compact M the algebra $C_0(M)$ has a unit (the constant function $f = 1$) and it coincides with the algebra $C(M)$ of all continuous functions on M. One can prove that $C_0(M)$ (and a fortiori $C(M)$ if M is compact) is complete in the supremum norm. Indeed, it is the closure in the above norm of the algebra of functions with compact support [65].

Given *any* commutative C^*-algebra $\mathcal{C}$ one can reconstruct a Hausdorff topological space M such that $\mathcal{C}$ is isometrically $*$-isomorphic to the algebra of (complex valued) continuous functions $C_0(M)$ [32, 34]. The space M is the space of characters of $\mathcal{C}$. A *character* of the commutative C^*-algebra $\mathcal{C}$ is a one dimensional irreducible representation, that is, a (non-zero) *-linear functional $x : \mathcal{C} \to \mathbb{C}$ which is multiplicative, i.e. $x(fg) = x(f)x(g)$ for any $f, g \in \mathcal{C}$. The space $\widehat{\mathcal{C}}$ of all characters is called the *structure space* (or *Gel'fand spectrum*) of $\mathcal{C}$. It is made into a topological space, called the *Gel'fand space* of $\mathcal{C}$, by endowing it with the *Gel'fand topology*, i.e. the topology of pointwise convergence on $\mathcal{C}$. In such a topology, a sequence $\{x_n\}$ of elements of $\widehat{\mathcal{C}}$ converges to $x \in \widehat{\mathcal{C}}$ if and only if for any $g \in \mathcal{C}$, the sequence $\{x_n(g)\}$ converges to $x(g)$ in the topology of $\mathbb{C}$. It turns out that $\widehat{\mathcal{C}}$ is a compact Hausdorff space if the algebra $\mathcal{C}$ has a unit, otherwise it is only locally compact. If $f \in \mathcal{C}$, its *Gel'fand transform* $\hat{f} : \widehat{\mathcal{C}} \to \mathbb{C}$ is the function on $\widehat{\mathcal{C}}$ given by

$$\hat{f}(x) = x(f) \, , \quad \forall \, x \in \widehat{\mathcal{C}} \, ,$$

which is clearly continuous. We thus get the interpretation of elements in $\mathcal{C}$ as $\mathbb{C}$-valued continuous functions on $\widehat{\mathcal{C}}$. The Gel'fand-Naimark theorem states that all continuous functions on $\widehat{\mathcal{C}}$ are of the form above for some $f \in \mathcal{C}$. In fact, the Gel'fand transform $f \to \hat{f}$ is an isometric $*$-isomorphism of $\mathcal{C}$ onto $\mathcal{C}(\widehat{\mathcal{C}})$; isometric meaning that $\|\hat{f}\|_\infty = \|f\|$, for any $f \in \mathcal{C}$, with $\| \cdot \|_\infty$ the supremum norm on $\mathcal{C}(\widehat{\mathcal{C}})$ as in Ex. 2.1.

Viceversa, if M is a (locally) compact topological space the spaces $\widehat{C_0(M)}$ and M can be identified both setwise and topologically. Therefore, there is a one-to-one correspondence between the $*$-isomorphism classes of commutative C^*-algebras and the homeomorphism classes of locally compact Hausdorff spaces. It is a complete duality between the category of (locally) compact Hausdorff spaces and (proper and) continuous maps and the category of commutative (not necessarily) unital C^*-algebras and *-homomorphisms.

In fact, if M is a (compact) differentiable manifold, one would like to reconstruct it from the algebra $C^\infty(M)$ of smooth functions which is only a Frèchet algebra, although it is dense in the C^*-algebra $C(M)$ of continuous functions. However, all characters of $C^\infty(M)$ are evaluations at points of M and any such a character extends to a character of $C(M)$. Thus, the Frèchet algebra $C^\infty(M)$ does select M in the sense that two manifolds are diffeomorphic if and only if the corresponding algebras of differentiable functions are isomorphic. However, it is not yet known how to algebraically characterize algebras which are isomorphic to Frèchet algebras of smooth functions (see also [71]).

The above topological reconstruction scheme cannot be directly generalized to a noncommutative C^*-algebra $\mathcal{A}$. In this case, there is more than one candidate for the analogue of the topological space M. One is the *structure space* $\widehat{\mathcal{A}}$, the space of all unitary equivalence classes of irreducible *-representations with the regional topology. A second possibility is the *primitive spectrum* $\mathrm{Prim}(\mathcal{A})$ of $\mathcal{A}$, the space of kernels of irreducible *-representations with the Jacobson topology. While for a commutative C^*-algebra these spaces agree, this is no longer true for a general C^*-algebra $\mathcal{A}$, not even setwise. For instance, $\widehat{\mathcal{A}}$ may be very complicate while $\mathrm{Prim}(\mathcal{A})$ consisting of a single point.

A very fruitful way to generalize the duality between commutative algebras and spaces is to associate noncommutative algebras to quotient spaces. There are plenty of examples, notably leaf spaces of foliations, for which this association gives highly nontrivial algebras (see for instance [22]). The prototype example of quotient space in which this fact is clearly illustrated is the space of leaves of the foliation associated with the irrational rotations on an ordinary torus, and the resulting space is the noncommutative torus which plays a key role in several instances (see [62] for a thorough survey).

3 Modules as bundles

The algebraic analogue of vector bundles has its origin in a second classical duality: a vector bundle $E \to M$ over a manifold M is completely characterized by the space $\mathcal{E} = \Gamma(E, M)$ of its smooth sections. Since the algebra acts on the sections, the space of sections can be thought of as a (right) module over the algebra $C^\infty(M)$ of smooth functions over M. Indeed by the Serre-Swan theorem [69, 16], locally trivial, finite-rank complex vector bundles over a compact Hausdorff space M correspond canonically to finite projective modules over the algebra $\mathcal{A} = C^\infty(M)$.

Let $\mathcal{A}$ be an algebra over $\mathbb{C}$. A vector space $\mathcal{E}$ over $\mathbb{C}$ is a right module over $\mathcal{A}$ if it carries a right representation of $\mathcal{A}$:

$$\mathcal{E} \times \mathcal{A} \ni (\eta, a) \mapsto \eta a \in \mathcal{E} \ ,$$
$$\eta(ab) = (\eta a)b \ , \ \ \eta(a+b) = \eta a + \eta b \ , \ \ (\eta + \xi)a = \eta a + \xi a \ ,$$

for any $\eta, \xi \in \mathcal{E}$ and $a, b \in \mathcal{A}$. A *left module* is defined in a similar way. A *bimodule* over the algebra $\mathcal{A}$ is a vector space $\mathcal{E}$ which carries both a left and a right $\mathcal{A}$-module structure and the two structures are required to be compatible, namely

$$(a\eta)b = a(\eta b) \ , \quad \forall \, \eta \in \mathcal{E} \ , \ a, b \in \mathcal{A} \ .$$

Any right (respectively left) $\mathcal{A}$-module $\mathcal{E}$ can be regarded as a left (respectively right) module over the opposite algebra $\mathcal{A}^o$ by setting $a^o\eta = \eta a$ (respectively $a\eta = \eta a^o$), for any $\eta \in \mathcal{E}, a \in \mathcal{A}$. Any $\mathcal{A}$-bimodule $\mathcal{E}$ can be regarded as a right module over the enveloping algebra $\mathcal{A}^e := \mathcal{A} \otimes \mathcal{A}^o$by setting $\eta(a \otimes b^o) = b\eta a$, for any $\eta \in \mathcal{E}, a \in \mathcal{A}, b^o \in \mathcal{A}^o$. One can also regard $\mathcal{E}$ as a left $\mathcal{A}^e$-module by setting $(a \otimes b^o)\eta = a\eta b$, for any $\eta \in \mathcal{E}, a \in \mathcal{A}, b^o \in \mathcal{A}^o$.

A family (e_n) is a *generating family* for the right module $\mathcal{E}$ if any element of $\mathcal{E}$ can be written (possibly in more than one way) as a finite combination $\sum_n e_n a_n$, with $a_n \in \mathcal{A}$. The family (e_n) is *free* if it is made of linearly independent elements (over $\mathcal{A}$), and it is a *basis* for the module $\mathcal{E}$ if it is a free generating family, so that any $\eta \in \mathcal{E}$ can be written uniquely as a combination $\sum_n e_n a_n$, with $a_n \in \mathcal{A}$. A module is called *free* if it admits a basis. A module is said to be of *finite type* if it is finitely generated, namely if it admits a generating family of finite cardinality.

Consider the free module $\mathbb{C}^N \otimes \mathcal{A} := \mathcal{A}^N$. Any element $\eta \in \mathcal{A}^N$ can be thought of as an N-dimensional vector with entries in $\mathcal{A}$ and can be written uniquely as a linear combination $\eta = \sum_{j=1}^N e_j a_j$, with $a_j \in \mathcal{A}$ and the basis $\{e_j, \ j = 1, \ldots, N\}$ being identified with the canonical basis of $\mathbb{C}^N$. This module is both free and of finite type. A general free module (of finite type) *might* admit bases of different cardinality and so it does not make sense to talk of dimension. If the free module is such that any two bases have the same cardinality (this is, for instance, the case if $\mathcal{A}$ is commutative [6]), the latter is called the *dimension* of the module. However, if the module $\mathcal{E}$ is of finite type there is always an integer N and a (module) surjection $\rho : \mathcal{A}^N \to \mathcal{E}$. Then one has a basis $\{\varepsilon_j, \ j = 1, \ldots, N\}$ for $\mathcal{E}$ which is the image of the free basis of $\mathcal{A}^N$, $\varepsilon_j = \rho(e_j)$. In general one cannot solve the constraints among the basis elements so as to get a free basis.

3.1 Projective modules of finite type

By the Serre-Swan theorem, the particular kind of modules which correspond to vector bundles are not only of finite type but also projective.

Definition 3.1 A right $\mathcal{A}$-module $\mathcal{E}$ is projective if it is a direct summand in a free module, that is there exists a free module $\mathcal{F}$ and a module $\mathcal{E}'$ (which is then projective as well) such that

$$\mathcal{F} = \mathcal{E} \oplus \mathcal{E}' \ .$$

An equivalent and useful definition of projective modules is that given a surjective homomorphism $\rho : \mathcal{M} \to \mathcal{N}$ of right $\mathcal{A}$-modules, any homomorphism $\lambda : \mathcal{E} \to \mathcal{N}$ can be lifted to a homomorphism $\widetilde{\lambda} : \mathcal{E} \to \mathcal{M}$ such that $\rho \circ \widetilde{\lambda} = \lambda$. Then, if the module $\mathcal{E}$ is both projective and of finite type with surjection $\rho : \mathcal{A}^N \to \mathcal{E}$, there exists a lift $\widetilde{\lambda} : \mathcal{E} \to \mathcal{A}^N$ such that $\rho \circ \widetilde{\lambda} = \mathrm{id}_{\mathcal{E}}$. We can then construct an idempotent $e \in \mathrm{End}_{\mathcal{A}}\mathcal{A}^N \simeq \mathbb{M}_N(\mathcal{A})$, $\mathbb{M}_N(\mathcal{A})$ being the algebra of $N \times N$ matrices with entries in $\mathcal{A}$, given by

$$e = \widetilde{\lambda} \circ \rho \ .$$

Indeed, $e^2 = \widetilde{\lambda} \circ \rho \circ \widetilde{\lambda} \circ \rho = \widetilde{\lambda} \circ \rho = e$. The idempotent e allows one to decompose the free module $\mathcal{A}^N$ as a direct sum of submodules,

$$\mathcal{A}^N = e\mathcal{A}^N \oplus (\mathbb{I} - e)\mathcal{A}^N ,$$

and in this way ρ and $\widetilde{\lambda}$ are isomorphisms (inverses of each other) between $\mathcal{E}$ and $e\mathcal{A}^N$. The module $\mathcal{E}$ is then projective of finite type over $\mathcal{A}$ if and only if there exits an idempotent $e \in \mathbb{M}_N(\mathcal{A}),\ e^2 = e$, such that $\mathcal{E} = e\mathcal{A}^N$. We shall use the term *finite projective* to mean *projective of finite type*. The crucial link between finite projective modules and vector bundles is provided by the Serre-Swan theorem.

Theorem 3.1 *Let M be a compact finite dimensional manifold. Any $C^\infty(M)$-module $\mathcal{E}$ is isomorphic to the module $\Gamma(E, M)$ of smooth sections of a bundle $E \to M$ if and only if $\Gamma(E, M)$ is finite projective.*

This theorem was first established for the continuous category, i.e. for continuous functions and sections in [69], and extended to the smooth case in [16] (see also [39, 47]). It says that with $\mathcal{A} = C^\infty(M)$, one can find an integer N and an idempotent $p \in \mathbb{M}_N(\mathcal{A})$ such that the module $\Gamma(E, M)$ is written as $\Gamma(E, M) = p\mathcal{A}^N$.

4 Homology and cohomology

We shall take $\mathcal{A}$ to be a generic associative algebra over $\mathbb{C}$ with unit $\mathbb{I}$.

4.1 Differential calculi

Given the algebra $\mathcal{A}$, let Γ be a bimodule over $\mathcal{A}$, and let $\mathrm{d} : \mathcal{A} \to \Gamma$ be an additive map. We say that the pair (Γ, d) is a *first order differential calculus* over $\mathcal{A}$ if it happens that

(1) there is a Leibniz rule:

$$\mathrm{d}(ab) = (\mathrm{d}a)b + a\mathrm{d}b , \quad \forall\, a, b \in \mathcal{A} ;$$

(2) any element $\omega \in \Gamma$ is of the form,

$$\omega = \sum_i a_i \mathrm{d}b_i , \quad a_i, b_i \in \mathcal{A} .$$

From the Leibniz rule it follows that $\mathrm{d}(\mathbb{I}) = 0$ but a generic element of $\mathbb{C}$ need not be killed by d. For simplicity one asks that $\mathrm{d}\mathbb{C} = 0$, which is equivalent to the additional requirement that $\mathrm{d} : \mathcal{A} \to \Gamma$ is a linear map. Any two first order differential calculi (Γ, d) and (Γ', d') are be isomorphic if there is a bimodule isomorphism $\phi : \Gamma \to \Gamma'$ such that

$$\phi(\mathrm{d}a) = \mathrm{d}'a , \quad \forall\, a \in \mathcal{A} .$$

Example 4.1 There is a universal first order differential calculus associated with any $\mathcal{A}$. Consider first the submodule of $\mathcal{A} \otimes \mathcal{A}$ given by

$$\Omega^1\mathcal{A} := \ker(m : \mathcal{A} \otimes \mathcal{A} \to \mathcal{A}) , \ \ m(a \otimes b) = ab ,$$

and define the differential δ by

$$\delta : \mathcal{A} \to \Omega^1\mathcal{A}\,, \quad \delta a := \mathbb{I} \otimes a - a \otimes \mathbb{I}\,.$$

The submodule $\Omega^1\mathcal{A}$ is generated by elements of the form $1 \otimes a - a \otimes 1$ with $a \in \mathcal{A}$. Indeed, if $\sum_i a_i b_i = m(\sum_i a_i \otimes b_i) = 0$, then one gets

$$\sum_i a_i \otimes b_i = \sum_i a_i (1 \otimes b_i - b_i \otimes 1) = \sum_i a_i \delta b_i$$

as it should be. Thus $(\Omega^1\mathcal{A}, \delta)$ is a first order differential calculus over $\mathcal{A}$.

Any first order differential calculus over $\mathcal{A}$ can be obtained from the universal one.

Proposition 4.1 *Let $\mathcal{N}$ be any sub bimodule of $\Omega^1\mathcal{A}$ with canonical projection given by $\pi : \Omega^1\mathcal{A} \to \Gamma = \Omega^1\mathcal{A}/\mathcal{N}$ and define $\mathrm{d} = \pi \circ \delta$. Then (Γ, d) is a first order differential calculus over $\mathcal{A}$ and any such a calculus can be obtained in this way.*

The first statement is obvious. Conversely, if (Γ, d) is a first order differential calculus over $\mathcal{A}$, define $\pi : \Omega^1\mathcal{A} \to \Gamma$ by

$$\pi\left(\sum_i a_i \otimes b_i\right) := \sum_i a_i \mathrm{d} b_i\,.$$

Then, using the fact that $\sum_i a_i b_i = m(\sum_i a_i \otimes b_i) = 0$, one easily proves that π is a bimodule morphism. Moreover, π is surjective, since given $\omega = \sum_i a_i \mathrm{d} b_i \in \Gamma$, the element $\widetilde{\omega} = \sum_i a_i \otimes b_i - (\sum_i a_i b_i) \otimes \mathbb{I}$ belongs to $\Omega^1\mathcal{A}$, $m(\widetilde{\omega}) = 0$, and projects to ω, $\pi(\widetilde{\omega}) = \omega - (\sum_i a_i b_i)\mathrm{d}\mathbb{I} = \omega$. Define then the sub-bimodule $\mathcal{N}$ of $\Omega^1\mathcal{A}$ by

$$\mathcal{N} := \ker \pi = \left\{ \sum_i a_i \otimes b_i \in \Omega^1\mathcal{A} \mid \sum_i a_i \mathrm{d} b_i = 0 \right\}\,.$$

Finally, $\pi(\delta a) = \pi(\mathbb{I} \otimes a - a \otimes \mathbb{I}) = \mathbb{I}\mathrm{d}a - a\mathrm{d}\mathbb{I} = \mathrm{d}a$, which shows that $\pi \circ \delta = \mathrm{d}$ and concludes the proof that (Γ, d) and $(\Omega^1\mathcal{A}/\mathcal{N}, \delta)$ are isomorphic.

With any given algebra $\mathcal{A}$, there is associated a *universal graded differential algebra of forms* $\Omega\mathcal{A} = \bigoplus_p \Omega^p\mathcal{A}$. In degree 0, symply $\Omega^0\mathcal{A} = \mathcal{A}$. The space $\Omega^1\mathcal{A}$ of *one-forms* has been constructed explicitly in terms of tensor products in Ex. 4.1. One thinks of $\Omega^1\mathcal{A}$ as generated, as a left $\mathcal{A}$-module, by symbols δa for $a \in \mathcal{A}$ with relations

$$\delta(ab) = (\delta a)b + a\delta b\,, \quad \forall\, a, b \in \mathcal{A}\,. \tag{4.1}$$

$$\delta(\alpha a + \beta b) = \alpha\delta a + \beta\delta b\,, \quad \forall\, a, b \in \mathcal{A}\,, \quad \alpha, \beta \in \mathbb{C}\,. \tag{4.2}$$

A generic element $\omega \in \Omega^1\mathcal{A}$ is a finite sum of the form $\omega = \sum_i a_i \delta b_i\,,\; a_i, b_i \in \mathcal{A}$. The left $\mathcal{A}$-module $\Omega^1\mathcal{A}$ can also be endowed with a structure of a right $\mathcal{A}$-module by using (4.1),

$$\left(\sum_i a_i \delta b_i\right) c := \sum_i a_i (\delta b_i) c = \sum_i a_i \delta(b_i c) - \sum_i a_i b_i \delta c\,.$$

The relation (4.1) is just the Leibniz rule for the map $\delta : \mathcal{A} \to \Omega^1\mathcal{A}$, which therefore is a derivation of $\mathcal{A}$ with values in the bimodule $\Omega^1\mathcal{A}$. The requirement (4.2) gives $\delta\mathbb{C} = 0$.

The space $\Omega^p\mathcal{A}$ of p-forms is defined as $\Omega^p\mathcal{A} = \Omega^1\mathcal{A}\,\Omega^1\mathcal{A}\cdots\Omega^1\mathcal{A}\,\Omega^1\mathcal{A}$ (p factors), with the product of any two one-forms defined by "juxtaposition",

$$(a_0\delta a_1)(b_0\delta b_1) := a_0(\delta a_1)b_0\delta b_1 = a_0\delta(a_1 b_0)\delta b_1 - a_0 a_1 \delta b_0 \delta b_1 ,$$

for any $a_0, a_1, b_0, b_1 \in \mathcal{A}$. Thus, elements of $\Omega^p\mathcal{A}$ are finite linear combinations of monomials of the form

$$\omega = a_0\delta a_1\delta a_2\cdots\delta a_p , \quad a_k \in \mathcal{A} .$$

The product $\Omega^p\mathcal{A}\times\Omega^q\mathcal{A} \to \Omega^{p+q}\mathcal{A}$ of any p-form with any q-form produces a $p+q$ form and is again defined by juxtaposition and rearranging the result by using (4.1). Notice that there is nothing like graded commutativity of forms. The algebra $\Omega^p\mathcal{A}$ is a left $\mathcal{A}$-module by construction. Similarly to $\Omega^1\mathcal{A}$, it can also be made into a right $\mathcal{A}$-module.

One makes the algebra $\Omega\mathcal{A}$ a differential algebra by extending the *differential* δ to a linear operator $\delta : \Omega^p\mathcal{A} \to \Omega^{p+1}\mathcal{A}$, unambiguously by

$$\delta(a_0\delta a_1\cdots\delta a_p) := \delta a_0\delta a_1\cdots\delta a_p .$$

It is nilpotent, $\delta^2 = 0$, and a graded derivation,

$$\delta(\omega_1\omega_2) = \delta(\omega_1)\omega_2 + (-1)^p\omega_1\delta\omega_2 , \quad \forall\, \omega_1 \in \Omega^p\mathcal{A} ,\ \omega_2 \in \Omega\mathcal{A} .$$

The Prop. 4.1 is a manifestation of the fact that the graded differential algebra $(\Omega\mathcal{A}, \delta)$ is universal in the following sense [5, 9, 41].

Proposition 4.2 *Let* $(\Gamma = \oplus_p\Gamma^p, \mathrm{d})$ *be a graded differential algebra, and let* $\rho : \mathcal{A} \to \Gamma^0$ *be a morphism of unital algebras. Then there exists a unique extension of* ρ *to a morphism of graded differential algebras* $\widetilde{\rho} : \Omega\mathcal{A} \to \Gamma$ *such that* $\widetilde{\rho}\circ\delta = \mathrm{d}\circ\widetilde{\rho}$.

As a consequence, just as any first order differential calculus over $\mathcal{A}$ can be obtained as a quotient of the universal one $\Omega^1\mathcal{A}$, any graded differential algebra is a quotient of the universal $(\Omega\mathcal{A}, \delta)$. One should remark that the latter is not very interesting from the cohomological point of view; all cohomology spaces

$$H^p(\Omega\mathcal{A}) := \ker(\delta : \Omega^p\mathcal{A} \to \Omega^{p+1}\mathcal{A})/\mathrm{im}(\delta : \Omega^{p-1}\mathcal{A} \to \Omega^p\mathcal{A})$$

vanish, except in degree zero, $H^0(\Omega\mathcal{A}) = \mathbb{C}$. Indeed, there is a contracting homotopy $k : \Omega^p\mathcal{A} \to \Omega^{p+1}\mathcal{A}$, giving $k\delta + \delta k = \mathbb{I}$, and defined by

$$k(a_0\delta a_1\cdots\delta a_p) := (-1)^{p+1}a_0\delta a_1\cdots\delta a_{p-1}\, a_p .$$

4.2 Hochschild and cyclic homology

Given an algebra $\mathcal{A}$, consider the chain complex $(C_*(\mathcal{A}) = \bigoplus_n C_n(\mathcal{A}), b)$ with chains $C_n(\mathcal{A}) = \mathcal{A}^{\otimes(n+1)}$ and the boundary map $b : C_n(\mathcal{A}) \to C_{n-1}(\mathcal{A})$ defined by

$$\begin{aligned} b(a_0\otimes a_1\otimes\cdots\otimes a_n) := &\sum_{j=0}^{n-1}(-1)^j a_0\otimes\cdots\otimes a_j a_{j+1}\otimes\cdots\otimes a_n \\ &+ (-1)^n a_n a_0\otimes a_1\otimes\cdots\otimes a_{n-1} . \end{aligned} \tag{4.3}$$

It is easy to prove that $b^2 = 0$. The *Hochschild homology* $HH_*(\mathcal{A})$ of the algebra $\mathcal{A}$ is the homology of this complex, $HH_n(\mathcal{A}) := H_n(C_*(\mathcal{A}), b) = Z_n/B_n$, with the cycles given by$Z_n := \ker(b : C_n(\mathcal{A}) \to C_{n-1}(\mathcal{A}))$ and boundaries by $B_n := \mathrm{im}(b : C_{n+1}(\mathcal{A}) \to C_n(\mathcal{A}))$.

Remark 4.1 One can generalize the previous constructions by taking chains with values in any $\mathcal{A}$-bimodue $\mathcal{E}$. One defines $C_n(\mathcal{A}, \mathcal{E}) := \mathcal{E} \otimes \mathcal{A}^{\otimes n}$ on which the formula (4.3) makes perfect sense when $a_a \in \mathcal{E}$ due to the bimodule structure of $\mathcal{E}$. The homology of this complex is then denoted by $HH_*(\mathcal{A}, \mathcal{E})$.

Let us go back to the original case when $\mathcal{E} = \mathcal{A}$. Besides b we have another operator which increases the degree, $B : C_n(\mathcal{A}) \to C_{n+1}(\mathcal{A})$, written $B = B_0 A$, where

$$\begin{aligned} B_0(a_0 \otimes a_1 \otimes \cdots \otimes a_n) &:= \mathbb{I} \otimes a_0 \otimes a_1 \otimes \cdots \otimes a_n \\ A(a_0 \otimes a_1 \otimes \cdots \otimes a_n) &:= \sum_{j=0}^{n} \frac{(-1)^{nj}}{n+1} a_j \otimes a_{j+1} \otimes \cdots \otimes a_{j-1} \,, \end{aligned} \tag{4.4}$$

with the obvious cyclic identification $n+1 = 0$. One checks that $B^2 = 0$ and $bB + Bb = 0$.

Putting together these two operators one gets a bi-complex $(C_*(\mathcal{A}), b, B)$ with $C_{p-q}(\mathcal{A})$ in bi-degree p, q. The *cyclic homology* $HC_*(\mathcal{A})$ of the algebra $\mathcal{A}$ is the homology of the total complex $(CC(\mathcal{A}), b + B)$, whose n-th term is given by $CC_n(\mathcal{A}) := \oplus_{p+q=n} C_{p-q}(\mathcal{A}) = \oplus_{0 \le q \le [n/2]} C_{2n-q}(\mathcal{A})$. Then,

$$HC_n(\mathcal{A}) := H_n(CC(\mathcal{A}), b + B) = Z_n^\lambda / B_n^\lambda \,,$$

with the *cyclic cycles* given by $Z_n^\lambda := \ker(b + B : CC_n(\mathcal{A}) \to CC_{n-1}(\mathcal{A}))$ and the *cyclic boundaries* given by $B_n^\lambda := \mathrm{im}(b + B : CC_{n+1}(\mathcal{A}) \to CC_n(\mathcal{A}))$.

Example 4.2 If M is a compact manifold, the Hochschild homology of the algebra $C^\infty(M)$ of smooth functions gives the deRham complex (Hochschild-Konstant-Rosenberg theorem),

$$\Omega^k_{dR}(M) \simeq HH_k(C^\infty(M)) \,,$$

with $\Omega^k_{dR}(M)$ the space of deRham forms of degree k on M. If d denotes the deRham exterior differential, this isomorphisms is implemented by

$$a_0 \mathrm{d}a_1 \wedge \cdots \wedge \mathrm{d}a_k \mapsto \varepsilon_k(a_0 \otimes a_1 \otimes \cdots \otimes \mathrm{d}a_k)$$

where ε_k is the *antisymmetrization map*

$$\varepsilon_k(a_0 \otimes a_1 \otimes \cdots \otimes \mathrm{d}a_k) := \sum_{\sigma \in S_k} sign(\sigma)(a_0 \otimes a_{\sigma(1)} \otimes \cdots \otimes \mathrm{d}a_{\sigma(k)})$$

and S_k is the symmetric group of degree k. In particular one checks that $b \circ \varepsilon_k = 0$. The deRham differential d corresponds to the operator B_* (the lift of B to homology):

$$\varepsilon_{k+1} \circ \mathrm{d} = (k+1) B_* \circ \varepsilon_k \,.$$

On the other hand, the cyclic homology gives [13, 51]

$$HC_k(C^\infty(M)) = \Omega^k_{dR}(M)/\mathrm{d}\Omega^{k-1}_{dR}(M) \oplus H^{k-2}_{dR}(M) \oplus H^{k-4}_{dR}(M) \oplus \cdots \,, \tag{4.5}$$

where $H^j_{dR}(M)$ is the j-th deRham cohomology group. The last term in the sum is $H^0_{dR}(M)$ or $H^1_{dR}(M)$ according to whether k is even or odd. Since $C^\infty(M)$ is commutative there is a natural decomposition (the λ-decomposition) of cyclic homology into smaller pieces,

$$HC_0(C^\infty(M)) = HC_0^{(0)}(C^\infty(M)) \, ,$$
$$HC_k(C^\infty(M)) = HC_k^{(k)}(C^\infty(M)) \oplus \cdots \oplus HC_k^{(1)}(C^\infty(M)) \, ,$$

obtained by suitable idempotents $e_k^{(i)}$ commuting with the operator B: $Be_k^{(i)} = e_{k+1}^{(i+1)}B$. It corresponds to the decomposition above and permits to extract the deRham cohomology

$$\begin{aligned} &HC_k^{(k)}(C^\infty(M)) = \Omega^k_{dR}(M)/\mathrm{d}\Omega^{k-1}_{dR}(M) \, , && \\ &HC_k^{(i)}(C^\infty(M)) = H^{2i-k}_{dR}(M) \, , && \text{for} \quad [n/2] \le i < n \, , \\ &HC_k^{(i)}(C^\infty(M)) = 0 \, , && \text{for} \quad i < [n/2] \, . \end{aligned}$$

This example shows that it is possible to think of cyclic homology as a generalization of deRham cohomology to the noncommutative setting.

4.3 Hochschild and cyclic cohomology

Similarly, one describes the dual theories. A Hochschild k cochain on the algebra $\mathcal{A}$ is an $(n+1)$-linear functional on $\mathcal{A}$ or a linear form on $\mathcal{A}^{\otimes(n+1)}$. Let

$$C^n(\mathcal{A}) = \mathrm{Hom}(\mathcal{A}^{\otimes(n+1)}, \mathbb{C})$$

be the collection of such cochains. We have a cochain complex $(C^*(\mathcal{A}) = \bigoplus_n C^n(\mathcal{A}), b)$ with a coboundary map, $b : C^n(\mathcal{A}) \to C^{n+1}(\mathcal{A})$, defined by

$$\begin{aligned} b\varphi(a_0, a_1, \cdots, a_{n+1}) :=& \sum_{j=0}^{n} (-1)^j \varphi(a_0, \cdots, a_j a_{j+1}, \cdots, a_{n+1}) \\ & + (-1)^{n+1} \varphi(a_{n+1} a_0 a_1, \cdots, a_n) \, . \end{aligned} \tag{4.6}$$

Clearly $b^2 = 0$ and the *Hochschild cohomology* $HH^*(\mathcal{A})$ of the algebra $\mathcal{A}$ is the cohomology of this complex, $HH^n(\mathcal{A}) := H^n(C^*(\mathcal{A}), b) = Z^n/B^n$, with the cocycles given by $Z^n := \ker(b : C^n(\mathcal{A}) \to C^{n+1}(\mathcal{A}))$ and the coboundaries by $B^n := \mathrm{im}(b : C^{n-1}(\mathcal{A}) \to C^n(\mathcal{A}))$.

Remark 4.2 The previous constructions can be generalized by taking cochains with values in any $\mathcal{A}$-bimodue $\mathcal{E}$: $C^n(\mathcal{A}, \mathcal{E})$ is the space of n-linear maps $\varphi : \mathcal{A}^{\otimes n} \to \mathcal{E}$ with an $\mathcal{A}$-bimodule structure given by $(a' \varphi a'')(a_0, a_1, \cdots, a_n) = a' \varphi(a_0, a_1, \cdots, a_n) a''$. The coboundary map (4.6) is generalized to

$$\begin{aligned} b\varphi(a_1, \cdots, a_{n+1}) \quad := \quad & a_1 \varphi(a_2, \cdots, a_{n+1}) \\ & + \sum_{j=1}^{n} (-1)^j \varphi(a_1, \cdots, a_j a_{j+1}, \cdots, a_{n+1}) \\ & + (-1)^{n+1} \varphi(a_1, \cdots, a_n) a_{n+1} \, . \end{aligned}$$

The cohomology of this complex is then denoted by $HH^*(\mathcal{A}, \mathcal{E})$.

A Hochschild 0-cocycle τ is a *trace* on $\mathcal{A}$, since $\tau \in \mathrm{Hom}(\mathcal{A}, \mathbb{C})$ $b\tau = 0$ reads now,

$$\tau(a_0 a_1) - \tau(a_1 a_0) = b\tau(a_0, a_1) = 0 \, .$$

The trace property is extended to higher orders by saying that an n-cochain φ is *cyclic* if $\lambda\varphi = \varphi$, with the map λ defined by,

$$\lambda\varphi(a_0, a_1, \cdots, a_n) = (-1)^n \varphi(a_n, a_0, \cdots, a_{n-1}) \, .$$

A *cyclic cocycle* is a cyclic cochain φ for which $b\varphi = 0$. A straightforward computation shows that the sets of cyclic n-cochains $C^n_\lambda(\mathcal{A}) = \{\varphi \in C^n(\mathcal{A}) \mid \lambda\varphi = \varphi\}$ are preserved by the Hochschild boundary operator: $(1-\lambda)\varphi = 0$ implies that $(1-\lambda)b\varphi = 0$. Thus we get a subcomplex $(C^*_\lambda(\mathcal{A}) = \bigoplus_n C^n_\lambda(\mathcal{A}), b)$ of the complex $(C^*(\mathcal{A}) = \bigoplus_n C^n(\mathcal{A}), b)$. The *cyclic cohomology* $HC^*(\mathcal{A})$ of the algebra $\mathcal{A}$ is the cohomology of this subcomplex,

$$HC^n(\mathcal{A}) := H^n(C^*_\lambda(\mathcal{A}), b) = Z^n_\lambda / B^n_\lambda \, ,$$

with the cyclic cocycles given by $Z^n_\lambda := \ker(b : C^n_\lambda(\mathcal{A}) \to C^{n+1}_\lambda(\mathcal{A}))$ and the cyclic coboundaries given by $B^n_\lambda := \mathrm{im}(b : C^{n-1}_\lambda(\mathcal{A}) \to C^n_\lambda(\mathcal{A}))$.

One can also define an operator B dual to the one in (4.4) for the homology and give a bicomplex description of cyclic cohomology; we shall not use this description and we only refer to [16] for all details. An additional important operator is the *periodicity operator*, a degree two map between cyclic cocycles, $S : Z^{n-1}_\lambda \longrightarrow Z^{n+1}_\lambda$, given by

$$S\varphi(a_0, a_1, \cdots, a_{n+1}) := -\frac{1}{n(n+1)} \sum_{j=1}^{n} \varphi(a_0, \cdots, a_{j-1}a_j a_{j+1}, \cdots, a_{n+1})$$
$$- \frac{1}{n(n+1)} \sum_{1 \le i < j \le n}^{n} (-1)^{i+j} \varphi(a_0, \cdots, a_{i-1}a_i, \cdots, a_j a_{j+1}, \cdots, a_{n+1}) \, . \tag{4.7}$$

In fact, one has the stronger result $S(Z^{n-1}_\lambda) \subseteq B^{n+1}$, the latter being the Hochschild coboundaries; and cyclicity is easy to show. The induced morphisms in cohomology $S : HC^n \to HC^{n+2}$ define two directed systems of abelian groups. Their inductive limits

$$HP^0(\mathcal{A}) := \lim_{\to} HC^{2n}(\mathcal{A}) \, , \quad HP^1(\mathcal{A}) := \lim_{\to} HC^{2n+1}(\mathcal{A}) \, ,$$

form a $\mathbb{Z}_2$-graded group which is called the *periodic cyclic cohomology* $HP^*(\mathcal{A})$ of the algebra $\mathcal{A}$. There is also a *periodic cyclic homology* [16, 51].

5 The Chern characters

There are two kinds of Chern characters, dual to each other. The first one, $\mathrm{ch}_*(\cdot)$, leads to well defined maps from the K-theory groups $K_*(\mathcal{A})$ of projections and unitaries to (period) cyclic homology of $\mathcal{A}$. The dual Chern character, $\mathrm{ch}^*(\cdot)$, of even and odd Fredholm modules provides similar maps to (period) cyclic cohomology of $\mathcal{A}$. These characters are coupled by index theorems.

5.1 The Chern character of idempotents and unitaries

The Chern character of a projection $e \in \mathbb{M}_m(\mathcal{A})$ is a formal sum, $\mathrm{ch}_*(e) = \sum_k \mathrm{ch}_k(e)$, with

$$\mathrm{ch}_k(e) = N_k \left(e^{i_0}_{i_1} - \frac{1}{2}\delta^{i_0}_{i_1} \right) \otimes e^{i_1}_{i_2} \otimes e^{i_2}_{i_3} \otimes \cdots \otimes e^{i_{2k}}_{i_0}, \tag{5.1}$$

and N_k suitable normalization constants. The crucial property of $\mathrm{ch}_*(e)$ is that it defines a cycle in the (b, B) bicomplex of cyclic homology [13, 16, 51],

$$(b + B)\mathrm{ch}_*(e) = 0 \,, \quad B\,\mathrm{ch}_k(e) = b\,\mathrm{ch}_{k+1}(e) \,,$$

and the map $e \mapsto \mathrm{ch}_*(e)$ leads to a well defined map from the K-theory group $K_0(\mathcal{A})$ to the cyclic homology of $\mathcal{A}$.

In the odd case unitary elements are used, instead of projections. The Chern character of unitary $u \in \mathbb{M}_r(\mathcal{A})$ is a formal, $\mathrm{ch}_*(u) = \sum_k \mathrm{ch}_{k+\frac{1}{2}}(u)$, with

$$\mathrm{ch}_{k+\frac{1}{2}}(u) = N_k \Big(u^{i_0}_{i_1} \otimes (u^*)^{i_1}_{i_2} \otimes u^{i_2}_{i_3} \otimes \cdots \otimes (u^*)^{i_{2k+1}}_{i_0} - (u^*)^{i_0}_{i_1} \otimes u^{i_1}_{i_2} \otimes (u^*)^{i_2}_{i_3} \otimes \cdots \otimes u^{i_{2k+1}}_{i_0} \Big) , \tag{5.2}$$

and N_k suitable normalization constants. Again $\mathrm{ch}_*(u)$ defines a cycle in the (b, B) bicomplex of cyclic homology [13, 16, 51],

$$(b + B)\mathrm{ch}_*(u) = 0 \,, \quad B\,\mathrm{ch}_{k+\frac{1}{2}}(e) = b\,\mathrm{ch}_{k+\frac{1}{2}+1}(e)$$

and the map $u \mapsto \mathrm{ch}_*(u)$ leads to a well defined map from the K-theory group $K_1(\mathcal{A})$ to the cyclic homology of $\mathcal{A}$. In fact, in both even and odd cases the correct receptacle is periodic cyclic homology, but we shall not dwell upon this point here and refer to [51].

5.2 Fredholm modules and index theorems

A Fredholm module can be thought of as an abstract elliptic operator. The full fledged theory started with Atiyah and culminated in the KK-theory of Kasparov and the cyclic cohomology of Connes. We shall only mention the few facts that we shall need later on.

Definition 5.1 ([13]) Let $\mathcal{A}$ be an involutive algebra. An *odd* Fredholm module over $\mathcal{A}$ is the datum of

(1) a representation ψ of the algebra $\mathcal{A}$ on a Hilbert space $\mathcal{H}$;

(2) an operator F on $\mathcal{H}$ such that

$$F^2 = \mathbb{I} \,, \quad F^* = F \,, \qquad [F, \psi(a)] \in \mathcal{K}(\mathcal{H}) \,, \quad \forall\, a \in \mathcal{A} \,,$$

where $\mathcal{K}(\mathcal{H})$ is the algebra of compact operators on $\mathcal{H}$.

An *even* Fredholm module has, in addition, a $\mathbb{Z}_2$-grading γ of $\mathcal{H}$, $\gamma^* = \gamma$, $\gamma^2 = \mathbb{I}$, with

$$F\gamma + \gamma F = 0 \,, \qquad \psi(a)\gamma - \gamma\psi(a) = 0 \,, \quad \forall\, a \in \mathcal{A} \,.$$

With an even module we shall indicate with $\mathcal{H}^{\pm}$ and $\psi^{\pm}$ the components of the Hilbert space and of the representation with respect to the grading.

Given any positive integer r, one can extend the previous modules to a Fredholm module $(\mathcal{H}_r, F_r)$ over the algebra $\mathbb{M}_r(\mathcal{A}) = \mathcal{A} \otimes \mathbb{M}_r(\mathbb{C})$ by a simple procedure:

$$\mathcal{H}_r = \mathcal{H} \otimes \mathbb{C}^r \ , \quad \psi_r = \psi \otimes \mathbb{I} \ , \quad \mathbb{F}_r = F \otimes \mathbb{I}_r \ ,$$

and $\gamma_r = \gamma \otimes \mathbb{I}_r$ for an even module.

The importance of Fredholm modules is witnessed by the following theorem which can be associated with the names of Atiyah and Kasparov [1, 42],

Theorem 5.1 *a) Let $(\mathcal{H}, F, \gamma)$ be an even Fredholm module over the algebra $\mathcal{A}$, and let $e \in \mathbb{M}_r(\mathcal{A})$ be a projection $e^2 = e = e^*$. Then, the operator*

$$\psi_r^-(e) F_r \psi_r^+(e) : \psi_r^+(e)\mathcal{H}_r \to \psi_r^-(e)\mathcal{H}_r \ ,$$

is a Fredholm operator whose index depends only on the class of the projection e in the K-theory of $\mathcal{A}$. Thus we get an additive map

$$\varphi : K_0(\mathcal{A}) \to \mathbb{Z} \ , \qquad \varphi([e]) = \mathit{Index}\left(\psi_r^-(e) F_r \psi_r^+(e)\right) \ . \tag{5.3}$$

b) Let $(\mathcal{H}, F)$ be an odd Fredholm module over the algebra $\mathcal{A}$, and take the projection $E = \frac{1}{2}(\mathbb{I} + F)$. Let $u \in \mathbb{M}_r(\mathcal{A})$ be unitary $uu^* = u^*u = \mathbb{I}$. Then the operator

$$E_r \psi_r(u) E_r : E_r\mathcal{H}_r \to E_r\mathcal{H}_r \ ,$$

is a Fredholm operator whose index depends only on the class of the unitary u in the K-theory of $\mathcal{A}$. Thus we get an additive map

$$\varphi : K_1(\mathcal{A}) \to \mathbb{Z} \ , \qquad \varphi([u]) = \mathit{Index}\left(E_r \psi_r(u) E_r\right) \ . \tag{5.4}$$

If $\mathcal{A}$ is a C^*-algebra, then in both even and odd cases the index map φ depends only on the *K-homology* class $[(\mathcal{H}, F)] \in KK(\mathcal{A}, \mathbb{C})$, of the Fredholm module in the Kasparov KK-group, $K^*(\mathcal{A}) = KK(\mathcal{A}, \mathbb{C})$, which is the abelian group of stable homotopy classes of Fredholm modules over $\mathcal{A}$ [42]. The index pairings (5.3) and (5.4) can be given as [16]

$$\varphi(x) = \langle \mathrm{ch}^*(\mathcal{H}, F), \mathrm{ch}_*(x) \rangle \ , \quad x \in K_*(\mathcal{A}) \ ,$$

via the Chern characters

$$\mathrm{ch}^*(\mathcal{H}, F) \in HC^*(\mathcal{A}) \ , \quad \mathrm{ch}_*(x) \in HC_*(\mathcal{A}) \ ,$$

and the pairing between cyclic cohomology $HC^*(\mathcal{A})$ and cyclic homology $HC_*(\mathcal{A})$ of the algebra $\mathcal{A}$. The Chern character $\mathrm{ch}_*(x)$ in homology is given by (5.1) and (5.2) in the even and odd case respectively. As for the Chern character $\mathrm{ch}^*(x)$ in cohomology we shall give some fundamentals in the next Subsection.

5.3 The Chern characters of Fredholm modules

We recall [67] that on a Hilbert space $\mathcal{H}$, with $\mathcal{K}(\mathcal{H})$ the algebra of compact operators, one defines for $p \in [1, \infty[$ the Schatten p-class $\mathcal{L}^p$ as the ideal of compact operators for which $\mathrm{Tr}\, T^p$ is finite: $\mathcal{L}^p = \{T \in \mathcal{K}(\mathcal{H}) \mid \mathrm{Tr}\, T^p < \infty\}$. The Hölder inequality implies $\mathcal{L}^{p_1} \cdots \mathcal{L}^{p_k} \subset \mathcal{L}^p$, with $p^{-1} = \sum_{j=1}^k p_j^{-1}$.

The Fredholm module $(\mathcal{H}, F)$ over the algebra $\mathcal{A}$ is said to be *p-summable* if

$$[F, \psi(a)] \in \mathcal{L}^p\ , \quad \forall\, a \in \mathcal{A}\ .$$

This is the same for a even or odd module. For simplicity, in the rest of this section, we shall drop the symbol ψ which indicates the representation on $\mathcal{A}$ on $\mathcal{H}$. The idea is then to construct 'quantized differential forms' and integrate (via a trace) forms of sufficiently high degree, so that they belong to $\mathcal{L}^1$. In fact, one needs to introduce a conditional trace. Given an operator T on $\mathcal{H}$ such that $FT + TF \in \mathcal{L}^1$, define

$$\mathrm{Tr}'\, T := \frac{1}{2}\,\mathrm{Tr}\, F(FT + TF)\ ;$$

note that, if $T \in \mathcal{L}^1$, then $\mathrm{Tr}'\, T = \mathrm{Tr}\, T$ by cyclicity of the trace.

Let n be a nonnegative integer and let $(\mathcal{H}, F)$ be a Fredholm module over the algebra $\mathcal{A}$. We take this module to be *even* or *odd* according to whether n is even or odd. We shall also take it to be $(n+1)$-summable. We construct an n-dimensional *cycle* $(\Omega^* = \oplus_k \Omega^k, \mathrm{d}, \int)$ over the algebra $\mathcal{A}$. Elements of Ω^k are *quantized differential forms*: $\Omega^0 = \mathcal{A}$ and for $k > 0$, Ω^k is the linear span of operators of the form

$$\omega = a_0 [F, a_1] \cdots [F, a_n]\ , \quad a_j \in \mathcal{A}\ .$$

By the assumption of summability, the Hölder inequality gives $\Omega^k \subset \mathcal{L}^{\frac{n+1}{k}}$. The product in Ω^* is just the product of operators $\omega\omega' \in \Omega^{k+k'}$ for any $\omega \in \Omega^k$ and $\omega' \in \Omega^{k'}$. The differential $\mathrm{d} : \Omega^k \to \Omega^{k+1}$ is defined by

$$\mathrm{d}\omega = F\omega - (-1)^k \omega F\ , \quad \omega \in \Omega^k\ ,$$

and $F^2 = 1$ implies both $\mathrm{d}^2 = 0$ and the fact that d is a graded derivation

$$\mathrm{d}(\omega\omega') = (\mathrm{d}\omega)\omega' + (-1)^k \omega \mathrm{d}\omega'\ , \quad \omega \in \Omega^k\ , \ \ \omega' \in \Omega^{k'}\ .$$

Finally, one defines a trace $\int : \Omega^n \to \mathbb{C}$ in degree n, which is both closed ($\int \mathrm{d}\omega = 0$) and graded ($\int \omega\omega' = (-1)^{kk'} \int \omega'\omega$). First, take n to be odd. With $\omega \in \Omega^n$,

$$\int \omega := \mathrm{Tr}'\, \omega = \frac{1}{2}\,\mathrm{Tr}\, F(F\omega + \omega F)) = \frac{1}{2}\,\mathrm{Tr}\, F\mathrm{d}\omega\ , \tag{5.5}$$

is well defined since $F\mathrm{d}\omega \in \mathcal{L}^1$. If n is even and γ is the grading, with $\omega \in \Omega^n$ one defines

$$\int \omega := \mathrm{Tr}'\, \gamma\omega = \frac{1}{2}\,\mathrm{Tr}\, F(F\gamma\omega + \gamma\omega F)) = \frac{1}{2}\,\mathrm{Tr}\, \gamma F\mathrm{d}\omega\ , \tag{5.6}$$

(recall that $F\gamma = -\gamma F$); this is again well defined since $\gamma F \mathrm{d}\omega \in \mathcal{L}^1$. One straightforwardly proves closure and graded cyclicity of both the integrals (5.5) and (5.6).

The *character* of the Fredholm module is the cyclic cocycle $\tau^n \in Z^n_\lambda(\mathcal{A})$ given by, $\tau^n(a_0, a_1, \cdots, a_n) := \int a_0 \mathrm{d}a_1 \cdots \mathrm{d}a_n$, with $a_j \in \mathcal{A}$. Explicitly,

$$\begin{aligned} \tau^n(a_0, a_1, \cdots, a_n) &= \mathrm{Tr}'\, a_0 [F, a_1], \cdots, [F, a_n]\ , \ \ n \text{ odd}\ , \\ \tau^n(a_0, a_1, \cdots, a_n) &= \mathrm{Tr}'\, \gamma\, a_0 [F, a_1], \cdots, [F, a_n]\ , \ \ n \text{ even}\ . \end{aligned} \tag{5.7}$$

In both cases one checks closure, $b\tau^n = 0$, and cyclicity, $\lambda\tau^n = (-1)^n\tau^n$.

There is an ambiguity in the choice of the integer n. Given a Fredholm module $(\mathcal{H}, F)$ over $\mathcal{A}$, the parity of n is fixed. However for its precise value there is only a lower bound, determined by the $(n+1)$-summability: since $\mathcal{L}^{p_1} \subset \mathcal{L}^{p_2}$ if $p_1 \leq p_2$, one could replace n by $n + 2k$ with k any integer. Thus one gets a sequence of cyclic cocycles $\tau^{n+2k} \in Z_\lambda^{n+2k}(\mathcal{A}), k \geq 0$, with the same parity. The crucial fact is that the cyclic cohomology classes of these cocycles are related by the periodicity operator S in (4.7). The characters τ^{n+2k} satisfy the recursive relation

$$S[\tau^m]_\lambda = c_m[\tau^{m+2}]_\lambda \quad \text{in} \quad HC^{m+2}(\mathcal{A}) \, , \quad m = n + 2k \, , \; k \geq 0 \, ,$$

with c_m a constant depending on m (one could get rid of these constants by suitably normalizing the characters in (5.7)). Therefore, the sequence $\{\tau^{n+2k}\}_{k\geq 0}$ determines a well defined class $[\tau^F]$ in the periodic cyclic cohomology $HP^0(\mathcal{A})$ or $HP^1(\mathcal{A})$ according to whether n is even or odd. The class $[\tau^F]$ is the Chern character of the Fredholm module $(\mathcal{A}, \mathcal{H}, F)$ in periodic cyclic cohomology.

More details of the general theory are in [16]. Some examples are in [40].

6 Connections and gauge transformations

The notion of a (gauge) connection on a (finite projective) module $\mathcal{E}$ over an algebra $\mathcal{A}$ and with respect to a given calculus makes perfectly sense and one can develop several related concepts algebraically. We take a right module structure.

6.1 Connections on modules

Let us suppose we have an algebra $\mathcal{A}$ with a differential calculus $(\Omega\mathcal{A} = \oplus_p \Omega^p\mathcal{A}, \mathrm{d})$. A *connection* on the right $\mathcal{A}$-module $\mathcal{E}$ is a $\mathbb{C}$-linear map

$$\nabla : \mathcal{E} \otimes_\mathcal{A} \Omega^p\mathcal{A} \longrightarrow \mathcal{E} \otimes_\mathcal{A} \Omega^{p+1}\mathcal{A},$$

defined for any $p \geq 0$, and satisfying the Leibniz rule

$$\nabla(\omega\rho) = (\nabla\omega)\rho + (-1)^p\omega\mathrm{d}\rho, \;\; \forall\, \omega \in \mathcal{E} \otimes_\mathcal{A} \Omega^p\mathcal{A}, \; \rho \in \Omega\mathcal{A}.$$

A connection is completely determined by its restriction

$$\nabla : \mathcal{E} \to \mathcal{E} \otimes_\mathcal{A} \Omega^1\mathcal{A},$$

which satisfies

$$\nabla(\eta a) = (\nabla\eta)a + \eta \otimes_\mathcal{A} \mathrm{d}a, \;\; \forall\, \eta \in \mathcal{E}, \; a \in \mathcal{A} \, ,$$

and which is extended to all of $\mathcal{E} \otimes_\mathcal{A} \Omega^p\mathcal{A}$ using Leibniz rule. It is the latter rule that implies the $\Omega\mathcal{A}$-linearity of the composition,

$$\nabla^2 = \nabla \circ \nabla : \mathcal{E} \otimes_\mathcal{A} \Omega^p\mathcal{A} \longrightarrow \mathcal{E} \otimes_\mathcal{A} \Omega^{p+2}\mathcal{A}.$$

The restriction of ∇^2 to $\mathcal{E}$ is the *curvature*

$$F : \mathcal{E} \to \mathcal{E} \otimes_\mathcal{A} \Omega^2\mathcal{A},$$

of the connection. It is $\mathcal{A}$-linear, $F(\eta a) = F(\eta)a$ for any $\eta \in \mathcal{E}, a \in \mathcal{A}$, and satisfies

$$\nabla^2(\eta \otimes_{\mathcal{A}} \rho) = F(\eta)\rho, \ \forall\, \eta \in \mathcal{E}, \ \rho \in \Omega\mathcal{A}.$$

Thus, $F \in \mathrm{Hom}_{\mathcal{A}}(\mathcal{E}, \mathcal{E} \otimes_{\mathcal{A}} \Omega^2\mathcal{A})$, the latter being the collection of (right) $\mathcal{A}$-linear homomorphisms from $\mathcal{E}$ to $\mathcal{E} \otimes_{\mathcal{A}} \Omega^2\mathcal{A}$ (for this collection an alternative notation that is used in the literature, is $\mathrm{End}_{\mathcal{A}}(\mathcal{E}, \mathcal{E} \otimes_{\mathcal{A}} \Omega^2\mathcal{A})$). In order to have a Bianchi identity we need some natural generalization. Let $\mathrm{End}_{\Omega\mathcal{A}}(\mathcal{E} \otimes_{\mathcal{A}} \Omega\mathcal{A})$ be the collection of all $\Omega\mathcal{A}$-linear endomorphisms of $\mathcal{E} \otimes_{\mathcal{A}} \Omega\mathcal{A}$. It is an algebra under composition. The curvature F can be thought of as an element of $\mathrm{End}_{\Omega\mathcal{A}}(\mathcal{E} \otimes_{\mathcal{A}} \Omega\mathcal{A})$. There is a well defined map

$$[\nabla, \cdot\,] \ : \ \mathrm{End}_{\Omega\mathcal{A}}(\mathcal{E} \otimes_{\mathcal{A}} \Omega\mathcal{A}) \longrightarrow \mathrm{End}_{\Omega\mathcal{A}}(\mathcal{E} \otimes_{\mathcal{A}} \Omega\mathcal{A}),$$
$$[\nabla, T] := \nabla \circ T - (-1)^{|T|}\, T \circ \nabla,$$

where $|T|$ denotes the degree of T with respect to the $\mathbb{Z}^2$-grading of $\Omega\mathcal{A}$. It is straightforwardly checked that $[\nabla, \cdot\,]$ is a graded derivation for the algebra $\mathrm{End}_{\Omega\mathcal{A}}(\mathcal{E} \otimes_{\mathcal{A}} \Omega\mathcal{A})$,

$$[\nabla, S \circ T] = [\nabla, S] \circ T + (-1)^{|S|} S \circ [\nabla, T].$$

Proposition 6.1 *The curvature F satisfies the Bianchi identity,*

$$[\nabla, F] = 0.$$

Since F is an even element in $\mathrm{End}_{\Omega\mathcal{A}}(\mathcal{E} \otimes_{\mathcal{A}} \Omega\mathcal{A})$, the map $[\nabla, F]$ makes sense. Furthermore,

$$[\nabla, F] = \nabla \circ \nabla^2 - \nabla^2 \circ \nabla = \nabla^3 - \nabla^3 = 0.$$

In Sect. II.2 of [13], such a Bianchi identity was implicitly used in the construction of a so-called canonical cycle from a connection on a finite projective $\mathcal{A}$-module $\mathcal{E}$.

Connections always exist on a projective module. On the module $\mathcal{E} = \mathbb{C}^N \otimes_{\mathbb{C}} \mathcal{A} \simeq \mathcal{A}^N$, which is free, a connection is given by the operator

$$\nabla_0 = \mathbb{I} \otimes \mathrm{d} : \mathbb{C}^N \otimes_{\mathbb{C}} \Omega^p\mathcal{A} \longrightarrow \mathbb{C}^N \otimes_{\mathbb{C}} \Omega^{p+1}\mathcal{A}.$$

With the canonical identification $\mathbb{C}^N \otimes_{\mathbb{C}} \Omega\mathcal{A} = (\mathbb{C}^N \otimes_{\mathbb{C}} \mathcal{A}) \otimes_{\mathcal{A}} \Omega\mathcal{A} \ \simeq \ (\Omega\mathcal{A})^N$, one thinks of ∇_0 as acting on $(\Omega\mathcal{A})^N$ as the operator $\nabla_0 = (\mathrm{d}, \mathrm{d}, \cdots, \mathrm{d})$ (N-times). Next, take a projective module $\mathcal{E}$ with inclusion map, $\lambda : \mathcal{E} \to \mathcal{A}^N$, which identifies $\mathcal{E}$ as a direct summand of the free module $\mathcal{A}^N$ and idempotent $p : \mathcal{A}^N \to \mathcal{E}$ which allows one to identify $\mathcal{E} = p\mathcal{A}^N$. Using these maps and their natural extensions to $\mathcal{E}$-valued forms, a connection ∇_0 on $\mathcal{E}$ (called *Levi-Civita* or *Grassmann*) is the composition,

$$\mathcal{E} \otimes_{\mathcal{A}} \Omega^p\mathcal{A} \ \xrightarrow{\lambda}\ \mathbb{C}^N \otimes_{\mathbb{C}} \Omega^p\mathcal{A} \ \xrightarrow{\mathbb{I}\otimes\mathrm{d}}\ \mathbb{C}^N \otimes_{\mathbb{C}} \Omega^{p+1}\mathcal{A} \ \xrightarrow{p}\ \mathcal{E} \otimes_{\mathcal{A}} \Omega^{p+1}\mathcal{A},$$

that is

$$\nabla_0 = p \circ (\mathbb{I} \otimes \mathrm{d}) \circ \lambda, \tag{6.1}$$

which is simply written as $\nabla_0 = p\mathrm{d}$. The space $C(\mathcal{E})$ of all connections on $\mathcal{E}$ is an affine space modeled on $\mathrm{Hom}_{\mathcal{A}}(\mathcal{E}, \mathcal{E} \otimes_{\mathcal{A}} \Omega^1\mathcal{A})$. Indeed, if ∇_1, ∇_2 are two connections on $\mathcal{E}$, their difference is $\mathcal{A}$-linear,

$$(\nabla_1 - \nabla_2)(\eta a) = ((\nabla_1 - \nabla_2)(\eta))a, \quad \forall\, \eta \in \mathcal{E}, \ a \in \mathcal{A},$$

so that $\nabla_1 - \nabla_2 \in \mathrm{Hom}_{\mathcal{A}}(\mathcal{E}, \mathcal{E} \otimes_{\mathcal{A}} \Omega^1\mathcal{A})$. Thus, any connection can be written as

$$\nabla = p\mathrm{d} + \alpha, \tag{6.2}$$

where α is any element in $\mathrm{Hom}_{\mathcal{A}}(\mathcal{E}, \mathcal{E} \otimes_{\mathcal{A}} \Omega^1\mathcal{A})$. The "matrix of 1-forms" α as in (6.2) is called the *gauge potential* of the connection ∇. The corresponding curvature F of ∇ is

$$F = p\mathrm{d}p\mathrm{d}p + p\mathrm{d}\alpha + \alpha^2.$$

Next, let the algebra $\mathcal{A}$ have an involution *; this is extended to the whole of $\Omega\mathcal{A}$ by the requirement $(\mathrm{d}a)^* = \mathrm{d}a^*$ for any $a \in \mathcal{A}$. A *Hermitian structure* on the module $\mathcal{E}$ is a map $\langle\cdot,\cdot\rangle : \mathcal{E} \times \mathcal{E} \to \mathcal{A}$ with the properties

$$\begin{aligned} &\langle\eta, \xi a\rangle = \langle\xi, \eta\rangle\, a, \qquad \langle\eta, \xi\rangle^* = \langle\xi, \eta\rangle\,, \\ &\langle\eta, \eta\rangle \geq 0, \qquad \langle\eta, \eta\rangle = 0 \Leftrightarrow \eta = 0, \end{aligned} \tag{6.3}$$

for any $\eta, \xi \in \mathcal{E}$ and $a \in \mathcal{A}$ (an element $a \in \mathcal{A}$ is positive if it is of the form $a = b^*b$ for some $b \in \mathcal{A}$). We shall also require the Hermitian structure to be *self-dual*, i.e. every right $\mathcal{A}$-module homomorphism $\phi : \mathcal{E} \to \mathcal{A}$ is represented by an element of $\eta \in \mathcal{E}$, by the assignment $\phi(\cdot) = \langle\eta, \cdot\rangle$, the latter having the correct properties by the first of (6.3). The Hermitian structure is naturally extended to an $\Omega\mathcal{A}$-valued linear map on the product $\mathcal{E} \otimes_{\mathcal{A}} \Omega\mathcal{A} \times \mathcal{E} \otimes_{\mathcal{A}} \Omega\mathcal{A}$ by

$$\langle\eta \otimes_{\mathcal{A}} \omega, \xi \otimes_{\mathcal{A}} \rho\rangle = (-1)^{|\eta||\omega|}\omega^* \langle\eta, \xi\rangle\, \rho, \quad \forall\, \eta, \xi \in \mathcal{E} \otimes_{\mathcal{A}} \Omega\mathcal{A},\ \omega, \rho \in \Omega\mathcal{A}. \tag{6.4}$$

A connection ∇ on $\mathcal{E}$ and a Hermitian structure $\langle\cdot,\cdot\rangle$ on $\mathcal{E}$ are called compatible if,

$$\langle\nabla\eta, \xi\rangle + \langle\eta, \nabla\xi\rangle = \mathrm{d}\,\langle\eta, \xi\rangle\,, \quad \forall\, \eta, \xi \in \mathcal{E}.$$

It follows directly from the Leibniz rule and (6.4) that this extends to

$$\langle\nabla\eta, \xi\rangle + (-1)^{|\eta|}\,\langle\eta, \nabla\xi\rangle = \mathrm{d}\,\langle\eta, \xi\rangle\,, \quad \forall\, \eta, \xi \in \mathcal{E} \otimes_{\mathcal{A}} \Omega\mathcal{A}.$$

On the free module $\mathcal{A}^N$ there is a canonical Hermitian structure given by

$$\langle\eta, \xi\rangle = \sum_{j=1}^{N} \eta_j^* \xi_j, \tag{6.5}$$

with $\eta = (\eta_1, \cdots, \eta_N)$ and $\eta = (\eta_1, \cdots, \eta_N)$ any two elements of $\mathcal{A}^N$.

Under suitable regularity conditions on the algebra $\mathcal{A}$ all Hermitian structures on a given finite projective module $\mathcal{E}$ over $\mathcal{A}$ are isomorphic to each other and are obtained from the canonical structure (6.5) on $\mathcal{A}^N$ by restriction [16, II.1]. Moreover, if $\mathcal{E} = p\mathcal{A}^N$, then p is self-adjoint: $p = p^*$, with p^* obtained by the composition of the involution * in the algebra $\mathcal{A}$ with the usual matrix transposition. The Grassmann connection (6.1) is easily seen to be compatible with this Hermitian structure,

$$\mathrm{d}\,\langle\eta, \xi\rangle = \langle\nabla_0\eta, \xi\rangle + \langle\eta, \nabla_0\xi\rangle\,.$$

For a general connection (6.2), the compatibility with the Hermitian structure reduces to

$$\langle\alpha\eta, \xi\rangle + \langle\eta, \alpha\xi\rangle = 0, \quad \forall\, \eta, \xi \in \mathcal{E},$$

which just says that the gauge potential is skew-hermitian,

$$\alpha^* = -\alpha.$$

We still use the symbol $C(\mathcal{E})$ to denote the space of compatible connections on $\mathcal{E}$.

6.2 Gauge transformations

We now add the additional requirement that the algebra $\mathcal{A}$ is a Fréchet algebra and that $\mathcal{E}$ a right Fréchet module. That is, both $\mathcal{A}$ and $\mathcal{E}$ are complete in the topology defined by a family of seminorms $\|\cdot\|_i$ such that the following condition is satisfied: for all j there exists a constant c_j and an index k such that

$$\|\eta a\|_j \leq c_j \|\eta\|_k \|a\|_k .$$

The collection $\mathrm{End}_{\mathcal{A}}(\mathcal{E})$ of all $\mathcal{A}$-linear maps is an algebra with involution; its elements are also called endomorphisms of $\mathcal{E}$. It becomes a Fréchet algebra with the following family of seminorms: for $T \in \mathrm{End}_{\mathcal{A}}(\mathcal{E})$,

$$\|T\|_i = \sup_{\eta} \left\{ \|T\eta\|_i : \|\eta\|_i \leq 1 \right\}.$$

Since we are taking a self-dual Hermitian structure (see the discussion after (6.3)), any $T \in \mathrm{End}_{\mathcal{A}}(\mathcal{E})$ is adjointable, that is it admits an adjoint, an $\mathcal{A}$-linear map $T^* : \mathcal{E} \to \mathcal{E}$ such that

$$\langle T^* \eta, \xi \rangle = \langle \eta, T\xi \rangle , \quad \forall\, \eta, \xi \in \mathcal{E}.$$

The group $\mathcal{U}(\mathcal{E})$ of unitary endomorphisms of $\mathcal{E}$ is given by

$$\mathcal{U}(\mathcal{E}) := \{ u \in \mathrm{End}_{\mathcal{A}}(\mathcal{E}) \mid uu^* = u^* u = \mathbb{I}_{\mathcal{E}} \}.$$

This group plays the role of the *infinite dimensional group of gauge transformations*. It naturally acts on compatible connections by

$$(u, \nabla) \mapsto \nabla^u := u^* \nabla u, \quad \forall\, u \in \mathcal{U}(\mathcal{E}),\ \nabla \in C(\mathcal{E}), \tag{6.6}$$

where u^* is really $u^* \otimes \mathbb{I}_{\Omega\mathcal{A}}$; this will always be understood in the following. Then the curvature transforms in a covariant way

$$(u, F) \mapsto F^u = u^* F u,$$

since, evidently, $F^u = (\nabla^u)^2 = u^* \nabla u u^* \nabla u^* = u^* \nabla^2 u = u^* F u$.
As for the gauge potential, one has the usual affine transformation,

$$(u, \alpha) \mapsto \alpha^u := u^* p \mathrm{d} u + u^* \alpha u. \tag{6.7}$$

Indeed, $\nabla^u(\eta) = u^*(p\mathrm{d} + \alpha) u\eta = u^* p \mathrm{d}(u\eta) + u^* \alpha u \eta = u^* p u \mathrm{d}\eta + u^* p (\mathrm{d}u)\eta + u^* \alpha u \eta = p\mathrm{d}\eta + (u^* p \mathrm{d}u + u^* \alpha u)\eta$ for any $\eta \in \mathcal{E}$, which yields (6.7) for the transformed potential.

The "tangent vectors" to the gauge group $\mathcal{U}(\mathcal{E})$ constitute the vector space of infinitesimal gauge transformations. Suppose $\{u_t\}_{t \in \mathbb{R}}$ is a continuous family of elements in $\mathrm{End}_{\mathcal{A}}(\mathcal{E})$ (in the topology defined by the above sup-norms) and define $X := (\partial u_t / \partial t)_{t=0}$. Unitarity of u_t then induces that $X = -X^*$. Thus, for u_t to be a gauge transformation, X should be a skew-hermitian endomorphisms of $\mathcal{E}$. In this way, we understand the real vector space $\mathrm{End}^s_{\mathcal{A}}(\mathcal{E})$ of all skew-hermitian endomorphisms of $\mathcal{E}$ as made of *infinitesimal gauge transformations*. The complexification $\mathrm{End}^s_{\mathcal{A}}(\mathcal{E}) \otimes_{\mathbb{R}} \mathbb{C}$ is identified with $\mathrm{End}_{\mathcal{A}}(\mathcal{E})$.

Infinitesimal gauge transformations act on a connection in a natural way. Let the gauge transformation u_t, with $X = (\partial u_t/\partial t)_{t=0}$, act on ∇ as in (6.6). From the fact that $(\partial(u_t \nabla u_t^*)/\partial t)_{t=0} = [\nabla, X]$, we conclude that an element $X \in \mathrm{End}^s_{\mathcal{A}}(\mathcal{E})$ acts infinitesimally on a connection ∇ by the addition of $[\nabla, X]$,

$$(X, \nabla) \mapsto \nabla^X = \nabla + t[\nabla, X] + O(t^2), \quad \forall\, X \in \mathrm{End}^s_{\mathcal{A}}(\mathcal{E}),\ \nabla \in C(\mathcal{E}).$$

As a consequence, for the transformed curvature one finds

$$(X, F) \mapsto F^X = F + t[F, X] + O(t^2),$$

since $F^X = (\nabla + t[\nabla, X]) \circ (\nabla + t[\nabla, X]) = \nabla^2 + t[\nabla^2, X] + O(t^2)$.

7 Noncommutative manifolds

A noncommutative geometry is described by a spectral triple $(\mathcal{A}, \mathcal{H}, D)$. The operator D plays in general the role of the Dirac operator [50] in ordinary Riemannian geometry. It specifies both the K-homology fundamental class [16], as well as the metric on the state space of $\mathcal{A}$ (see formula (7.6) later on). The 'nontriviality' of the spectral geometry is measured by the nontriviality of the pairing between the K-theory of the algebra $\mathcal{A}$ and the K-homology class of D. There are index maps with the Fredholm module $(\mathcal{H}, F)$, with $D = |D|F$, described in Sect. 5.2,

$$\varphi : K_*(\mathcal{A}) \longrightarrow \mathbb{Z} \tag{7.1}$$

by the expressions like (5.3) and (5.4) for the even and odd case respectively.

7.1 The Dixmier trace

An algebraic generalization of the integral is via the Dixmier trace. We need a few facts about compact operators – which we take mainly from [58, 67] – and show their use an 'infinitesimals'. An operator T on a (separable, infinite dimensional) Hilbert space $\mathcal{H}$ is said to be of *finite rank* if the orthogonal complement of its null space is finite dimensional. This is equivalent to T having a finite dimensional range and what this is saying is that such an operator is a finite dimensional matrix, even if $\mathcal{H}$ is infinite dimensional. An operator T on $\mathcal{H}$ is *compact* if it can be approximated in norm by finite rank operators. An equivalent way to characterize a compact operator T is that

$$\forall\, \varepsilon > 0\,,\ \exists \ \text{ a finite dimensional subspace } \ E \subset \mathcal{H} \mid \|T|_{E^\perp}\| < \varepsilon\,. \tag{7.2}$$

The algebra $\mathcal{K}(\mathcal{H})$ is the largest two-sided ideal in the C^*-algebra $\mathcal{B}(\mathcal{H})$. In fact, it is the only norm closed and two-sided ideal; and it is an essential ideal [34]. It is a C^*-algebra without a unit, since the operator $\mathbb{I}$ on an infinite dimensional Hilbert space is not compact. The defining representation of $\mathcal{K}(\mathcal{H})$ by itself is irreducible and it is the *only* irreducible representation of $\mathcal{K}(\mathcal{H})$ up to equivalence.

Remark 7.1 In general compact operators need not admit any nonzero eigenvalue. If T is a self-adjoint compact operator there is a complete orthonormal basis for $\mathcal{H}$ made of eigenvectors, $\{\phi_k\}_{k\in\mathbb{N}}$, with eigenvalues $\lambda_k \to 0$ as $k \to \infty$. On the other hand, any compact operator T has a uniformly convergent (i.e. convergent in norm) expansion $T = \sum_{k\geq 0} \mu_k(T) \, |\psi_k\rangle \, \langle\phi_k|$, with $0 \leq \mu_{j+1} \leq \mu_j$, and $\{\psi_k\}_{k\in\mathbb{N}}, \{\phi_k\}_{k\in\mathbb{N}}$ are orthonormal sets (not necessarily complete). For this, one writes the polar decomposition $T = U|T|$, with $|T| = \sqrt{T^*T}$. Then, $\{\mu_k(T)\}$ are the *characteristic values* of T (with $\mu_0(T) = \|T\|$), that is the non-vanishing eigenvalues of the (compact self-adjoint) operator $|T|$ arranged with repeated multiplicity; $\{\phi_k\}$ are the corresponding eigenvectors and $\psi_k = U\phi_k$.

Due to condition (7.2) compact operators are in a sense 'small' and they play the role of *infinitesimals*. The size of the infinitesimal $T \in \mathcal{K}(\mathcal{H})$ is governed by the rate of decay of the sequence $\{\mu_k(T)\}$ of characteristic values as $k \to \infty$.

Definition 7.1 An operator $T \in \mathcal{K}(\mathcal{H})$ is said to be an infinitesimal of order $\alpha \in \mathbb{R}^+$ if

$$\mu_k(T) = O(k^{-\alpha}) \text{ as } k \to \infty \,, \qquad \text{i.e. } \exists\, C < \infty \mid \mu_k(T) \leq Ck^{-\alpha} \,, \quad \forall\, k \geq 1 \,.$$

The collection of all infinitesimals of order α form a (not closed) two-sided ideal in $\mathcal{B}(\mathcal{H})$, since for any $T \in \mathcal{K}(\mathcal{H})$ and $B \in \mathcal{B}(\mathcal{H})$, one has

$$\mu_k(TB) \leq \|B\|\mu_k(T) \,, \quad \mu_k(BT) \leq \|B\|\mu_k(T) \,.$$

The Dixmier trace is constructed in such a way that infinitesimals of order 1 are in the domain of the trace, while higher order infinitesimals have vanishing trace. The usual trace is not appropriate. Its domain is the two-sided ideal $\mathcal{L}^1$ of trace class operators. For any $T \in \mathcal{L}^1$, the trace, defined as $\mathrm{tr}(T) := \sum_k \langle\xi_k, T\xi_k\rangle$, is independent of the orthonormal basis $\{\xi_k\}_{k\in\mathbb{N}}$ of $\mathcal{H}$ and is, indeed, the sum of eigenvalues of T. When the operator is positive and compact, one has that $\mathrm{tr}(T) = \sum_{k=0}^{\infty} \mu_k(T)$. In general, an infinitesimal of order 1 is not in $\mathcal{L}^1$, since the only control on its characteristic values is that $\mu_k(T) \leq Ck^{-1}$ for some positive constant C. Moreover, $\mathcal{L}^1$ contains infinitesimals of order higher than 1. However, for (positive) infinitesimals of order 1, the usual trace is at most logarithmically divergent since $\sum_{k=0}^{N-1} \mu_k(T) \leq C \log N$. The Dixmier trace is just a way to extract the coefficient of the logarithmic divergence [33].

Let $\mathcal{L}^{(1,\infty)}$ be the ideal of infinitesimal of order 1; it is also denoted $\mathcal{L}^{1+}$ and named the Dixmier ideal. If $T \in \mathcal{L}^{(1,\infty)}$, one could try to define a positive functional by the limit

$$\lim_{N\to\infty} \frac{1}{\log N} \sum_{k=0}^{N-1} \mu_k(T) \,. \tag{7.3}$$

There are two problems with this formula: non-linearity and lack of convergence. For any compact operator T, consider the partial normalized sums

$$\gamma_N(T) = \frac{1}{\log N} \sum_{k=0}^{N-1} \mu_k(T) \,.$$

Given any two positive operators T_1 and T_2, they satisfy the inequality

$$\gamma_N(T_1 + T_2) \leq \gamma_N(T_1) + \gamma_N(T_2) \leq \gamma_{2N}(T_1 + T_2) \left(1 + \frac{\log 2}{\log N}\right) \,. \tag{7.4}$$

From this, linearity would follow from convergence. In general, however, the sequence $\{\gamma_N\}$, although bounded, is not convergent. Dixmier [33] proved that there exists an uncountable worth of scale invariant linear forms $\lim_\omega$ on the space $\ell^\infty(\mathbb{N})$ of bounded sequences. For each such form one gets a positive trace on the positive part of $\mathcal{L}^{(1,\infty)}$,

$$\mathrm{tr}_\omega(T) := \lim_\omega \gamma_N \ , \quad \forall\, T \in \mathcal{L}^{(1,\infty)} \ , \ T \geq 0 \ .$$

This trace is called the *Dixmier trace* and, since the eigenvalues $\mu_k(T)$ together with the sequence $\{\gamma_N\}$ are invariant under unitary tranformations, the trace is invariant as well. From (7.4), it also follows that tr_ω is additive on positive operators,

$$\mathrm{tr}_\omega(T_1 + T_2) = \mathrm{tr}_\omega(T_1) + \mathrm{tr}_\omega(T_2) \ , \quad \forall\, T_1, T_2 \ \geq 0 \ , \ T_1, T_2 \in \mathcal{L}^{(1,\infty)} \ .$$

This, together with the fact that $\mathcal{L}^{(1,\infty)}$ is generated by its positive part ([47, 39]), implies that tr_ω extends by linearity to the entire $\mathcal{L}^{(1,\infty)}$ with the property that,

$$\mathrm{tr}_\omega(BT) = \mathrm{tr}_\omega(TB) \ , \ \ \forall\, B \in \mathcal{B}(\mathcal{H}) \ .$$

Now, the Dixmier trace is explicitly computable only for operators for which all values tr_ω coincides. Operators for which this happens are called *measurable* [16, 39]. An operator T for which the sequence $\{\gamma_N\}$ itself converges, that is the ordinary limit (7.3) exists, is indeed mesurable. It is proved in [52] that a positive compact operator $T \in \mathcal{L}^{(1,\infty)}$ is measurable if and only if this ordinary limit exists.

One has that $\mathrm{tr}_\omega(T) = 0$ for an operator T of order higher than 1, . This follows from the fact that the space of all infinitesimals of order higher than 1 forms a two-sided ideal whose elements satisfy the condition $k\mu_k(T) \to 0$ as $k \to \infty$; then the corresponding sequence $\{\gamma_N\}$ converges to zero giving a vanishing Dixmier trace. For a similar reason, trace class operator also satisfy $\mathrm{tr}_\omega(T) = 0$.

Explicit examples of computations of Dixmier traces are in [39, 47].

An important result is the fact that for (a class of) pseudodifferential operators the Dixmier trace coincides with a residue found by Wodzicki. This residue is a unique trace on the algebra of pseudodifferential operators of any order. For operators of order at most $-n$ it coincides with the corresponding Dixmier trace.

Definition 7.2 Let (M, g) be an n-dimensional compact Riemannian manifold. Let T be a pseudodifferential operator of order $-n$ acting on sections of a complex vector bundle $E \to M$. Its Wodzicki residue is defined by

$$\mathrm{Res}_W T := \frac{1}{n(2\pi)^n} \int_{\mathbb{S}^* M} \mathrm{tr}_E \ \sigma_{-n}(T)\, \mathrm{d}\mu \ .$$

Here $\sigma_{-n}(T)$ is the principal symbol of T, a matrix-valued function on T^*M which is homogeneous of degree $-n$ in the fibre coordinates. The trace tr_E is a matrix trace over 'internal indices', and the integral is taken over the unit co-sphere $\mathbb{S}^*M = \{(x, \xi) \in T^*M \mid g^{\mu\nu}\xi_\mu\xi_\nu = 1\} \subset T^*M$, with measure $\mathrm{d}\mu = \mathrm{d}\mu_g(x)\mathrm{d}\xi$. Note that the constant in front of the integral is not 'universally' agreed upon.

Wodzicki has extended the above formula to a unique trace on the algebra of pseudodifferential operators of any order [73, 74] acting on sections of a vector bundle over a compact Riemannian manifold. The trace of any operator T is given by the right-hand side of the same formula, with now $\sigma_{-n}(T)$ the symbol of order $-n$ of T. In particular, one puts $\mathrm{Res}_W T = 0$ if the order of T is less than $-n$. For any pseudodifferential operator of order $\leq -n$, the Wodzicki residue coincides (up to a multiplicative constant) with the Dixmier trace, as shown by the following (Connes' trace theorem).

Theorem 7.1 *Let M be an n dimensional closed Riemannian manifold. Let T be a pseudodifferential operator of order $-n$ acting on sections of a vector bundle $E \to M$. Then,*

(1) *The corresponding operator T on the Hilbert space $\mathcal{H} = L^2(M, E)$ of square integrable sections belongs to $\mathcal{L}^{(1,\infty)}$.*

(2) *The trace $\mathrm{tr}_\omega T$ does not depend on ω (thus T is measurable) and coincides with the residue:*

$$\mathrm{tr}_\omega T = \mathrm{Res}_W T.$$

This result has been given in [14] (see also [39]). In the course of the proof one also establishes that the trace $\mathrm{tr}_\omega T$ depends only on the conformal class of the metric on M.

7.2 Spectral triples

We introduce the concept of spectral triple, the main ingredient in Connes' machinery.

Definition 7.3 A spectral triple $(\mathcal{A}, \mathcal{H}, D)$ is given by a complex unital $*$-algebra $\mathcal{A}$ with a faithfully representation $\pi : \mathcal{A} \to \mathcal{B}$ as bounded operators on the Hilbert space $\mathcal{H}$, together with a self-adjoint operator $D = D^*$ on $\mathcal{H}$ with the following properties.

(1) The resolvent $(D + i)^{-1}$, is a compact operator on $\mathcal{H}$;

(2) $[D, \pi(a)] = D\pi(a) - \pi(a)D \in \mathcal{B}(\mathcal{H})$, for any $a \in \mathcal{A}$.

The triple is said to be even if there is a $\mathbb{Z}_2$ grading of $\mathcal{H}$, namely an operator γ on $\mathcal{H}$, $\gamma = \gamma^*, \gamma^2 = 1$, such that

$$\gamma D + D\gamma = 0 \,, \quad \gamma\pi(a) - \pi(a)\gamma = 0 \,, \quad \forall\, a \in \mathcal{A} \,.$$

If such a grading does not exist, the triple is said to be odd. In this case, for convenience one takes $\gamma = 1$.

In the following, the representation being faithfull, the symbol π will be omitted and $\mathcal{A}$ considered a subalgebra of $\mathcal{B}(\mathcal{H})$; then, its norm closure $\overline{\mathcal{A}}$ is a C^*-algebra.

By the first assumption above, the self-adjoint operator D has a real discrete spectrum made of eigenvalues, with each eigenvalue of finite multiplicity. Furthermore, $|\lambda_k| \to \infty$ as $k \to \infty$. Indeed, since $(D + i)^{-1}$ is compact, it has characteristic values $\mu_k((D + i)^{-1}) \to 0$, from which $|\lambda_k| = \mu_k(|D|) \to \infty$. As we will see, D is a generalization of the Dirac operator on an ordinary spin manifold, and we will symply call it the *Dirac operator*.

For simplicity we shall assume that D is invertible, simple modifications being needed were this not the case. As alluded to above, The polar decomposition $D = |D|F$ yields a Fredholm module $(\mathcal{H}, F)$ over $\mathcal{A}$ with the properties of Sect. 5.2 and defines the fundamental class in the K-homology of $\mathcal{A}$. In order to define the analogue of the measure integral, one needs the additional notion of the dimension of a spectral triple[1].

[1] For an infinite dimensional geometry one needs θ-summability, a spectral triple $(\mathcal{A}, \mathcal{H}, D)$ being θ-summable if $\mathrm{tr}(\mathrm{e}^{-tD^2}) < \infty$ for all $t > 0$.

Definition 7.4 With $0 < n < \infty$, a spectral triple $(\mathcal{A}, \mathcal{H}, D)$ is said to be n^+-*summable* if $|D|^{-1}$ is an infinitesimal (in the sense of Def. 7.1) of order $1/n$ (and this implies that $|D|^{-n}$ is an infinitesimal of order 1, that is $|D|^{-n} \in \mathcal{L}^{(1,\infty)}$). We shall also call such a n the metric dimension of the triple.

Various degrees of regularity of elements of $\mathcal{A}$ are defined using the operator D and its modulus $|D|$ [23]. To the unbounded operator D on $\mathcal{H}$ one associates an unbounded derivation δ on $\mathcal{B}(\mathcal{H})$, defined for all $a \in \mathcal{B}(\mathcal{H})$ by the rule,

$$\delta(a) = [|D|, a] \ ,$$

A spectral triple is called *regular* if the following inclusion holds,

$$\mathcal{A} \cup [D, \mathcal{A}] \subset \bigcap\nolimits_{j \in \mathbb{N}} \mathrm{dom}\, \delta^j \ ,$$

and $\mathrm{OP}^0 := \bigcap_{j \in \mathbb{N}} \mathrm{dom}\, \delta^j$ is referred to as the as the 'smooth domain' of the operator δ. The class Ψ^0 of pseudodifferential operators of order less or equal that zero is defined as the algebra generated by $\bigcup_{k \in \mathbb{N}} \delta^k(\mathcal{A} \cup [D, \mathcal{A}])$. Then, if the triple has finite metric dimension n, the 'zeta-type' function

$$\zeta_a(s) := \mathrm{tr}_{\mathcal{H}}(a|D|^{-s})$$

associated to $a \in \Psi^0$ is defined and holomorphic for $s \in \mathbb{C}$ with $\mathrm{Re}\, s > n$. And for a regular finite-dimensional spectral triple it makes sense the following definition.

Definition 7.5 A spectral triple has *dimension spectrum* Σ iff $\Sigma \subset \mathbb{C}$ is a countable set and, for all $a \in \Psi^0$, the function $\zeta_a(s)$ extends to a meromorphic function on $\mathbb{C}$ with poles in Σ as unique singularities.

If Σ is made only of simple poles the Wodzicki-type residue functional,

$$⨍ T := \mathrm{Res}_{s=0}\, \mathrm{tr}(T|D|^{-s}), \tag{7.5}$$

is tracial on Ψ^0 (see [11, 39, 8]). We also recall the definition of 'smoothing operators':

$$\mathrm{OP}^{-\infty} := \{T \in \mathrm{OP}^0 \mid |D|^k T \in \mathrm{OP}^0 \ \forall\, k \in \mathbb{N}\} \ .$$

The class $\mathrm{OP}^{-\infty}$ is a two-sided $*$-ideal in the $*$-algebra OP^0, is δ-invariant and in the smooth domain of δ. If T is a smoothing operator, $\zeta_T(s)$ is holomorphic on $\mathbb{C}$. Also, the integral (7.5) vanishes if T is a smoothing operator. Thus, elements in $\mathrm{OP}^{-\infty}$ can be neglected when computing the dimension spectrum and residues.

When restricted to elements of $\mathcal{A}$, the integral (7.5) is given by the Dixmier trace,

$$⨍ a = \mathrm{tr}_\omega(a|D|^{-n}) \ , \quad \forall\, a \in \mathcal{A} \ .$$

The role of the operator $|D|^{-n}$ is to bring the bounded operator a into $\mathcal{L}^{(1,\infty)}$ so that the Dixmier trace makes sense; and $|D|^{-n}$ is the analogue of the volume form of the space.

Given a spectral triple $(\mathcal{A}, \mathcal{H}, D)$, there is a natural distance function on the space $\mathcal{S}(\overline{\mathcal{A}})$ of states of the C^*-algebra $\overline{\mathcal{A}}$ (the norm closure of $\mathcal{A}$), which is defined by [15],

$$d(\phi, \chi) := \sup_{a \in \mathcal{A}} \{|\phi(a) - \chi(a)| \mid \|[D, a]\| \leq 1\} \ , \quad \forall\, \phi, \chi \in \mathcal{S}(\overline{\mathcal{A}}) \ . \tag{7.6}$$

This formula has been the starting point for interesting work on compact quantum metric spaces (for additional material see [64]).

7.3 Real structures

There are many interesting examples of spectral triples just satisfying the conditions in Def. 7.3. However, these are also interesting examples for which one has additional properties, for instance a real structure.

A *real structure* is given by an antilinear isometry $J : \mathcal{H} \to \mathcal{H}$ with a list of conditions [17]. This may be thought of as a generalization of the CPT operator (in fact only CP, since we are considering Euclidean signature). Indeed, the canonical triple associated with any (Riemannian spin) manifold in Sect. 7.6 has a canonical real structure in the sense of Def. 7.6 below, the antilinear isometry J being given by

$$J\psi := C\bar{\psi} \,, \quad \forall\, \psi \in \mathcal{H} \,,$$

where C is the charge conjugation operator [7].

Definition 7.6 Let $(\mathcal{A}, \mathcal{H}, D)$ be a spectral triple, with γ the $\mathbb{Z}_2$-grading when n is even. A *real structure of KO-dimension* n is an antilinear isometry $J : \mathcal{H} \to \mathcal{H}$, such that:

$$J^2 = \varepsilon(n)\mathbb{I} \,, \qquad JD = \varepsilon'(n) DJ \,, \qquad J\gamma = (-1)^{n/2}\gamma J \,, \quad \text{if } n \text{ is even},$$

and the mod 8 periodic functions $\varepsilon(n)$ and $\varepsilon'(n)$ given by

$$\begin{aligned} \varepsilon(n) &= (1, 1, -1, -1, -1, -1, 1, 1) \,, \\ \varepsilon'(n) &= (1, -1, 1, 1, 1, -1, 1, 1) \,, \qquad n = 0, 1, \ldots, 7 \,. \end{aligned}$$

Furthermore, the map

$$b \mapsto b^o = Jb^*J^{-1} \,,$$

determines a representation of the opposite algebra $\mathcal{A}^o$ on $\mathcal{H}$ which commutes with $\mathcal{A}$,

$$[a, Jb^*J^{-1}] = 0 \,, \quad \forall\, a, b \in \mathcal{A} \tag{7.7}$$

and the operator D satisfies the order one condition,

$$[[D, a], Jb^*J^{-1}] = 0 \,, \quad \forall\, a, b \in \mathcal{A} \,. \tag{7.8}$$

A map J satisfying condition (7.7) also turns the Hilbert space $\mathcal{H}$ into a bimodule over $\mathcal{A}$, the bimodule structure being given by

$$a\,\xi\,b := aJb^*J^{-1}\,\xi \,, \quad \forall\, a, b \in \mathcal{A} \,.$$

If $a \in \mathcal{A}$ acts on $\mathcal{H}$ by *left* multiplication, then Ja^*J^{-1} is the corresponding *right* multiplication. For commutative algebras, these actions can be identified and one writes $a = Ja^*J^{-1}$. In this case, condition (7.8) reads $[[D, a], b] = 0$ for any $a, b \in \mathcal{A}$, which is just the statement that D is a differential operator of order 1. Thus, the general condition (7.8) may be thought of as the statement that D is a 'generalized differential operator' of order 1. Since a and b^o commute by condition (7.7), condition (7.8) is symmetric, namely it is equivalent to the condition $[[D, b^o], a] = 0$, for any $a, b \in \mathcal{A}$.

Recent interesting examples of spectral triples on spaces coming from quantum groups show that the conditions (7.7) and (7.8) above are too restrictive. It was suggested in [28] that one should modify these in order to obtain a meaningful noncommutative geometry on quantum groups. One needs to replace (7.7) and (7.8) in Def. 7.6 by

$$[a, Jb^*J^{-1}] \in \mathcal{I} \ , \qquad [[D, a], Jb^*J^{-1}] \in \mathcal{I}, \quad \forall\, a, b \in \mathcal{A} \ ,$$

where $\mathcal{I}$ is a suitable operator ideal of infinitesimals (see Def. 7.1). We shall describe these examples in Sect. 9.

In general, a good starting point for a real structure J is the Tomita-Takesaki involution [70]. Let us recall some definitions. If $\mathcal{M}$ is an involutive subalgebra of $\mathcal{B}(\mathcal{H})$, a vector $\xi \in \mathcal{H}$ is called *cyclic* for $\mathcal{M}$ if $\mathcal{M}\xi$ is dense in $\mathcal{H}$; it is called *separating* for $\mathcal{M}$ if for any $T \in \mathcal{M}$, the condition $T\xi = 0$ implies $T = 0$. One finds that a cyclic vector for $\mathcal{M}$ is separating for the commutant

$$\mathcal{M}' := \{T \in \mathcal{B}(\mathcal{H}) \mid Ta = aT \ , \ \ \forall\, a \in \mathcal{M}\} \ .$$

If $\mathcal{M}$ is a von Neumann algebra – that is $\mathcal{M} = \mathcal{M}''$ –, the converse is also true, namely a cyclic vector for $\mathcal{M}'$ is separating for $\mathcal{M}$ [32]. Tomita's theorem then states that for any weakly closed[2] *-algebra of operators $\mathcal{M}$ on a Hilbert space $\mathcal{H}$ which admits a cyclic and separating vector ξ, there exists a canonical antilinear isometric involution $J : \mathcal{H} \to \mathcal{H}$ which leaves ξ invariant, i.e. $J\xi = \xi$, and which conjugates $\mathcal{M}$ to its commutant, i.e. $J\mathcal{M}J^{-1} = \mathcal{M}'$. As a consequence, $\mathcal{M}$ is anti-isomorphic to $\mathcal{M}'$, the anti-isomorphism being given by the map $\mathcal{M} \ni a \mapsto Ja^*J^{-1} \in \mathcal{M}'$. We sketch how this works. The densely defined antilinear operator $a\xi \mapsto a^*\xi$ is closable and its closure S has the polar decomposition $S = J\Delta^{1/2}$, with Δ a positive self-adjoint operator and J is the above antiunitary operator. The operator Δ reduces to the identity when $a\xi \mapsto a^*\xi$ is an isometry; then J is the extension of this map to an antilinear isometry of $\mathcal{H}$. This happens when the state $\mathcal{M} \ni a \mapsto (\xi, a\xi) \in \mathbb{C}$ is tracial.

7.4 Some additional conditions

There are some additional properties for a noncommutative geometry that we briefly list here. A (spin) noncommutative geometry of dimension n is given by a regular spectral triple $(\mathcal{A}, \mathcal{H}, D)$ as in Sect. 7.2 (with a grading γ if the triple is even. There may be a real structure J as in Sect. 7.3. In addition to those given above, they satisfiey the conditions [18]:

a) **(Finiteness).** The space $\mathcal{H}^\infty = \bigcap_k \mathrm{Dom}(D^k)$ is a finitely generated projective left $\mathcal{A}$ module. It follows that the algebra $\mathcal{A}$ is a Fréchet pre-C^*-algebra.

b) **(Orientation).** There exists a Hochschild cycle $c \in Z_n(\mathcal{A}, \mathcal{A} \otimes \mathcal{A}^o)$ in degree n (see Sect. 4.2) which, when represented on $\mathcal{H}$, gives the grading γ of $\mathcal{H}$,

$$\pi_D(c) = \gamma \ . \tag{7.9}$$

[2] We also recall that the sequence $\{T_\lambda\}_{\lambda\in\Lambda}$ is said to converge *weakly to* T, $T_\lambda \to T$, if and only if, for any $\xi, \eta \in \mathcal{H}$, $\langle (T_\lambda - T)\xi, \eta\rangle \to 0$.

Recall that in odd dimensions this means that $\pi_D(c) = 1$. Here $\mathcal{A} \otimes \mathcal{A}^o$ is given the $\mathcal{A}$-bimodule structure $a_1(a \otimes b^o)a_2 := (a_1 a a_2) \otimes b^o$ and any Hochschild k-chain in $Z_k(\mathcal{A}, \mathcal{A} \otimes \mathcal{A}^o)$ is represented on $\mathcal{H}$ by

$$\pi_D((a \otimes b^o)a_1 \otimes \cdots \otimes a_k) := aJb^* J^{-1}[D, a_1] \cdots [D, a_k] .$$

This condition gives an abstract volume form.

c) **(Poincaré duality).** The intersection form

$$K_*(\mathcal{A}) \times K_*(\mathcal{A}) \longrightarrow \mathbb{Z}$$

determined by the Fredholm index maps (7.1) of the operator D and on the K-theory $K_*(\mathcal{A} \otimes \mathcal{A}^o)$ is nondegenerate.

Recall from Ex. 4.2 that Hochschild homology classes of the algebra $C^\infty(M)$ give differential forms on the manifold M; in particular an n-cycle gives a form of top degree. The nondegeneracy of the volume form is the algebraic requirement (7.9). Also, recall that ordinary Poincaré duality for an n dimensional manifold M is the existence of an isomorphism $H^p(M) \simeq H_{n-p}(M)$ (cohomology with homology); this gives a pairing between the cohomology groups $H^p(M)$ and $H^{n-p}(M)$.

The above conditions are satisfied by classical commutative geometry for a smooth Riemannian manifold. Also, they are satisfied by the toric deformation that we shall describe in Sect. 8. But we have already mentioned that some of them – the commutant and first order conditions – need to be modified for quantum groups (see Sect. 9) and are valid only up to compact operators. It is still unclear what are the consequences of thee modifications on the last two conditions, i.e. on the orientation and Poincaré duality conditions. Finally, we mention that the above conditions are for compact geometries. For noncompact geometries, that is for non unital algebras a general strategy is still missing (but see [36] for an interesting example and useful suggestions).

7.5 Differential forms for spectral triples

We shall now describe how to construct a differential algebra of forms out of a spectral triple $(\mathcal{A}, \mathcal{H}, D)$. Recall the universal calculus $(\Omega\mathcal{A}, \delta)$ of Sect. 4.1. The map

$$\pi_D : \Omega\mathcal{A} \longrightarrow \mathcal{B}(\mathcal{H}) ,$$
$$\pi_D(a_0 \delta a_1 \cdots \delta a_p) := a_0[D, a_1] \cdots [D, a_p] , \quad a_j \in \mathcal{A} ,$$

is a homomorphism since both δ and $[D, \cdot]$ are derivations on $\mathcal{A}$. Furthermore, with $\delta(a^*) = -\delta(a)$, from $[D, a]^* = -[D, a^*]$ one has also that $\pi_D(\omega)^* = \pi_D(\omega^*)$ for any form $\omega \in \Omega\mathcal{A}$ and π_D is a *-representation of $\Omega\mathcal{A}$ as bounded operators on $\mathcal{H}$.

It is not possible to define the space of forms as the image $\pi_D(\Omega\mathcal{A})$, since in general $\pi_D(\omega) = 0$ does not imply that $\pi_D(\delta\omega) = 0$. Such forms ω, for which $\pi_D(\omega) = 0$ while $\pi_D(\delta\omega) \neq 0$, are called *junk forms*. They have to be disposed of in order to construct a true differential algebra and make π_D into a homomorphism of differential algebras. If $J_0 := \bigoplus_p J_0^p$ is the graded two-sided ideal of $\Omega\mathcal{A}$ given by

$$J_0^p := \{\omega \in \Omega^p\mathcal{A} \mid \pi_D(\omega) = 0\} ,$$

then $J = J_0 + \delta J_0$ is a graded differential two-sided ideal of $\Omega\mathcal{A}$, and the following makes sense.

Definition 7.7 The graded differential algebra of Connes' forms over the algebra $\mathcal{A}$ is

$$\Omega_D\mathcal{A} := \Omega\mathcal{A}/J \simeq \pi_D(\Omega\mathcal{A})/\pi_D(\delta J_0) .$$

This algebra is naturally graded by the degrees of $\Omega\mathcal{A}$ and J, the space of p-forms being

$$\Omega_D^p\mathcal{A} = \Omega^p\mathcal{A}/J^p .$$

Since J is a differential ideal, the exterior differential δ defines a differential on $\Omega_D\mathcal{A}$,

$$\mathrm{d} : \Omega_D^p\mathcal{A} \longrightarrow \Omega_D^{p+1}\mathcal{A} , \quad \mathrm{d}[\omega] := [\delta\omega] \simeq [\pi_D(\delta\omega)] ,$$

with $[\omega]$ the class in $\Omega_D^p\mathcal{A}$ of $\omega \in \Omega^p\mathcal{A}$. Explicitly, the $\mathcal{A}$-bimodule $\Omega_D^p\mathcal{A}$ of p-forms is

$$\Omega_D^p\mathcal{A} \cong \pi_D(\Omega^p\mathcal{A})/\pi_D(\delta(J_0 \cap \Omega^{p-1}\mathcal{A})) ,$$

and is made of classes of operators of the form

$$\omega_p = \sum_j a_0^j[D, a_1^j][D, a_2^j] \cdots [D, a_p^j] , \quad a_i^j \in \mathcal{A} ,$$

modulo the sub-bimodule of operators

$$\left\{ \sum_j [D, b_0^j][D, b_1^j] \cdots [D, b_{p-1}^j] \mid b_i^j \in \mathcal{A} , \sum_j b_0^j[D, b_1^j] \cdots [D, b_{p-1}^j] = 0 \right\} .$$

On them, the exterior differential is given by

$$\mathrm{d}\left[\sum_j a_0^j[D, a_1^j] \cdots [D, a_p^j] \right] = \left[\sum_j [D, a_0^j][D, a_1^j] \cdots [D, a_p^j] \right] .$$

7.6 The canonical triple over a manifold

The basic example of spectral triple is the *canonical triple* on a closed n-dimensional Riemannian spin manifold (M, g), with g denoting a Riemannian metric. We recall that a spin manifold is a manifold on which it is possible to construct principal bundles having the groups $Spin(n)$ as structure groups. Then one defines spinor fields (i.e. fields which describe fermions) as sections of suitable associated bundles over M. There are topological obstructions to the existence of spin structure. A manifold admits a spin structure if and only if its second Stiefel-Whitney class vanishes [50].

The canonical spectral triple $(\mathcal{A}, \mathcal{H}, D)$ over the closed n-dimensional manifold M is:

(1) $\mathcal{A} = C^\infty(M)$ is the algebra of $\mathbb{C}$-valued smooth functions on M.

(2) $\mathcal{H} = L^2(M, S)$ is the Hilbert space of square integrable sections of the irreducible spinor bundle $S \to M$. The rank of the bundle is $2^{[n/2]}$ with $[k]$ indicating the

integer part in k and the fibre at the point $x \in M$ is $S_x \simeq \mathbb{C}^{2^{[n/2]}}$. The Hilbert space $L^2(M, S)$ is the completion of the dense $\mathcal{A}$-module $\mathcal{S} = \Gamma^\infty(M, S)$ of spinor fields with respect to the scalar product defined by the measure $\mathrm{d}\mu_g$ of the metric g,

$$(\psi, \phi) = \int \mathrm{d}\mu_g \, \psi \cdot \phi \, ,$$

with $(\psi \cdot \phi)(x) = \sum_j \overline{\psi}_j(x)\phi_j(x)$ the natural scalar product in the spinor space S_x at the point $x \in M$.

(3) D is the Dirac operator associated with the Levi-Civita connection $\omega = \mathrm{d}x^\mu \omega_\mu$ of the metric g.

We can assume that $\ker D$ is trivial. If not, since M compact, $\ker D$ is finite dimensional and can thus be subtracted from the spinor space.

Elements of the algebra $\mathcal{A}$ act as multiplicative operators on $\mathcal{H}$,

$$(f\psi)(x) := f(x)\psi(x) \, , \ \forall \, f \in \mathcal{A} \, , \ \psi \in \mathcal{H} \, . \tag{7.10}$$

Let $(e_a, a = 1, \dots, n)$ be an orthonormal basis of vector fields; it is related to the natural basis $(\partial_\mu, \mu = 1, \dots, n)$ via the n-beins, the latter having components e^μ_a. The components $\{g^{\mu\nu}\}$ and $\{\eta^{ab}\}$ of the tangent and of the frame bundle metrics are related by

$$g^{\mu\nu} = e^\mu_a e^\nu_b \eta^{ab} \, , \quad \eta_{ab} = e^\mu_a e^\nu_b g_{\mu\nu} \, .$$

Indices $\{\mu\}$ and $\{a\}$ will be lowered and raised by the metric g and η respectively. As usual we sum over repeated indices.

The coefficients $(\omega_{\mu a}^{\ \ b})$ of the Levi-Civita (that is, metric and torsion-free) connection of the metric g, defined by $\nabla_\mu e_a = \omega_{\mu a}^{\ \ b} e_b$, are the solutions of the equations

$$\partial_\mu e^a_\nu - \partial_\nu e^a_\mu - \omega_{\mu b}^{\ \ a} e^b_\nu + \omega_{\nu b}^{\ \ a} e^b_\mu = 0 \, .$$

Let $\mathrm{Cl}(M)$ be the Clifford bundle over M whose fibre at $x \in M$ is just the complexified Clifford algebra $\mathrm{Cl}_{\mathbb{C}}(T^*_x M)$, and let $\Gamma(M, \mathrm{Cl}(M))$ be the module of corresponding sections. There is an algebra morphism

$$\gamma : \Gamma(M, \mathrm{Cl}(M)) \to \mathcal{B}(\mathcal{H}) \, , \qquad \gamma(\mathrm{d}x^\mu) := \gamma^\mu(x) = \gamma^a e^\mu_a \, , \quad \mu = 1, \dots, n \, ,$$

and extended as an $\mathcal{A}$-linear algebra map. The gamma matrices $\{\gamma^\mu(x)\}$ and $\{\gamma^a\}$, taken to be Hermitian, satisfy

$$\begin{aligned} \gamma^\mu(x)\gamma^\nu(x) + \gamma^\nu(x)\gamma^\mu(x) &= 2g(\mathrm{d}x^\mu, \mathrm{d}x^\nu) = 2g^{\mu\nu} \, , \quad \mu, \nu = 1, \dots, n \, ; \\ \gamma^a\gamma^b + \gamma^b\gamma^a &= 2\eta^{ab} \, , \quad a, b = 1, \dots, n \, . \end{aligned}$$

The lift ∇^S of the Levi-Civita connection to the bundle of spinors is then

$$\nabla^S_\mu = \partial_\mu + \omega^S_\mu = \partial_\mu + \frac{1}{4}\omega_{\mu ab}\gamma^a\gamma^b \, .$$

The Dirac operator, defined by $D = -\mathrm{i}\gamma \circ \nabla^S$, can be written locally as

$$D = -\mathrm{i}\gamma(\mathrm{d}x^\mu)\nabla^S_\mu = -\mathrm{i}\gamma^\mu(x)(\partial_\mu + \omega^S_\mu) = -\mathrm{i}\gamma^a e^\mu_a(\partial_\mu + \omega^S_\mu) \, . \tag{7.11}$$

From the action (7.10) of $\mathcal{A}$ as multiplicative operators, one finds that

$$[D, f]\psi = -\mathrm{i}(\gamma^\mu \partial_\mu f)\psi , \quad \forall\, f \in \mathcal{A} ,$$

and the commutator $[D, f]$ is a multiplicative operator as well and is then bounded.

The factor $(-\mathrm{i})$ in the definition (7.11) is introduced to make the Dirac operator D a *symmetric* (instead of skew-symmetric) operator on the dense domain $\mathcal{S}$, that is,

$$(D\psi, \phi) = (\psi, D\phi) , \qquad \forall\, \psi, \phi \in \mathcal{S} .$$

The operator D is in fact *essentially self-adjoint* on $\mathcal{S}$, as first proved in [75] (see also [35]). For a densely defined operator T on a Hilbert space $\mathcal{H}$, its *adjoint* T^* has domain

$$\mathrm{Dom}(T^*) := \{\psi \in \mathcal{H} \,:\, \exists \chi \in \mathcal{H} \text{ with } (T\psi, \phi) = (\psi, \chi) \;\; \forall\, \psi \in \mathrm{Dom}(T)\} .$$

Then $T^*\phi = \chi$ or $(T\psi, \phi) = (\psi, T^*\phi)$. If T is symmetric, then T^* is an extension of T, that is $\mathrm{Dom}(T) \subseteq \mathrm{Dom}(T^*)$ with $T^* = T$ on $\mathrm{Dom}(T)$. The *closure* of T is the second adjoint $\overline{T} := T^{**}$ and its domain is determined by the fact that the graph of $\overline{T}$ is the closure of the graph of T. Symmetric operators always admit a closure with $\mathrm{Dom}(T) \subseteq \mathrm{Dom}(\overline{T}) \subseteq \mathrm{Dom}(T^*)$. One says that T is *self-adjoint* if $T = T^*$, that is T is symmetric and closed; one says that T is *essentially self-adjoint* if it is symmetric and its closure $\overline{T}$ is self-adjoint. A merely symmetric unbounded operator may have non real elements in its spectrum; a self-adjoint operator has a real spectrum, $\mathrm{sp}(T) \in \mathbb{R}$ instead.

Finally, we mention the Lichnérowicz formula [2] for the square of D,

$$D^2 = \Delta^S + \frac{1}{4} R ,$$

where R is the scalar curvature of the metric and Δ^S is the Laplacian operator lifted to the bundle of spinors,

$$\Delta^S = -g^{\mu\nu}(\nabla^S_\mu \nabla^S_\nu - \Gamma^\rho_{\mu\nu} \nabla^S_\rho) ,$$

with $\Gamma^\rho_{\mu\nu}$ the Christoffel symbols of the connection.

If the dimension n of M is even, the previous spectral triple is even. For the grading operator one just takes the product of all flat gamma matrices,

$$\gamma = \gamma^{n+1} = \mathrm{i}^{-n/2} \gamma^1 \cdots \gamma^n ,$$

which, when n is even, anticommutes with the Dirac operator, $\gamma D + D\gamma = 0$. Furthermore, the factor $\mathrm{i}^{n/2}$ ensures that $\gamma^2 = \mathbb{I}$, $\gamma^* = \gamma$.

Proposition 7.1 *Let $(\mathcal{A}, \mathcal{H}, D)$ be the canonical triple over the manifold (M, g). Then,*

(1) *The space M is the structure space of the algebra $\overline{\mathcal{A}}$ of continuous functions on M.*

(2) *The geodesic distance between any two points on M is given by*

$$d(p, q) = \sup_{f \in \mathcal{A}} \{ |f(p) - f(q)| \mid \|[D, f]\| \leq 1 \} , \;\; \forall\, p, q \in M .$$

(3) *The Riemannian measure on M is given by*

$$\int_M \mathrm{d}\mu_g f = \frac{n(2\pi)^n}{2^{[n/2]}\Omega_n}\ \mathrm{tr}_\omega(f|D|^{-n})\ ,\ \ \forall\, f \in \mathcal{A}\,,$$

where $\Omega_n = 2\pi^{n/2}/\Gamma(\frac{n}{2})$ is the usual volume of the sphere $\mathbb{S}^{n-1}$.

For a full proof we refer to [16]. We only mention that point (3). above is a simple consequence of Thm. 7.1: Since any function f acts as a multiplicative operator, the operator $f|D|^{-n}$ is pseudodifferential of order $-n$. On the co-sphere bundle, its principal symbol, $\sigma_{-n}(x,\xi) = f(x)\|\gamma^\mu \xi_\mu\|^{-n}$, reduces to the matrix $f(x)\mathbb{I}_{2^{[n/2]}}$, with $2^{[n/2]}$ the rank of the spinor bundle. From Thm. 7.1 and Def. 7.2 if follows that

$$\mathrm{tr}_\omega(f|D|^{-n}) = \frac{2^{[n/2]}}{n(2\pi)^n}\int_{\mathbb{S}^{n-1}} \mathrm{d}\xi \int_M \mathrm{d}\mu_g f = \frac{2^{[n/2]}\Omega_n}{n(2\pi)^n}\int_M \mathrm{d}\mu_g f\,.$$

The metric dimension of the canonical spectral triple over the manifold M coincides with the dimension of M. Indeed, the Weyl formula [37] for the eigenvalues of $|D|$ gives

$$\mu_k(|D|)\ \sim\ 2\pi\left(\frac{n}{\Omega_n \mathrm{vol}(M)}\right)^{1/n} k^{1/n} \qquad \text{as} \quad k \to \infty\,.$$

Here $\mathrm{vol}(M) = \int_M \mathrm{d}\mu_g$ is the volume of the manifold M. In turn, this gives for the operator $|D|^n$ a linear growth: $\mu_k(|D|^n) \sim k$ and $|D|^{-n}$ is an infinitesimal of order 1. Finally, the dimension spectrum for the canonical spectral triple is the set $\{0, 1, \ldots, n\}$ and the corresponding singularities are simple. Multiplicities occur for singular manifolds [23].

One does not need spinor structures to recover spectral properties like the dimension since for this the Laplacian operator suffices. For the closed Riemannian manifold (M, g) the Laplacian operator is written is local coordinates as

$$\Delta = -g^{\mu\nu}(\partial_\mu\partial_\nu - \Gamma^\rho_{\mu\nu}\partial_\rho)\ ,$$

and $\Gamma^\rho_{\mu\nu}$ are the Christoffel symbols of the Levi-Civita connection. The operator Δ extends to a positive self-adjoint operator on the Hilbert space $L^2(M, \mathrm{d}\mu_g)$ and $(\Delta + 1)^{-1}$ is compact. The counting function for its spectrum is

$$N_\Delta(\lambda) := \#\{\lambda_k(\Delta) : \lambda_k(\Delta) \leq \lambda\}\,,$$

where the eigenvalues $\lambda_k(\Delta)$ are counted with multiplicity and an eigenvalue of multiplicity m appears m times in the counting function. Weyl theorem gives an asymptotic estimate for the counting function,

$$N_\Delta(\lambda) \sim \frac{\Omega_n \mathrm{vol}(M)}{n(2\pi)^n}\,\lambda^{n/2} \qquad \text{as} \quad \lambda \to \infty\,.$$

Finally, we mention that the construction of differential forms given in Sect. 7.5, when applied to the canonical triple over an ordinary manifold M, reproduces the usual exterior algebra over M: for more details we refer to [16, 47, 39].

7.7 Reconstructing commutative geometries

The previous Section shows how the usual Riemannian spin geometry is recovered when the algebra $\mathcal{A}$ the algebra of smooth functions on a manifold, $\mathcal{H}$ is Hilbert space of spinors, and D the Dirac operator for the spin structure the Riemannian metric. The question of reconstruction is whether the operator-theoretic framework described above could determine a manifold structure whenever the algebra $\mathcal{A}$ of the triple is commutative.This programme, started in [18] has proved quite formidable and has been completed after 10 years only recently in [59]. In the latter paper, using a slightly stronger set of conditions on a spectral triple, with the additional assumption of a commutative algebra $\mathcal{A}$ one recovers a closed manifold whose algebra of smooth functions coincides with $\mathcal{A}$.

8 Toric noncommutative manifolds

We briefly recal the general construction of toric noncommutative manifolds given in [21] where they were called isospectral deformations. These are deformations of a classical Riemannian manifold and satisfy all the properties of a noncommutative spin geometry. They are the content of the following result taken from [21],

Theorem 8.1 *Let M be a compact spin Riemannian manifold whose isometry group has rank $r \geq 2$. Then M admits a natural one parameter isospectral deformation to noncommutative geometries M_θ.*

The idea of the construction is to deform the standard spectral triple describing the Riemannian geometry of M along a torus embedded in the isometry group, thus obtaining a family of spectral triples describing noncommutative geometries. On this class of noncommutative manifolds, gauge theories have been constructed in [48, 49].

8.1 Deforming a torus action

Let M be an m dimensional compact Riemannian manifold equipped with an isometric smooth action σ of an n-torus $\mathbb{T}^n$, $n \geq 2$. We denote by σ also the corresponding action of $\mathbb{T}^n$ by automorphisms – obtained by pull-backs – on the algebra $C^\infty(M)$ of smooth functions on M.

The algebra $C^\infty(M)$ may be decomposed into spectral subspaces which are indexed by the dual group $\mathbb{Z}^n = \widehat{\mathbb{T}}^n$. With $s = (s_1, \cdots, s_n) \in \mathbb{T}^n$, each $r \in \mathbb{Z}^n$ yields a character of $\mathbb{T}^n$, $e^{2\pi i s} \mapsto e^{2\pi i r \cdot s}$, with the scalar product $r \cdot s := r_1 s_1 + \cdots + r_n s_n$. The r-th spectral subspace for the action σ of $\mathbb{T}^n$ on $C^\infty(M)$ consists of those smooth functions f_r for which

$$\sigma_s(f_r) = e^{2\pi i r \cdot s} f_r, \tag{8.1}$$

and each $f \in C^\infty(M)$ is the sum of a unique series $f = \sum_{r \in \mathbb{Z}^n} f_r$, which is rapidly convergent in the Fréchet topology of $C^\infty(M)$ [63]. Let now $\theta = (\theta_{jk} = -\theta_{kj})$ be a real antisymmetric $n \times n$ matrix. The θ-deformation of $C^\infty(M)$ may be defined by replacing the ordinary product by a deformed product, given on spectral subspaces by

$$f_r \times_\theta g_{r'} := f_r \, \sigma_{\frac{1}{2} r \cdot \theta}(g_{r'}) = e^{\pi i r \cdot \theta \cdot r'} f_r g_{r'}, \tag{8.2}$$

where $r \cdot \theta$ is the element in $\mathbb{R}^n$ with components $(r \cdot \theta)_k = \sum r_j \theta_{jk}$ for $k = 1, \ldots, n$. The product in (8.2) is then extended linearly to all functions in $C^\infty(M)$. We denote the space $C^\infty(M)$ endowed with the product $\times_\theta$ by $C^\infty(M_\theta)$. The action σ of $\mathbb{T}^n$ on $C^\infty(M)$ extends to an action on $C^\infty(M_\theta)$ given again by (8.1) on the homogeneous elements.

Next, let M be a spin manifold with $\mathcal{H} := L^2(M, \mathcal{S})$ the Hilbert space of spinors and D the usual Dirac operator of the metric of M. We know from Sect. 7.6 that functions act on spinors by pointwise multiplication thus giving a representation $\pi : C^\infty(M) \to \mathcal{B}(\mathcal{H})$.

There is a double cover $c : \tilde{\mathbb{T}}^n \to \mathbb{T}^n$ and a representation of $\tilde{\mathbb{T}}^n$ on $\mathcal{H}$ by unitary operators $U(s)$, $s \in \tilde{\mathbb{T}}^n$, for which $U(s)\pi(f)U(s)^{-1} = \pi(\sigma_{c(s)}(f))$ for all $f \in C^\infty(M)$ and – since the torus action is assumed to be isometric – such that

$$U(s)DU(s)^{-1} = D,$$

From its very definition, α_s coincides on $\pi(C^\infty(M)) \subset \mathcal{B}(\mathcal{H})$ with the automorphism $\sigma_{c(s)}$. Recall that an element $T \in \mathcal{B}(\mathcal{H})$ is called smooth for the action of $\tilde{\mathbb{T}}^n$ if the map

$$\tilde{\mathbb{T}}^n \ni s \mapsto \alpha_s(T) := U(s)TU(s)^{-1},$$

is smooth for the norm topology. Much as it was done before for the smooth functions, we shall use the torus action to give a spectral decomposition of smooth elements of $\mathcal{B}(\mathcal{H})$. Any such a smooth element T is written as a (rapidly convergent) series $T = \sum T_r$ with $r \in \mathbb{Z}^n$ and each T_r is homogeneous of degree r under the action of $\tilde{\mathbb{T}}^n$, i.e.

$$\alpha_s(T_r) = e^{2\pi i r \cdot s} T_r, \quad \forall\, s \in \tilde{\mathbb{T}}^n.$$

Let $(P_1, P_2, \ldots, P_n)$ be the infinitesimal generators of the action of $\tilde{\mathbb{T}}^n$ so that we can write $U(s) = \exp 2\pi i s \cdot P$. Now, with θ a real $n \times n$ anti-symmetric matrix as above, one defines a twisted representation of the smooth elements of $\mathcal{B}(\mathcal{H})$ on $\mathcal{H}$ by

$$L_\theta(T) := \sum_r T_r U(\tfrac{1}{2} r \cdot \theta) = \sum_r T_r \exp\left\{\pi i\, r_j \theta_{jk} P_k\right\},$$

Taking smooth functions on M as elements of $\mathcal{B}(\mathcal{H})$ – via the representation π – the previous definition gives an algebra $L_\theta(C^\infty(M))$ which we may think of as a representation (as bounded operators on $\mathcal{H}$) of the algebra $C^\infty(M_\theta)$. Indeed, by the very definition of the product $\times_\theta$ in (8.2) one establishes that

$$L_\theta(f \times_\theta g) = L_\theta(f) L_\theta(g),$$

proving that the algebra $C^\infty(M)$ equipped with the product $\times_\theta$ is isomorphic to the algebra $L_\theta(C^\infty(M))$. It is shown in [63] that there is a natural completion of the algebra $C^\infty(M_\theta)$ to a C^*-algebra $C(M_\theta)$ whose smooth subalgebra – under the extended action of $\mathbb{T}^n$ – is precisely $C^\infty(M_\theta)$. Thus, we can understand L_θ as a *quantization map* from

$$L_\theta : C^\infty(M) \to C^\infty(M_\theta),$$

which provides a strict deformation quantization in the sense of Rieffel. More generally, in [63] one considers a (not necessarily commutative) C^*-algebra A carrying an action of $\mathbb{R}^n$. For an anti-symmetric $n \times n$ matrix θ, one defines a star product $\times_\theta$ between elements in A much as we did before. The algebra A equipped with the product $\times_\theta$ gives rise to a C^*-algebra denoted by A_θ. Then the collection $\{A_{\hbar\theta}\}_{\hbar\in[0,1]}$ is a continuous family of C^*-algebras providing a strict deformation quantization in the direction of the Poisson structure on the algebra A defined by the matrix θ.

Our case corresponds to the choice $A = C(M)$ with an action of $\mathbb{R}^n$ that is periodic or, in other words, an action of $\mathbb{T}^n$. The smooth elements in the deformed algebra make up the algebra $C^\infty(M_\theta)$. The quantization map will play a key role in what follows, allowing us to extend differential geometric techniques from M to the noncommutative space M_θ.

It was shown in [21] that the datum $(L_\theta(C^\infty(M)), \mathcal{H}, D)$ satisfies all properties of a noncommutative spin geometry as listed in Sect. 7; there is also a grading γ (for the even case) and a real structure J. In particular, boundedness of the commutators $[D, L_\theta(f)]$ for $f \in C^\infty(M)$ follows from $[D, L_\theta(f)] = L_\theta([D, f])$, D being of degree 0 (since $\mathbb{T}^n$ acts by isometries, each P_k commutes with D). This noncommutative geometry is an isospectral deformation of the classical Riemannian geometry of M, in that the spectrum of the operator D coincides with that of the Dirac operator D on M. Thus all spectral properties are unchanged. In particular, the triple is m^+-summable and there is a noncommutative integral as a Dixmier trace [33],

$$\int\!\!\!\!\!\!- L_\theta(f) := \mathrm{Tr}_\omega\big(L_\theta(f)|D|^{-m}\big),$$

with $f \in C^\infty(M_\theta)$ understood in its representation on $\mathcal{H}$.

8.2 The manifold M_θ as a fixed point algebra

A different but equivalent approach to these noncommutative manifolds M_θ was introduced in [20]. In there the algebra $C^\infty(M_\theta)$ is identified as a fixed point subalgebra of $C^\infty(M) \otimes C^\infty(\mathbb{T}^n_\theta)$ where $C^\infty(\mathbb{T}^n_\theta)$ is the algebra of smooth functions on the noncommutative torus. This identification was shown to be useful in extending techniques from commutative differential geometry on M to the noncommutative space M_θ.

We recall the definition of the noncommutative n-torus $\mathbb{T}^n_\theta$ (see for instance [62]). Let $\theta = (\theta_{jk} = -\theta_{kj})$ be a real $n \times n$ anti-symmetric matrix as before, and let $\lambda^{jk} = e^{2\pi i\theta_{jk}}$. The unital $*$-algebra $\mathcal{A}(\mathbb{T}^n_\theta)$ of polynomial functions on $\mathbb{T}^n_\theta$ is generated by n unitary elements U^k, $k = 1, \ldots, n$, with relations

$$U^j U^k = \lambda^{jk} U^k U^j, \quad j, k = 1, \ldots, n.$$

The polynomial algebra is extended to the universal C^*-algebra with the same generators. There is a natural action of $\mathbb{T}^n$ on $\mathcal{A}(\mathbb{T}^n_\theta)$ by $*$-automorphisms given by $\tau_s(U^k) = e^{2\pi i s_k} U^k$ with $s = (s_k) \in \mathbb{T}^n$. The corresponding infinitesimal generators X_k of the action are algebra derivations given explicitly on the generators by $X_k(U^j) = 2\pi i\delta^j_k$. They are used [12] to construct the pre-C^*-algebra $C^\infty(\mathbb{T}^n_\theta)$ of smooth functions on $\mathbb{T}^n_\theta$, the completion of $\mathcal{A}(\mathbb{T}^n_\theta)$ with respect to the locally convex topology generated by the seminorms,

$$|u|_r := \sup_{r_1+\cdots+r_n\le r} \|X_1^{r_1} \cdots X_n^{r_n}(u)\|,$$

and $\|\cdot\|$ is the C^*-norm. The algebra $C^\infty(\mathbb{T}^n_\theta)$ turns out to be a nuclear Fréchet space and one can unambiguously take the completed tensor product $C^\infty(M)\overline{\otimes}C^\infty(\mathbb{T}^n_\theta)$. Then, one defines $\big(C^\infty(M)\overline{\otimes}C^\infty(\mathbb{T}^n_\theta)\big)^{\sigma\otimes\tau^{-1}}$ as the fixed point subalgebra of $C^\infty(M)\overline{\otimes}C^\infty(\mathbb{T}^n_\theta)$ consisting of elements a in the tensor product that are invariant under the diagonal action of $\mathbb{T}^n$, i.e. such that $\sigma_s\otimes\tau_{-s}(a)=a$ for all $s\in\mathbb{T}^n$. The noncommutative manifold M_θ is defined by "duality" by setting for its functions,

$$C^\infty(M_\theta):=\big(C^\infty(M)\overline{\otimes}C^\infty(\mathbb{T}^n_\theta)\big)^{\sigma\otimes\tau^{-1}}.$$

As the notation suggests, the algebra $C^\infty(M_\theta)$ is isomorphic to the algebra $L_\theta(C^\infty(M))$ defined in the previous section.

Next, let $\mathcal{S}$ be a spin bundle over M and D the Dirac operator on $\Gamma^\infty(M,\mathcal{S})$, the $C^\infty(M)$-module of smooth sections of $\mathcal{S}$. The action of the group $\mathbb{T}^n$ on M does not lift directly to the spinor bundle. Rather, there is a double cover $c:\widetilde{\mathbb{T}}^n\to\mathbb{T}^n$ and a group homomorphism $\tilde{s}\to V_{\tilde{s}}$ of $\widetilde{\mathbb{T}}^n$ into $\mathrm{Aut}(\mathcal{S})$ covering the action of $\mathbb{T}^n$ on M,

$$V_{\tilde{s}}(f\psi)=\sigma_{c(s)}(f)V_{\tilde{s}}(\psi),$$

for $f\in C^\infty(M)$ and $\psi\in\Gamma^\infty(M,\mathcal{S})$. According to [20], the proper notion of smooth sections $\Gamma^\infty(M_\theta,\mathcal{S})$ of a spinor bundle on M_θ are elements of $\Gamma^\infty(M,\mathcal{S})\widehat{\otimes}C^\infty(\mathbb{T}^n_{\theta/2})$ which are invariant under the diagonal action $V\times\tilde{\tau}^{-1}$ of $\tilde{\mathbb{T}}^n$. Here $\tilde{s}\mapsto\tilde{\tau}_{\tilde{s}}$ is the canonical action of $\tilde{\mathbb{T}}^n$ on $\mathcal{A}(\mathbb{T}^n_{\theta/2})$. Since the Dirac operator D commutes with $V_{\tilde{s}}$ (remember that the torus action is isometric) one can restrict $D\otimes\mathbb{I}$ to the fixed point elements $\Gamma^\infty(M_\theta,\mathcal{S})$.

Then, let $L^2(M,\mathcal{S})$ be the space of square integrable spinors on M and let $L^2(\mathbb{T}^n_{\theta/2})$ be the completion of $C^\infty(\mathbb{T}^n_{\theta/2})$ in the norm $a\mapsto\|a\|=\mathrm{tr}(a^*a)^{1/2}$, with tr the usual trace on $C^\infty(\mathbb{T}^n_{\theta/2})$. The diagonal action $V\times\tilde{\tau}^{-1}$ of $\tilde{\mathbb{T}}^n$ extends to $L^2(M,\mathcal{S})\otimes L^2(\mathbb{T}^n_{\theta/2})$ (where it becomes $U\times\tau$) and one defines $L^2(M_\theta,\mathcal{S})$ to be the fixed point Hilbert subspace. If D also denotes the closure of the Dirac operator on $L^2(M,\mathcal{S})$, one still denotes by D the operator $D\otimes\mathbb{I}$ on $L^2(M,\mathcal{S})\otimes L^2(\mathbb{T}^n_{\theta/2})$ when restricted to $L^2(M_\theta,\mathcal{S})$. The triple $(C^\infty(M_\theta),L^2(M_\theta,\mathcal{S}),D)$ is an m^+-summable noncommutative spin geometry.

9 The spectral geometry of the quantum group $SU_q(2)$

As mentioned in the Introduction, there are by now several examples of noncommutative geometries on spaces coming from quantum groups. These include quantum two spheres [31, 56, 66, 28, 25, 27], the quantum group $SU_q(2)$ [10, 19, 29, 30], the quantum flag manifolds [46] and a quantum Euclidead four-sphere [26]. To illustrate these examples, in this Section we shall briefly describe the isospectral spectral triple $(\mathcal{A}(SU_q(2)),\mathcal{H},D)$ for the quantum group $SU_q(2)$ given in [29].

The possibility of such a geometry was suggested in [21]. After some initial skepticism [38], the programme was completed in [29] with the construction of a 3-dimensional noncommutative geometry on the manifold of $SU_q(2)$. The spectrum of the operator D is the same as that of the Dirac operator on the 3-sphere $\mathbb{S}^3\simeq SU(2)$ with it rotation-invariant metric. In this sense the deformation, from $SU(2)$ to $SU_q(2)$, is isospectral and in particular the metric dimension of the spectral geometry is 3. The spectral triple is equivariant with respect to the symmetry algebra $\mathcal{U}_q(su(2)) \otimes \mathcal{U}_q(su(2))$ of the quantum group manifold $SU_q(2)$, implemented as a pair of commuting left and right actions of $\mathcal{U}_q(su(2))$ on the algebra $\mathcal{A} = \mathcal{A}(SU_q(2))$. The equivariance allows us to compute the spin representation of the algebra $\mathcal{A}$ and selects a class of possible Dirac operators D.

An equivariant real structure J is constructed by suitably lifting to the Hilbert space of spinors $\mathcal{H}$ the Tomita conjugation operator for the left regular representation of $\mathcal{A}$. However, unlike the Tomita operator for the spin representation, this J does not intertwine the spin representation of $\mathcal{A}$ with its commutant. It is thus incompatible with the full set of requirements for a real spectral triple as given in Sect. 7.3. It turns out that this commutant property, and the companion "first-order" property of D, still hold up to infinitesimals of arbitrarily high order.

9.1 The algebras of functions and of symmetries

We start with some algebraic preliminaries on the algebras of functions $\mathcal{A} = \mathcal{A}(SU_q(2))$ and of infinitesimal symmetries $\mathcal{U}_q(su(2))$.

Definition 9.1 For q real, $0 < q < 1$, we denote by $\mathcal{A} = \mathcal{A}(SU_q(2))$ the $*$-algebra generated by a and b, subject to the commutation rules

$$
\begin{aligned}
&ba = qab, \qquad b^*a = qab^*, \qquad bb^* = b^*b, \\
&a^*a + q^2 b^*b = 1, \qquad aa^* + bb^* = 1.
\end{aligned}
$$

At $q = 1$ this is just the algebra of polynomial functions on the manifold of the group $G = SU(2)$. The group structure is dualized into the one of a Hopf algebra structure for the algebra of representable functions on the group. And this is generalized to the notion of a *quantum groups*, that is a Hopf structure algebra on the algebra $\mathcal{A}(SU_q(2))$.

Definition 9.2 The algebra $\mathcal{A}(SU_q(2))$ comes with a Hopf $*$-algebra structure, with

coproduct: $\Delta a := a \otimes a - q\, b \otimes b^*, \quad \Delta b := b \otimes a^* + a \otimes b\,;$

counit: $\varepsilon(a) = 1, \varepsilon(b) = 0\,;$

antipode: $Sa = a^*$, $Sb = -qb$, $Sb^* = -q^{-1}b^*$, $Sa^* = a\,.$

Definition 9.3 The Hopf $*$-algebra $\mathcal{U} = \mathcal{U}_q(su(2))$ is generated as an algebra by elements e, f, k, with k invertible, satisfying the relations

$$
ek = qke, \qquad kf = qfk, \qquad k^2 - k^{-2} = (q - q^{-1})(fe - ef).
$$

and involution: $k^* = k, f^* = e, e^* = f$. The coproduct Δ is given by

$$
\Delta k = k \otimes k, \qquad \Delta e = e \otimes k + k^{-1} \otimes e, \qquad \Delta f = f \otimes k + k^{-1} \otimes f,
$$

while its counit ϵ and antipode S are given respectively by

$$
\begin{aligned}
&\epsilon(k) = 1, \qquad && \epsilon(f) = 0, \qquad && \epsilon(e) = 0, \\
&Sk = k^{-1}, \qquad && Sf = -qf, \qquad && Se = -q^{-1}e.
\end{aligned}
$$

With $k = q^H$, in the $q = 1$ limit the elements H, f, e generate the Lie algebra $su(2)$ and the universal enveloping algebra $\mathcal{U}(su(2))$. The action of $su(2)$ (and of $\mathcal{U}(su(2))$) on $SU(2)$ is generalized to a pairing between the corresponding deformed algebras.

Definition 9.4 The duality pairing between $\mathcal{U}$ and $\mathcal{A}$ is defined on generators by

$$\langle k, a\rangle = q^{\frac{1}{2}}, \quad \langle k, a^*\rangle = q^{-\frac{1}{2}}, \quad \langle e, -qb^*\rangle = \langle f, b\rangle = 1,$$

with all other couples of generators pairing to 0. It satisfies

$$\langle (Sh)^*, x\rangle = \overline{\langle h, x^*\rangle}, \quad \text{for all} \quad h \in \mathcal{U},\ x \in \mathcal{A}.$$

This pairing gives [76] canonical left and right $\mathcal{U}$-module algebra structures on $\mathcal{A}$ by

$$\langle g, h \triangleright x\rangle := \langle gh, x\rangle \quad \langle g, x \triangleleft h\rangle := \langle hg, x\rangle\,, \quad \text{for all} \quad g, h \in \mathcal{U},\ x \in \mathcal{A}.$$

These mutually commuting actions of $\mathcal{U}$ on $\mathcal{A}$ are given by

$$h \triangleright x := (\mathrm{id} \otimes h)\, \Delta x = x_{(1)}\, \big\langle h, x_{(2)}\big\rangle\,,$$
$$x \triangleleft h := (h \otimes \mathrm{id})\, \Delta x = \big\langle h, x_{(1)}\big\rangle\, x_{(2)},$$

using the Sweedler notation $\Delta x =: x_{(1)} \otimes x_{(2)}$ with implicit summation. It follows from the properties in Def. 9.4 that the star structure is compatible with both actions,

$$h \triangleright x^* = ((Sh)^* \triangleright x)^*, \quad x^* \triangleleft h = (x \triangleleft (Sh)^*)^*, \quad \text{for all} \quad h \in \mathcal{U},\ x \in \mathcal{A},$$

and they are linked through the antipodes: $S(Sh \triangleright x) = Sx \triangleleft h$.

Let us onsider the algebra automorphism ϑ of $\mathcal{U}$ defined on generators by

$$\vartheta(k) := k^{-1}, \quad \vartheta(f) := -e, \quad \vartheta(e) := -f,$$

and which is an antiautomorphism for the coalgebra structure of $\mathcal{U}$. Since $S^{-1} \circ \vartheta$ is an algebra antiautomorphism of $\mathcal{U}$, it converts a right action into a left action; and because both S^{-1} and ϑ are coalgebra antiautomorphisms, $S^{-1} \circ \vartheta$ preserves the coalgebra structure of $\mathcal{U}$. Thus, we get a second left action defined by

$$h \cdot x := x \triangleleft S^{-1}(\vartheta(h)).$$

and commuting with the first one. Both left actions are given on all generators by:

$$\begin{aligned}
&k \triangleright a = q^{\frac{1}{2}} a, && k \triangleright a^* = q^{-\frac{1}{2}} a^*, && k \triangleright b = q^{-\frac{1}{2}} b, && k \triangleright b^* = q^{\frac{1}{2}} b^*,\\
&f \triangleright a = 0, && f \triangleright a^* = -qb^*, && f \triangleright b = a, && f \triangleright b^* = 0,\\
&e \triangleright a = b, && e \triangleright a^* = 0, && e \triangleright b = 0, && e \triangleright b^* = -q^{-1} a^*,
\end{aligned}$$

and

$$\begin{aligned}
&k \cdot a = q^{\frac{1}{2}} a, && k \cdot a^* = q^{-\frac{1}{2}} a^*, && k \cdot b = q^{\frac{1}{2}} b, && k \cdot b^* = q^{-\frac{1}{2}} b^*,\\
&f \cdot a = 0, && f \cdot a^* = qb, && f \cdot b = 0, && f \cdot b^* = -a,\\
&e \cdot a = -b^*, && e \cdot a^* = 0, && e \cdot b = q^{-1} a^*, && e \cdot b^* = 0.
\end{aligned}$$

Together, they give a left action of $\mathcal{U}_q(su(2)) \otimes \mathcal{U}_q(su(2))$ on $\mathcal{A}(SU_q(2))$, that extends to the case $q \neq 1$ the (infinitesimal) classical action of $\mathrm{Spin}(4) = SU(2)) \times SU(2)$ on $SU(2) \approx \mathbb{S}^3$, realized as two commuting left actions of $SU(2)$.

Next, we recall [45] that $\mathcal{A}$ has a vector-space basis consisting of matrix elements of its irreducible corepresentations, $\{t^l_{mn} : 2l \in \mathbb{N},\ m, n = -l, \dots, l-1, l\}$, with

$$t^0_{00} = 1, \qquad t^{\frac{1}{2}}_{\frac{1}{2},\frac{1}{2}} = a, \qquad t^{\frac{1}{2}}_{\frac{1}{2},-\frac{1}{2}} = b.$$

The coproduct has the matricial form $\Delta t^l_{mn} = \sum_k t^l_{mk} \otimes t^l_{kn}$, while the product is

$$t^j_{rs} t^l_{mn} = \sum_{k=|j-l|}^{j+l} C_q \begin{pmatrix} j & l & k \\ r & m & r+m \end{pmatrix} C_q \begin{pmatrix} j & l & k \\ s & n & s+n \end{pmatrix} t^k_{r+m,s+n},$$

where the $C_q(-)$ factors are q-Clebsch–Gordan coefficients [4, 44].

The Haar state ψ on the C^*-completion $C(SU_q(2))$ is determined by setting $\psi(1) := 1$ and $\psi(t^l_{mn}) := 0$ if $l > 0$. Let $\mathcal{H}_\psi = L^2(SU_q(2), \psi)$ be the Hilbert space of its GNS representation. The GNS map $\eta : C(SU_q(2)) \to \mathcal{H}_\psi$ is injective and satisfies

$$\|\eta(t^l_{mn})\|^2 = \psi((t^l_{mn})^* \, t^l_{mn}) = \frac{q^{-2m}}{[2l+1]},$$

and the vectors $\eta(t^l_{mn})$ are mutually orthogonal. The involution in $C(SU_q(2))$ is given by

$$(t^l_{mn})^* = (-1)^{2l+m+n} q^{n-m} \, t^l_{-m,-n}, \tag{9.1}$$

and in particular, $t^{\frac{1}{2}}_{-\frac{1}{2},\frac{1}{2}} = -qb^*$ and $t^{\frac{1}{2}}_{-\frac{1}{2},-\frac{1}{2}} = a^*$. An orthonormal basis of $\mathcal{H}_\psi$ is given by

$$|lmn\rangle := q^m \, [2l+1]^{\frac{1}{2}} \, \eta(t^l_{mn}). \tag{9.2}$$

We denote by π_ψ the corresponding GNS representation of $C(SU_q(2))$ on $\mathcal{H}_\psi$,

$$\pi_\psi(x) \, |lmn\rangle := q^m \, [2l+1]^{\frac{1}{2}} \, \eta(x \, t^l_{mn}). \tag{9.3}$$

9.2 The equivariant representation of $\mathcal{A}(SU_q(2))$

The regular representation of the algebra $\mathcal{A}(SU_q(2))$ on its GNS space $\mathcal{H}_\psi$ is fully determined by its equivariance properties with respect to the left Hopf action of $\mathcal{U} \otimes \mathcal{U}$.

Definition 9.5 Let λ and ρ be mutually commuting representations of the Hopf algebra $\mathcal{U}$ on a vector space V. A representation π of the $*$-algebra $\mathcal{A}$ on V is (λ, ρ)-*equivariant* if the following compatibility relations hold [68]:

$$\lambda(h) \, \pi(x)\xi = \pi(h_{(1)} \cdot x) \, \lambda(h_{(2)})\xi, \qquad \rho(h) \, \pi(x)\xi = \pi(h_{(1)} \triangleright x) \, \rho(h_{(2)})\xi,$$

for all $h \in \mathcal{U}$, $x \in \mathcal{A}$ and $\xi \in V$.

For the case at hand, the two $\mathcal{U}_q(su(2))$ symmetries λ and ρ decompose into components according to the irreducible (involutive) representations of $\mathcal{U} = \mathcal{U}_q(su(2))$, which are well known [45]. The irreducible $*$-representations σ_l of $\mathcal{U}_q(su(2))$ are labelled by nonnegative half-integers $l = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$, acting on $\mathcal{U}$-modules V_l with $\dim V_l = 2l+1$, and having orthonormal bases $\{|lm\rangle : m = -l, -l+1, \dots, l-1, l\}$; they are given by

$$\begin{aligned} \sigma_l(k)\ |lm\rangle &= q^m\ |lm\rangle\,, \\ \sigma_l(f)\ |lm\rangle &= \sqrt{[l-m][l+m+1]}\ |l, m+1\rangle\,, \\ \sigma_l(e)\ |lm\rangle &= \sqrt{[l-m+1][l+m]}\ |l, m-1\rangle\,. \end{aligned} \tag{9.4}$$

Here, for each $n \in \mathbb{Z}$, $[n] =:= (q^n - q^{-n})/(q - q^{-1})$ is the corresponding "q-integer".

The equivariant representation of $\mathcal{A}(SU_q(2))$ acts on the pre-Hilbert space

$$V := \bigoplus_{2l=0}^{\infty} V_l \otimes V_l,$$

while the symmetries λ and ρ act on the first and the second leg of the tensor product respectively, via the irreducible representations (9.4),

$$\lambda(h) = \sigma_l(h) \otimes \mathrm{id}, \qquad \rho(h) = \mathrm{id} \otimes \sigma_l(h) \qquad \text{on } V_l \otimes V_l.$$

We abbreviate $|lmn\rangle := |lm\rangle \otimes |ln\rangle$, for $m, n = -l, \dots, l-1, l$. These form an orthonormal basis for $V_l \otimes V_l$, for each fixed l. Also, we adopt a shorthand notation,

$$l^{\pm} := l \pm \frac{1}{2}, \quad m^{\pm} := m \pm \frac{1}{2}, \quad n^{\pm} := n \pm \frac{1}{2}.$$

Proposition 9.1 *A (λ, ρ)-equivariant $*$-representation π of $\mathcal{A}(SU_q(2))$ on V must have the following form,*

$$\begin{aligned} \pi(a)\ |lmn\rangle &= A^{+}_{lmn}\left|l^{+}m^{+}n^{+}\right\rangle + A^{-}_{lmn}\left|l^{-}m^{+}n^{+}\right\rangle, \\ \pi(b)\ |lmn\rangle &= B^{+}_{lmn}\left|l^{+}m^{+}n^{-}\right\rangle + B^{-}_{lmn}\left|l^{-}m^{+}n^{-}\right\rangle, \end{aligned} \tag{9.5}$$

where, up to phase factors depending only on l, the constants $A^{\pm}_{lmn}$ and $B^{\pm}_{lmn}$ are,

$$\begin{aligned} A^{+}_{lmn} &= q^{(-2l+m+n-1)/2}\left(\frac{[l+m+1][l+n+1]}{[2l+1][2l+2]}\right)^{\frac{1}{2}}, \\ A^{-}_{lmn} &= q^{(2l+m+n+1)/2}\left(\frac{[l-m][l-n]}{[2l][2l+1]}\right)^{\frac{1}{2}}, \\ B^{+}_{lmn} &= q^{(m+n-1)/2}\left(\frac{[l+m+1][l-n+1]}{[2l+1][2l+2]}\right)^{\frac{1}{2}}, \\ B^{-}_{lmn} &= -q^{(m+n-1)/2}\left(\frac{[l-m][l+n]}{[2l][2l+1]}\right)^{\frac{1}{2}}. \end{aligned}$$

As shown in [29], the formulae (9.5) give precisely the left regular representation π_ψ of $\mathcal{A}(SU_q(2))$ in (9.3). The identification (9.2) embeds the pre-Hilbert space V densely in the Hilbert space $\mathcal{H}_\psi$, and the representation π_ψ extends to the GNS representation of $C(SU_q(2))$ on $\mathcal{H}_\psi$, as described by the Peter–Weyl theorem [45, 76].

9.3 The spin representation

Definition 9.6 The left regular representation π of $\mathcal{A}$ is amplified to $\pi' = \pi \otimes \mathrm{id}$ on

$$W := V \otimes \mathbb{C}^2 = V \otimes V_{\frac{1}{2}} \; .$$

In the commutative case when $q = 1$, this yields the spinor representation of $SU(2)$, because the spinor bundle is parallelizable, $S \simeq SU(2) \times \mathbb{C}^2$. The representation theory of $\mathcal{U}$ (and the corepresentation theory of $\mathcal{A}$) follows the same pattern: when $q \neq 1$ only the Clebsch–Gordan coefficients need to be modified [44]. The Clebsch–Gordan decomposition of W is the (algebraic) direct sum

$$\begin{aligned} W = \left(\bigoplus_{2l=0}^{\infty} V_l \otimes V_l \right) \otimes V_{\frac{1}{2}} \simeq V_{\frac{1}{2}} \oplus \bigoplus_{2j=1}^{\infty} (V_{j+\frac{1}{2}} \otimes V_j) \oplus (V_{j-\frac{1}{2}} \otimes V_j), \\ = W_0^{\uparrow} \oplus \bigoplus_{2j \geq 1} W_j^{\uparrow} \oplus W_j^{\downarrow}. \end{aligned} \tag{9.6}$$

Here $W_j^{\uparrow} \simeq V_{j+\frac{1}{2}} \otimes V_j$ and $W_j^{\downarrow} \simeq V_{j-\frac{1}{2}} \otimes V_j$, with $\dim W_j^{\uparrow} = (2j+1)(2j+2)$, and $\dim W_j^{\downarrow} = 2j(2j+1)$, for $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$ (there is no $W_0^{\downarrow}$).

Definition 9.7 We amplify the representation ρ of $\mathcal{U}$ on V to $\rho' = \rho \otimes \mathrm{id}$ on $W = V \otimes \mathbb{C}^2$. However, we replace λ on V by its tensor product with $\sigma_{\frac{1}{2}}$ on $\mathbb{C}^2$,

$$\lambda'(h) := (\lambda \otimes \sigma_{\frac{1}{2}})(\Delta h) = \lambda(h_{(1)}) \otimes \sigma_{\frac{1}{2}}(h_{(2)}).$$

It is straightforward to check that the representations λ' and ρ' on W commute, and that the representation π' of $\mathcal{A}$ on W is (λ', ρ')-equivariant,

$$\lambda'(h)\, \pi'(x)\psi = \pi'(h_{(1)} \cdot x)\, \lambda'(h_{(2)})\psi, \qquad \rho'(h)\, \pi'(x)\psi = \pi'(h_{(1)} \triangleright x)\, \rho'(h_{(2)})\psi,$$

for all $h \in \mathcal{U}$, $x \in \mathcal{A}$ and $\psi \in W$.

The Hilbert space of spinors is $\mathcal{H} := \mathcal{H}_\psi \otimes \mathbb{C}^2$, which is just the completion of W. It can be decomposed as $\mathcal{H} = \mathcal{H}^{\uparrow} \oplus \mathcal{H}^{\downarrow}$, where $\mathcal{H}^{\uparrow}$ and $\mathcal{H}^{\downarrow}$ are the respective completions of $\bigoplus_{2j \geq 0} W_j^{\uparrow}$ and $\bigoplus_{2j \geq 1} W_j^{\downarrow}$. An better explicit basis for W, is given as follows.

For $j = l + \frac{1}{2}$, $\mu = m - \frac{1}{2}$, with $\mu = -j, \dots, j$ and $n = -j^-, \dots, j^-$, let

$$|j\mu n \downarrow\rangle := C_{j\mu}\, |j^- \mu^+ n\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle + S_{j\mu}\, |j^- \mu^- n\rangle \otimes \left| \frac{1}{2}, +\frac{1}{2} \right\rangle;$$

and for $j = l - \frac{1}{2}$, $\mu = m - \frac{1}{2}$, with $\mu = -j, \dots, j$ and $n = -j^+, \dots, j^+$, let

$$|j\mu n \uparrow\rangle := -S_{j+1,\mu}\, |j^+ \mu^+ n\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle + C_{j+1,\mu}\, |j^+ \mu^- n\rangle \otimes \left| \frac{1}{2}, +\frac{1}{2} \right\rangle,$$

where the coefficients are

$$C_{j\mu} := q^{-(j+\mu)/2} \, \frac{[j-\mu]^{\frac{1}{2}}}{[2j]^{\frac{1}{2}}}, \qquad S_{j\mu} := q^{(j-\mu)/2} \, \frac{[j+\mu]^{\frac{1}{2}}}{[2j]^{\frac{1}{2}}}.$$

Notice that there are no $\downarrow$ vectors for $j = 0$. It is straightforward to verify that these vectors make up orthonormal bases for $W_j^{\downarrow}$ and $W_j^{\uparrow}$, respectively.

The representation π' can be computed in the new spinor basis by conjugating the form of $\pi \otimes \mathrm{id}$ found in Prop. 9.1 by the basis transformation (9.7). Equivalently, it can also be derived from the (λ', ρ')-equivariance.

Definition 9.8 For $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$, with $\mu = -j, \dots, j$ and $n = -j - \frac{1}{2}, \dots, j + \frac{1}{2}$, we juxtapose the pair of spinors

$$|j\mu n\rangle\!\rangle := \begin{pmatrix} |j\mu n \uparrow\rangle \\ |j\mu n \downarrow\rangle \end{pmatrix},$$

with the convention that the lower component is zero when $n = \pm(j + \frac{1}{2})$ or $j = 0$. Furthermore, a matrix with scalar entries,

$$A = \begin{pmatrix} A_{\uparrow\uparrow} & A_{\uparrow\downarrow} \\ A_{\downarrow\uparrow} & A_{\downarrow\downarrow} \end{pmatrix},$$

is understood to act on $|j\mu n\rangle\!\rangle$ by the rule:

$$A\,|j\mu n \uparrow\rangle = A_{\uparrow\uparrow}\,|j\mu n \uparrow\rangle + A_{\downarrow\uparrow}\,|j\mu n \downarrow\rangle\,, \qquad A\,|j\mu n \downarrow\rangle = A_{\downarrow\downarrow}\,|j\mu n \downarrow\rangle + A_{\uparrow\downarrow}\,|j\mu n \uparrow\rangle\,.$$

Proposition 9.2 *The spinor $*$-representation $\pi' := \pi \otimes \mathrm{id}$ of $\mathcal{A}$ on $\mathcal{H}$ is written as*

$$\pi(a) := a_+ + a_-, \qquad \pi(b) := b_+ + b_-,$$

where $a_\pm$ and $b_\pm$ are, up to phase factors depending only on j, the triangular operators,

$$a_+\,|j\mu n\rangle\!\rangle := q^{(\mu+n-\frac{1}{2})/2}[j+\mu+1]^{\frac{1}{2}} \cdot \begin{pmatrix} q^{-j-\frac{1}{2}}\,\frac{[j+n+\frac{3}{2}]^{1/2}}{[2j+2]} & 0 \\ q^{\frac{1}{2}}\,\frac{[j-n+\frac{1}{2}]^{1/2}}{[2j+1]\,[2j+2]} & q^{-j}\,\frac{[j+n+\frac{1}{2}]^{1/2}}{[2j+1]} \end{pmatrix} |j^+\mu^+n^+\rangle\!\rangle,$$

$$a_-\,|j\mu n\rangle\!\rangle := q^{(\mu+n-\frac{1}{2})/2}[j-\mu]^{\frac{1}{2}} \cdot \begin{pmatrix} q^{j+1}\,\frac{[j-n+\frac{1}{2}]^{1/2}}{[2j+1]} & -q^{\frac{1}{2}}\,\frac{[j+n+\frac{1}{2}]^{1/2}}{[2j]\,[2j+1]} \\ 0 & q^{j+\frac{1}{2}}\,\frac{[j-n-\frac{1}{2}]^{1/2}}{[2j]} \end{pmatrix} |j^-\mu^+n^+\rangle\!\rangle,$$

$$b_+\,|j\mu n\rangle\!\rangle := q^{(\mu+n-\frac{1}{2})/2}[j+\mu+1]^{\frac{1}{2}} \cdot \begin{pmatrix} \frac{[j-n+\frac{3}{2}]^{1/2}}{[2j+2]} & 0 \\ -q^{-j-1}\,\frac{[j+n+\frac{1}{2}]^{1/2}}{[2j+1]\,[2j+2]} & q^{-\frac{1}{2}}\,\frac{[j-n+\frac{1}{2}]^{1/2}}{[2j+1]} \end{pmatrix} |j^+\mu^+n^-\rangle\!\rangle,$$

$$b_-\,|j\mu n\rangle\!\rangle := q^{(\mu+n-\frac{1}{2})/2}[j-\mu]^{\frac{1}{2}} \cdot \begin{pmatrix} -q^{-\frac{1}{2}}\,\frac{[j+n+\frac{1}{2}]^{1/2}}{[2j+1]} & -q^{j}\,\frac{[j-n+\frac{1}{2}]^{1/2}}{[2j]\,[2j+1]} \\ 0 & -\frac{[j+n-\frac{1}{2}]^{1/2}}{[2j]} \end{pmatrix} |j^-\mu^+n^-\rangle\!\rangle.$$

9.4 The equivariant Dirac operator

The Casimir element of $\mathcal{U}_q(su(2))$ can be taken to be $C_q = qk^2 + q^{-1}k^{-2} + (q - q^{-1})^2 ef$. The symmetric operators $\lambda'(C_q)$ and $\rho'(C_q)$ on $\mathcal{H}$, defined with dense domain W, extend to selfadjoint operators on $\mathcal{H}$ and the subspaces $W_j^{\uparrow}$, $W_j^{\downarrow}$ are their joint eigenspaces,

$$\begin{aligned}
\lambda'(C_q)\,|j\mu n\uparrow\rangle &= (q^{2j+1} + q^{-2j-1})\;|j\mu n\uparrow\rangle\,,\\
\rho'(C_q)\,|j\mu n\uparrow\rangle &= (q^{2j+2} + q^{-2j-2})\;|j\mu n\uparrow\rangle\,,\\
\lambda'(C_q)\,|j\mu n\downarrow\rangle &= (q^{2j+1} + q^{-2j-1})\;|j\mu n\downarrow\rangle\,,\\
\rho'(C_q)\,|j\mu n\downarrow\rangle &= (q^{2j} + q^{-2j})\;|j\mu n\downarrow\rangle\,,
\end{aligned}$$

The finite-dimensional subspaces $W_j^{\uparrow}$ and $W_j^{\downarrow}$ will reduce any selfadjoint operator D on $\mathcal{H}$ which commutes strongly with $\lambda'(C_q)$ and $\rho'(C_q)$. If we require that D be invariant under the actions λ' and ρ', we get the following stronger condition.

Lemma 9.1 *Let D be a selfadjoint operator that commutes strongly with $\lambda'(h)$ and $\rho'(h)$, for each $h \in \mathcal{U}$. Then the subspaces $W_j^{\uparrow}$ and $W_j^{\downarrow}$ are eigenspaces for D,*

$$D\,|j\mu n\uparrow\rangle = d_j^{\uparrow}\;|j\mu n\uparrow\rangle\,, \qquad D\,|j\mu n\downarrow\rangle = d_j^{\downarrow}\;|j\mu n\downarrow\rangle\,, \tag{9.8}$$

where $d_j^{\uparrow}$ and $d_j^{\downarrow}$ are real eigenvalues of D which depend only on j. Their respective multiplicities are $(2j+1)(2j+2)$ and $2j(2j+1)$.

Additional natural restrictions on the eigenvalues $d_j^{\uparrow}, d_j^{\downarrow}$ of the operator D will come from the crucial requirement of boundedness of the commutators $[D, \pi'(x)]$ for $x \in \mathcal{A}$. The "q-Dirac" operator D proposed in [3] corresponds to taking, in our notation,

$$d_j^{\uparrow} = \frac{2\,[2j+1]}{q+q^{-1}}, \qquad d_j^{\downarrow} = -d_j^{\uparrow}.$$

These are q-analogues of the classical eigenvalues of $D\!\!\!\!/ - \frac{1}{2}$ where $D\!\!\!\!/$ is the classical Dirac operator on the sphere $\mathbb{S}^3$ (with the round metric). For this particular choice it follows directly from the explicit form of the representation in Prop. 9.2 that, for instance, the commutator $[D, \pi'(a)]$ is unbounded. This fact was already noted in [21] and it was suggested that one should instead consider an operator D whose spectrum is just that of the classical Dirac operator $D\!\!\!\!/$.

Proposition 9.3 *Let D be any selfadjoint operator with eigenspaces $W_j^{\uparrow}$ and $W_j^{\downarrow}$, and eigenvalues* (9.8). *If the eigenvalues $d_j^{\uparrow}$ and $d_j^{\downarrow}$ are linear in j,*

$$d_j^{\uparrow} = c_1^{\uparrow} j + c_2^{\uparrow}, \qquad d_j^{\downarrow} = c_1^{\downarrow} j + c_2^{\downarrow},$$

with $c_1^{\uparrow}$, $c_2^{\uparrow}$, $c_1^{\downarrow}$, $c_2^{\downarrow}$ not depending on j, then $[D, \pi'(x)]$ is a bounded operator for all $x \in \mathcal{A}$.

A selfadjoint operator D as in Prop. 9.3 is essentially the only possibility for a Dirac operator satisfying a (modified) first-order condition. It is necessary that we assume $c_1^{\downarrow} c_1^{\uparrow} < 0$ in order that the sign of the operator D be nontrivial; but up to irrelevant scaling factors the choice of $c_j^{\uparrow}$, $c_j^{\downarrow}$ is otherwise immaterial. With the particular choice $d_j^{\uparrow} = 2j + \frac{3}{2}$ and $d_j^{\downarrow} = -2j - \frac{1}{2}$, the spectrum of D, with multiplicity, coincides with that of the classical

Dirac operator $D\!\!\!\!/$ on the round sphere $\mathbb{S}^3$. Thus, we can regard our spectral triple as an isospectral deformation of $(C^\infty(\mathbb{S}^3), \mathcal{H}, D\!\!\!\!/)$, and in particular, its spectral dimension is 3.

The above particular choice of the classical eigenvalues has an extra benefit: it is easy to prove regularity. We summarize our conclusions in the following theorem.

Theorem 9.1 *The triple $(\mathcal{A}(SU_q(2)), \mathcal{H}, D)$, where the eigenvalues of D are given by*

$$d_j^\uparrow = 2j + \frac{3}{2}, \qquad d_j^\downarrow = -2j - \frac{1}{2},$$

is a regular 3^+-summable spectral triple.

9.5 The real structure

For the manifold of $SU_q(2)$, by adding an equivariant real structure J it is not possible to satisfy all usual properties of a real spectral triple, while having a nontrivial Dirac operator. The commutant properties: that J intertwines a left action and a commuting right action of the algebra on the Hilbert space; and the first order condition on D: that the commutators $[D, a]$, for any element a in the algebra, commute with the opposite action by any b, are satisfied only up to a certain ideal of compact operators.

On the GNS representation space $\mathcal{H}_\psi$, the natural involution $T_\psi : \eta(x) \mapsto \eta(x^*)$, is an unbounded (antilinear) operator on $\mathcal{H}_\psi$ with domain $\eta(C(SU_q(2)))$. From the Tomita–Takesaki theory [70], its closure has a polar decomposition $T_\psi =: J_\psi \Delta_\psi^{1/2}$ which defines both the positive "modular operator" Δ_ψ and the antiunitary "modular conjugation" J_ψ. From (9.1) and (9.2) it follows that

$$T_\psi \ |lmn\rangle = (-1)^{2l+m+n} q^{m+n} \ |l, -m, -n\rangle \,,$$

and the adjoint antilinear operator is given by

$$T_\psi^* \ |lmn\rangle = (-1)^{2l+m+n} q^{-m-n} \ |l, -m, -n\rangle \ .$$

Since $\Delta_\psi = T_\psi^* T_\psi$, it follows that every vector $|lmn\rangle$ lies in the domain $\operatorname{Dom}\Delta_\psi$ with $\Delta_\psi \ |lmn\rangle = q^{2m+2n} \ |lmn\rangle$. Consequently,

$$J_\psi \ |lmn\rangle = (-1)^{2l+m+n} \ |l, -m, -n\rangle \,.$$

It is clear that $J_\psi^2 = 1$ on $\mathcal{H}_\psi$.

Definition 9.9 Let $\pi^\circ(x) := J_\psi \, \pi(x^*) \, J_\psi^{-1}$, so that π° is a $*$-antirepresentation of $\mathcal{A}$ on $\mathcal{H}_\psi$. Equivalently, π° is a $*$-representation of the opposite algebra $\mathcal{A}(SU_{1/q}(2))$. By Tomita's theorem [70], π and π° are commuting representations.

The (λ, ρ)-equivariance of π is reflected in an analogous equivariance condition for π°.

Lemma 9.2 *The symmetry of the antirepresentation π° of $\mathcal{A}$ on $\mathcal{H}_\psi$ is given by the equivariance conditions,*

$$\lambda(h) \, \pi^\circ(x)\xi = \pi^\circ(\tilde{h}_{(2)} \cdot x) \, \lambda(h_{(1)})\xi, \qquad \rho(h) \, \pi^\circ(x)\xi = \pi^\circ(\tilde{h}_{(2)} \triangleright x) \, \rho(h_{(1)})\xi,$$

for all $h \in \mathcal{U}$, $x \in \mathcal{A}$ and $\xi \in \mathcal{H}_\psi$. Here $h \mapsto \tilde{h}$ is the automorphism of $\mathcal{U}$ determined on generators by $\tilde{k} := k$, $\tilde{f} := q^{-1} f$, and $\tilde{e} := qe$.

Recall that $T_\psi \eta(x) = \eta(x^*)$ for all $x \in \mathcal{A}$ and that $\eta(x) = \pi(x)\ |000\rangle$. From Def. 9.5 one finds that for generators h of $\mathcal{U}$,

$$T_\psi \lambda(h)\pi(x)\ |000\rangle = \pi(x^* \triangleleft \vartheta(h)^*)\ |000\rangle = \lambda(S(h)^*) T_\psi \pi(x)\ |000\rangle\,,$$

where we have used the relation $S(\vartheta(h)^*) = \vartheta(S(h)^*)$. Since the vector $|000\rangle$ is separating for the GNS representation, we conclude that

$$T_\psi\, \lambda(h)\, T_\psi^{-1} = \lambda((Sh)^*).$$

Similarly, we find that $T_\psi\, \rho(h)\, T_\psi^{-1} = \rho((Sh)^*)$. Thus, the antilinear involutory automorphism $h \mapsto (Sh)^*$ of the Hopf $*$-algebra $\mathcal{U}$ is implemented by the Tomita operator for the Haar state of the dual Hopf $*$-algebra $\mathcal{A}$. This is a known feature of quantum-group duality in the C^*-algebra framework [55].

The first step in defining an operator J on spinors is to construct the "right multiplication" representation of $\mathcal{A}$ on spinors from its symmetry alone, in close parallel with the equivariance conditions of Lem. 9.2 for the right regular representation π° of $\mathcal{A}$ on $\mathcal{H}_\psi$. Then, the conjugation operator J on spinors is constructed as the one that intertwines the left and the right spinor representations.

Proposition 9.4 *Let π'° be an antirepresentation of $\mathcal{A}$ on $\mathcal{H} = \mathcal{H}_\psi \oplus \mathcal{H}_\psi$ satisfying the following equivariance conditions:*

$$\lambda'(h)\, \pi'^\circ(x)\xi = \pi'^\circ(\tilde{h}_{(2)} \cdot x)\, \lambda'(h_{(1)})\xi, \qquad \rho'(h)\, \pi'^\circ(x)\xi = \pi'^\circ(\tilde{h}_{(2)} \triangleright x)\, \rho'(h_{(1)})\xi.$$

Then, up to some phase factors depending only on the index j in the decomposition (9.6)*, π'° is given on the spinor basis by $\pi'^\circ(a) = a^\circ_+ + a^\circ_-$ and $\pi'^\circ(b) = b^\circ_+ + b^\circ_-$, where in direct analogy with spinor representation in Prop. 9.2, the operators $a^\circ_\pm$ and $b^\circ_\pm$ have the triangular-matrix form as there, acting on the same respective basis vectors: but with the coefficients of $a_\pm$ modified by the replacement $q \mapsto q^{-1}$, while the coefficients of $b_\pm$ are modified by the replacement $q \mapsto q^{-1}$ and by multiplying the result by an overall factor q^{-1}.*

Definition 9.10 The conjugation operator J is the antilinear operator on $\mathcal{H}$ defined explicitly on the orthonormal spinor basis by

$$\begin{aligned} J\ |j\mu n\uparrow\rangle &:= i^{2(2j+\mu+n)}\ |j,-\mu,-n,\uparrow\rangle\,, \\ J\ |j\mu n\downarrow\rangle &:= i^{2(2j-\mu-n)}\ |j,-\mu,-n,\downarrow\rangle\,. \end{aligned}$$

It is immediate to see that J is antiunitary and that $J^2 = -1$, since each $4j \pm 2(\mu+n)$ is an odd integer.

Proposition 9.5 *The operator J intertwines the left and right spinor representations:*

$$J\, \pi'(x^*)\, J^{-1} = \pi'^\circ(x), \quad \textit{for all} \quad x \in \mathcal{A}.$$

For the invariant operator D of Sect. 9.4, from the diagonal form of both D and J on their common eigenspaces $W_j^\uparrow$ and $W_j^\downarrow$, given by the respective equations (9.8) and (9.10), it easily follows that D and J commute.

Proposition 9.6 *The invariant operator D of equation* (9.8) *commutes with the conjugation operator J of Def. 9.10:*

$$JDJ^{-1} = D.$$

The conjugation operator J of Def. 9.10 is *not* the Tomita modular conjugation for the spinor representation of $\mathcal{A}$, a fact having consequences on some of the requirements for a real spectral triple, as we shall see presently. The Tomita operator for spinor is $J_\psi \oplus J_\psi$, which does not have a diagonal form in our chosen spinor basis (unless $q = 1$). As mentioned above, conjugation of $\pi'(\mathcal{A}(SU_q(2))$ by the modular operator would yield a representation of the opposite algebra $\mathcal{A}(SU_{1/q}(2))$, and the commutation relation analogous to (9.6) would then require D to be equivariant under the corresponding symmetry of $U_{1/q}(su(2))$. This extra equivariance condition would force D to be merely a scalar operator, thereby ruling out the possibility of an equivariant 3^+-summable real spectral triple on $\mathcal{A}(SU_q(2))$ with the modular conjugation operator.

In order to get a nontrivial Dirac operator, the remedy is to modify J to a non-Tomita conjugation operator. As mentioned, the price to pay for this is that the conditions for a real spectral triple must be weakened: these are only satisfied up to certain trace-class operators, in fact infinitesimals of arbitrary high order.

Definition 9.11 We denote by $\mathcal{K}_q$ be the two-sided ideal of $\mathcal{B}(\mathcal{H})$ generated by the positive trace-class operators L_q given by

$$L_q|j\mu n\rangle\rangle := q^j \, |j\mu n\rangle\rangle \quad \text{for} \quad j \in \frac{1}{2}\mathbb{N},$$

The spectral triple over $\mathcal{A}(SU_q(2))$ is characterized by the folowing,

Theorem 9.2 *The real spectral triple $(\mathcal{A}(SU_q(2)), \mathcal{H}, D, J)$, with $\mathcal{A}(SU_q(2))$ acting on $\mathcal{H}$ via the spinor representation π' of Prop. 9.2, satisfies both the commutant property and the first order condition up to infinitesimals,*

$$[\pi'^{\circ}(x), \pi'(y)] \in \mathcal{K}_q, \qquad [\pi'^{\circ}(x), [D, \pi'(y)]] \in \mathcal{K}_q, \qquad \forall\, x, y \in \mathcal{A}(SU_q(2)).$$

The ideal $\mathcal{K}_q$ is made of trace-class operators and is contained in the ideal of infinitesimals of arbitrary high order. With Def. 7.1, a compact operator T is an infinitesimal of arbitrary high order if its singular values $\mu_j(T)$ satisfy $\lim_{j\to\infty} j^p \mu_j(A) = 0$ for all $p > 0$.

Bibliography

[1] M. F. Atiyah: Global theory of elliptic operators *In: Proc. Intn'l Conf. on Functional analysis and related topics* (Tokyo 1969, Univ. of Tokyo Press, 1970) 21–30

[2] N. Berline, E. Getzler and M. Vergne: *Heat Kernels and Dirac Operators* (Springer, 1991)

[3] P. N. Bibikov and P. P. Kulish: Dirac operators on the quantum group $SU_q(2)$ and the quantum sphere *J. Math. Sci.* **100** (2000) 2039–2050

[4] L. C. Biedenharn and M. A. Lohe: *Quantum Group Symmetry and q-Tensor Algebras* (World Scientific, 1995)
[5] N. Bourbaki: *Eléments de mathématique: algèbre* (Diffusion C.C.L.S., 1970)
[6] N. Bourbaki: *Eléments de mathématique: topologiegénérale* (Diffusion C.C.L.S., 1976)
[7] P. Budinich and A. Trautman: *The Spinorial Chessboard* (Springer, 1988)
[8] A. L. Carey, J. Phillips and F. Sukochev: Spectral flow and Dixmier traces *Adv. Math.* **173** (2003) 68–113
[9] H. Cartan, S. Eilenberg: *Homological Algebra* (Princeton University Press, 1973)
[10] P. S. Chakraborty, A. Pal: Equivariant spectral triples on the quantum $SU(2)$ group *K-Theory* **28** (2003) 107–126
[11] F. Cipriani, D. Guido, S. Scarlatti: A remark on trace properties of K-cycles *J. Oper. Theory* **35** (1996) 179–189
[12] A. Connes: C^*-algèbres et géométriedifférentielle *C.R. Acad. Sci. Paris Sér. A-B* **290** (1980) A599–A604; An english translation is: A. Connes, C^*-algebras and differential geometry, `arXiv:hep-th/0101093`
[13] A. Connes: Noncommutative differentialgeometry *I.H.E.S. Publ. Math.* **62** (1985) 257–360
[14] A. Connes: The action functional in non-commutative Geometry *Commun. Math. Phys.* **117** (1988) 673–683
[15] A. Connes: Compact metric spaces, Fredholm modules and hyperfiniteness *Ergod. Theory and Dynam. Syst.* **9** (1989) 208–220
[16] A. Connes: *Noncommutative Geometry* (Academic Press, 1994)
[17] A. Connes: Noncommutative geometry and reality *J. Math.Phys.* **36** (1995) 6194-6231
[18] A. Connes: Gravity coupled with matter andfoundation of noncommutative geometry *Commun. Math. Phys.* **182** (1996) 155–176
[19] A. Connes: Cyclic cohomology, quantum group symmetries and the local index formula for $SU_q(2)$ *J. Inst. Math. Jussieu* **3** (2004) 17–68
[20] A. Connes and M. Dubois-Violette: Noncommutative finite-dimensional manifolds. I. Spherical manifolds and related examples *Commun. Math. Phys.* **230** (2002) 539–579
[21] A. Connes, G. Landi: Noncommutative manifolds, the instanton algebraand isospectral deformations *Commun. Math. Phys.* **221** (2001) 141–159
[22] A. Connes, M. Marcolli: A walk in the noncommutative garden *Preprint* `arXiv:math.qa/0601054` (2006)
[23] A. Connes, H. Moscovici: The local index formula in noncommutative geometry *GAFA* **5** (1995) 174–243
[24] A. Connes, M. Rieffel: Yang-Mills for non-commutative two-tori *In: Operator Algebras and Mathematical Physics* (Contemp. Math. 62, 1987) 237–266
[25] F. D'Andrea, L. Dąbrowski: Local index formula on the equatorial Podleś Sphere *Lett. Math. Phys.* **75** (2006) 235–254
[26] F. D'Andrea, L. Dąbrowski, G. Landi: The isospectral Dirac operator on the four-dimensional quantum Euclidean sphere *Preprint* `arXiv:math.QA/0611100` (2006)

[27] F. D'Andrea, L. Dąbrowski, G. Landi, E. Wagner: Dirac operators on all Podleś spheres *J. Noncomm. Geom.* (in press) ; `arXiv:math.QA/0606480`

[28] L. Dąbrowski, G. Landi, M. Paschke, A. Sitarz: The spectral geometry of the equatorial Podleś sphere *C. R. Acad. Sci. Paris* **340** (2005) 819–822

[29] L. Dąbrowski, G. Landi, A. Sitarz, W. van Suijlekom, J. C. Várilly: The Dirac operator on $SU_q(2)$ *Commun. Math. Phys.* **259** (2005) 729–759

[30] L. Dąbrowski, G. Landi, A. Sitarz, W. van Suijlekom, J. C. Várilly: The local index formula for $SU_q(2)$ *K-Theory* **35** (2005) 375–394

[31] L. Dąbrowski, A. Sitarz: Dirac operator on the standard Podleś quantum sphere *In: Noncommutative Geometry and Quantum Groups* (Banach Centre Publications 61, P. M. Hajac et al. eds, IMPAN, 2003) 49–58

[32] J. Dixmier: *Les C^*-algèbres et leursreprésentations* (Gauthier-Villars, 1964)

[33] J. Dixmier: Existence de traces non normales *C.R. Acad.Sci. Paris, Ser. A-B* **262** (1966) A1107–A1108

[34] J. M. G. Fell, R. S. Doran: *Representations of*-Algebras, Locally Compact Groups and Banach *-Algebraic Bundles* (AcademicPress, 1988)

[35] T. Friedrich: *Dirac Operators in Riemannian Geometry* (Graduate Studies in Mathematics 25, AMS, 2000)

[36] V. Gayral, J. M. Gracia-Bondia, B. Iochum, T. Schücker, J. C. Varilly: Moyal Planes are spectral triples *Commun. Math. Phys.* **246** (2004) 569–623

[37] P. B. Gilkey: *Invariance Theory, The Heat Equation, And the Atiyah-Singer Index Theorem* (Studies in AdvancedMathematics, CRC Press, 1995)

[38] D. Goswami: Some noncommutative geometric aspects of $SU_q(2)$ *Preprint* `arXiv:math-ph/0108003` (2002)

[39] J. M. Gracia-Bondia, J. C. Várilly, H. Figueroa: *Elements of Noncommutative Geometry* (Birkhauser, 2000)

[40] E. Hawkins, G. Landi: Fredholm modules for quantum Euclidean spheres *J. Geom. Phys.* **49** (2004) 272–293

[41] M. Karoubi: Connexion, courbures et classes caracteristique en K-theorie algebrique *In: Can. Math. Soc. Conf. Proc. 2* (1982, 19–27)

[42] G. Kasparov: Topological invariants of elliptic operators. I. K-homology *Math. URSS Izv.* **9** (1975) 751–792

[43] M. Khalkhali: *Very Basic Noncommutative Geometry* (IPM Lecture Notes Series, IPM, Tehran, 2005)

[44] A. N. Kirillov, N. Yu. Reshetikhin: Representations of the algebra $U_q(sl(2))$, q-orthogonal polynomials and invariants of links *In: Infinite Dimensional Lie Algebras and Groups* (V G Kac ed., World Scientific, 1989) 285–339

[45] A. U. Klimyk, K. Schmüdgen: *Quantum Groups and their Representations* (Springer, 1998)

[46] U. Krähmer: Dirac operators on quantum flag manifolds *Lett. Math. Phys.* **67** (2004) 49–59

[47] G. Landi: *An introduction to NoncommutativeSpaces and their Geometries* (Springer, 1997)

[48] G. Landi, W. van Suijlekom: Principal fibrations from noncommutative spheres *Commun. Math. Phys.* **260** (2005) 203–225

[49] G. Landi, W. van Suijlekom: Noncommutative instantons from twisted conformal symmetries *Commun. Math. Phys.* (in press) `arXiv:math.qa/0601554`;
[50] B. Lawson, M. L. Michelson: *Spin Geometry* (Princeton University Press, 1989)
[51] J. L. Loday: *Cyclic Homology* (Springer, 1998)
[52] S. Lord, A. Sedaev, F. Sukochev: Dixmier traces as singular symmetric functionals andapplications to measurable operators *J. Funct. Anal.* **224** (2005) 72–106
[53] J. Madore: *An Introduction to Noncommutative Differential Geometry and its Physical Applications* (London Mathematical Society Lecture Note Series 257, Cambridge University Press, 1999)
[54] M. Marcolli: *Arithmetric Noncommutative Geometry* (University Lecture Series 36, AMS, 2005)
[55] T. Masuda, Y. Nakagami, S. L. Woronowicz: A C^*-algebraic framework for quantum groups *Int. J. Math.* **14** (2003) 903–1001
[56] S. Neshveyev, L. Tuset: A local index formula for the quantum sphere *Commun. Math. Phys.* **254** (2005) 323–341
[57] P. Podleś: Quantum spheres *Lett. Math. Phys.* **14** (1987) 521–531
[58] M. Reed, B. Simon: *II : FourierAnalysis, Self-Adjointness* (Academic Press, 1975)
[59] A. Rennie, J. C. Várilly: Reconstruction of Manifolds in Noncommutative Geometry *Preprint* `arXiv:math.OA/0610418` (2006)
[60] M. A. Rieffel: C^*-algebras Associatedwith Irrational Rotations *Pacific J. Math.* **93** (1981) 415–429
[61] M. A. Rieffel: The cancellation theorem for projective modules over irrational rotation C^*-algebras *Proc. London Math. Soc.* **47** (1983) 285–302
[62] M. A. Rieffel: Non-commutative Tori: a Case Study of Non-commutative Differential Manifolds *In: Geometric and Topological Invariants of Elliptic Operators* (Contemp. Math. 105, 1990) 191–211
[63] M. Rieffel: *Deformation Quantization for Actions of* $\mathbb{R}^d$ (Memoirs of the American Mathematical Society 506, AMS, 1993)
[64] M. Rieffel: Compact Quantum Metric Spaces *In: Operator algebras, quantization, and noncommutative geometry* (Contemp. Math. 365, 2004) 315–330
[65] W. Rudin: *Real and Complex Analysis* (McGraw-Hill, 1987)
[66] K Schmüdgen, E Wagner: Dirac operator and a twisted cyclic cocycle on the standard Podleś quantum sphere *J. reine angew. Math.* **574** (2004) 219–235
[67] B. Simon: *Trace Ideals and their Applications* (LondonMathematical Society Lecture Notes 35, Cambridge University Press, 1979)
[68] A. Sitarz: Equivariant spectral triples *In: Noncommutative Geometry and Quantum Groups* (Banach Centre Publications 61, P M Hajac et al. eds, IMPAN, 2003) 231–263
[69] R. G. Swan: Vector Bundles and ProjectiveModules *Trans. Am. Math. Soc.* **105** (1962) 264–277
[70] M. Takesaki: *Tomita's Theory of Modular Hilbert Algebras* (Lecture Notes in Mathematics 128, Springer, 1970)
[71] E. G. F. Thomas: *Two Notes on* $*$*-algebras* (Research Note W-9707, Department of Mathematics, University of Groningen, 1997)
[72] J. C. Várilly: *An Introduction to Noncommutative Geometry* (EMS Lectures in Mathematics, European Mathematical Society, 2006)

[73] M. Wodzicki: Local invariants of spectral asymmetry *Inv. Math.* **75** (1984) 143–177

[74] M. Wodzicki: Noncommutative residue. Chapter I: Fundamentals *In: K-theory, Arithmetic and Geometry* (Yu. I. Manin ed., Lectures Notes in Mathematics 1289, Springer, 1987) 320–399

[75] J. A. Wolf: Essential selfadjointness for the Dirac operator and its square *Indiana Univ. Math. J.* **22** (1973) 611–640

[76] S. L. Woronowicz: Compact quantum groups *In: Quantum Symmetries* (Les Houches, Session LXIV, Elsevier Science, 1998) 845–884

GIOVANNI LANDI

Dipartimento di Matematica e Informatica, Università di Trieste,
Via A. Valerio 12/1, I-34127 Trieste, Italy

E-mail: `landi@univ.trieste.it`

Handbook of Global Analysis
Demeter Krupka and David Saunders

De Rham cohomology[1]

M. A. Malakhaltsev

Contents

Introduction

A (co)homology theory is a functor from a subcategory of the category of topological spaces (e. g. the category of manifolds, the category of CW-complexes, etc.) to an algebraic category (e. g. the category of Abelian groups, the category of rings, etc) satisfying additional axioms. For the category of manifolds it is natural to have a cohomology theory which is constructed via the differentiable structure. The de Rham cohomology theory is a classical cohomology theory of this type. For the most important classes of differentiable manifolds, de Rham cohomology coincides with cohomology constructed in a pure topological way, and therefore gives us a possibility to establish relation between the invariants of differential geometric structures on a manifold and the topological invariants of this manifold. This fact results in numerous applications of de Rham cohomology: various characteristic classes, de Rham-like complexes associated to differential structures, etc.

The aim of the present paper is to give a brief introduction to the de Rham cohomology theory and to expose some relevant results in differential geometry. We do not give proofs, however we provide the reader with references to literature where he can find detailed exposition including proofs of results formulated here.

[1]This paper was written during the author's visit to Universidad de Los Andes (Bogota, Colombia). The author is thankful to Departamento de Matemáticas, Facultad de Ciencias de Universidad de los Andes (Bogotá, Colombia), and Grupo Sanford for invitation, hospitality, and financial support during his visit in September-October, 2006. Also he expresses his deep gratitude to professor José Ricardo Arteaga Bejarano for useful cooperation and constant support.

1 De Rham complex

1.1 De Rham complex. De Rham cohomology

1.1.1 The algebra of differential forms on smooth manifold

Let M be an n-dimensional smooth manifold. A *differential form* ω of degree k (k-form) on M is a skewsymmetric tensor field of type $(k,0)$: $p \to \omega(p)(X_1,\dots,X_k)$, where $p \in M$, and $X_1, \dots, X_k$ are vectors in T_pM. Note that a differential 0-form is a function, and, for $k > n$, each differential k-form vanishes.

For any differential k-form ω and m-form η, the *wedge product* $\omega \wedge \eta$ is defined as follows:

$$\omega \wedge \eta = \frac{1}{(k+m)!} \sum_{\sigma \in S_{k+m}} \varepsilon(\sigma)\omega(X_{\sigma(1)},\dots,X_{\sigma(k)})\eta(X_{\sigma(k+1)},\dots,X_{\sigma(k+m)}),$$

where S_{k+m} is the group of all permutations of the set $\{1,\dots,k+m\}$, and $\varepsilon(\sigma)$ is the sign of σ.

Let $\Omega^k(M)$ be the vector space of all differential k-forms on M, and $\Omega(M) = \bigoplus_{k=0}^{n} \Omega^k(M)$. Then the wedge product turns $\Omega(M)$ into a supercommutative associative graded algebra: for $\omega \in \Omega^k(M)$, $\eta \in \Omega^m(M)$, we have $\omega \wedge \eta = (-1)^{km}\eta \wedge \omega$.

With respect to local coordinates x^i, a k-form ω is written as follows:

$$\omega = \sum_{i_1 < i_2 \dots < i_k} \omega_{i_1 \dots i_k}(x^i) dx^{i_1} \wedge \dots \wedge dx^{i_k}.$$

Let M, M' be smooth manifolds, then any smooth map $f : M \to M'$ determines an algebra homomorphism $f^* : \Omega(M') \to \Omega(M)$,

$$f^*\omega'(X_1,\dots,X_k) = \omega(df X_1,\dots,df X_k),$$

where $df : TM \to TM'$ is the differential of f.

1.1.2 Exterior differential

The *exterior differential* $d : \Omega(M) \to \Omega(M)$ is uniquely defined by the following properties:

a) for any k, $d : \Omega^k(M) \to \Omega^{k+1}(M)$ is a linear operator;

b) d is a superderivation of the algebra $\Omega(M)$, i. e. for any $\omega \in \Omega^k(M)$, $\eta \in \Omega^l(M)$,

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta;$$

c) for any smooth map $f : M \to M'$ and $\omega \in \Omega(M')$, $df^*\omega = f^*d\omega$;

e) for any function $f \in \Omega^0(M)$, $df \in \Omega^1(M)$ is the differential of f;

d) $d \circ d = 0$;

The coordinates of $d\omega$ are expressed in terms of the coordinates of ω as follows:

$$(d\omega)_{i_1\dots i_{k+1}} = \sum_{a=1}^{k+1}(-1)^{a+1}\partial_{i_a}\omega_{i_1\dots\widehat{i_a}\dots i_{k+1}},$$

where the hat over index means that this index is dropped. Another formula for calculating the exterior differential $d\omega$ of a k-form $\omega \in \Omega^k(M)$ is

$$\begin{aligned} d\omega(X_1,\dots,X_{k+1}) &= \frac{1}{k+1}\sum_{a=0}^{k+1}(-1)^{a+1}X_a\omega(X_1,\dots,\widehat{X_a},\dots,X_{k+1}) \\ &+ \frac{1}{k+1}\sum_{1\le a<b\le k+1}(-1)^{a+b}\omega([X_a,X_b],X_1,\dots,\widehat{X_a},\dots,\widehat{X_b},\dots,X_{k+1}), \end{aligned} \quad (1)$$

where X_i, $i=\overline{1,k+1}$, are vector fields on M, and the hat again means that the corresponding argument is dropped. A differential form $\omega \in \Omega^k(M)$ is said to be *closed* if $d\omega = 0$, and *exact* if $\omega = d\eta$, $\eta \in \Omega^{k-1}(M)$.

1.1.3 De Rham complex. De Rham cohomology ([6], [20])

By definition, the exterior differential has property $d \circ d = 0$ (see 1.1.2). Hence we get the *de Rham complex of manifold M*:

$$0 \to \Omega^0(M) \xrightarrow{d} \Omega^1(M) \xrightarrow{d} \dots \xrightarrow{d} \Omega^n(M) \to 0$$

and the vector space

$$H^k(M) = H^k_{DR}(M) = \frac{\ker d : \Omega^k(M) \to \Omega^{k+1}(M)}{\operatorname{im} d : \Omega^{k-1}(M) \to \Omega^k(M)} \quad (2)$$

is called the *de Rham cohomology of M in dimension k*.

The algebra $(\Omega(M), \wedge)$ endowed with the exterior differential d is a differential algebra, therefore the de Rham cohomology $H(\Omega, d)$ also is an algebra with respect to the multiplication:

$$[\omega_1] \smile [\omega_2] = [\omega_1 \wedge \omega_2].$$

Thus we get the *de Rham cohomology algebra.*

Let $f : M \to M'$ be a smooth map. Then the inverse image map $f^* : \Omega^k(M') \to \Omega^k(M)$ gives us a chain map $(\Omega^*(M'), d) \to (\Omega^*(M), d)$, $\omega' \to f^*(\omega')$ because $df^* = f^*d$ (see 1.1.2). In addition, $f^* : \Omega(M') \to \Omega(M)$ is an algebra homomorphism (see 1.1.1). Hence, each smooth map $f : M \to M'$ determines an algebra morphism $f^* : H(M') \to H(M)$, thus we get a *functor H_{DR} from the category of smooth manifold to the category of algebras.*

1.1.4 De Rham cohomology with compact support ([6], [20])

We can obtain another functor considering the ring $\Omega_c(M)$ of forms with compact support. For each $\omega \in \Omega^k_c(M)$, we have $d\omega \in \Omega^{k+1}_c(M)$, hence $(\Omega_c(M), d)$ is a subcomplex in the de Rham complex (in fact, $(\Omega_c(M), d)$ is a differential ideal in the differential algebra $(\Omega(M), d)$). The cohomology $H_c(M) = H(\Omega_c(M), d)$ is called the *de Rham cohomology with compact support*. Obviously, if M is compact, we have $\Omega_c(M) = \Omega(M)$, and hence $H_c(M) = H(M)$.

1.1.5 Calculation of de Rham cohomology

In general, it is very difficult to calculate de Rham cohomology using only the definition. This can be done however in some special situations, for example, for a one-dimensional manifold M. The de Rham complex for M is $0 \to \Omega^0(M) \xrightarrow{d} \Omega^1(M) \to 0$. Let M be connected, then either $M = \mathbb{R}$, or $M = \mathbb{S}^1$. In both cases, from (2) it follows that $H^0(M) \cong \mathbb{R}$. If $M = \mathbb{R}$, we have $H^1(M) = 0$ because for any $\omega = g(x)dx$ we can find $f(x) = \int_0^x g(t)dt$ such that $\omega = df$. Now let $M = \mathbb{S}^1$ and x be the angle, then a solution f to the equation $\omega = df$ exists if and only if $\int_0^{2\pi} g(x)dx = 0$. From this follows that $H^1(M) = \mathbb{R}$.

Another case when one can directly calculate $H(M)$ is the case when $M = G/H$ is a compact symmetric space of a connected compact Lie group G. This calculation is based on the following statements: a) any invariant form on M is closed; b) for any closed form $\omega \in \Omega(M)$ a closed invariant form ω' exists such that $\omega' - \omega = d\eta$; c) If $\omega \in \Omega(M)$ is closed and invariant, then the cohomology class of ω is nonzero. From this follows that *the algebra of de Rham cohomology of $M = G/H$ is isomorphic to the algebra of exterior forms on the Lie algebra $\mathfrak{g}$ of G which vanish on the Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$ of H and are invariant with respect to the inner automorphisms ad_h, $h \in H$* (see, e. g., [22]).

1.1.6 Textbooks

The definition and basic properties of the de Rham cohomology are given in almost every textbook on differential geometry, therefore we mention only several books devoted mainly to the de Rham cohomology theory and related questions. First of all, it is the very clearly written classical book [20] by George de Rham himself. In [35] W. Greub, S. Halperin, and R. Vanstone give an exposition of the de Rham cohomology theory starting from the very beginning and avoiding "formal algebraic topology". For a very comprehensive introduction to this subject, we refer the reader to the book [6] by R. Bott and L. Tu, which gives an excellent presentation of algebraic topology via differential forms. The book [87] by I.Vaisman presents a unified exposition of the basic de Rham and harmonic theory used in the representation of the cohomology of differentiable, foliated and complex manifolds by differential forms.

1.2 Basic properties of de Rham cohomology (see, e. g.,[6])

1.2.1 Homotopy invariance of de Rham cohomology

Theorem 1 *Let $f_0, f_1 : M \to M'$ be homotopic smooth maps. Then the corresponding cohomology homomorphisms f_0^* and f_1^*, which map $H(M')$ to $H(M)$, coincide.*

Thus the de Rham cohomology is homotopy invariant, and, *if M is homotopically equivalent to M', then $H(M) \cong H(M')$*. In particular, from this follows

Theorem 2 (Poincaré Lemma) *Let U be a contractible open subset in $\mathbb{R}^n$. Then*

$$H^q(U) \cong H^q(pt) \cong \begin{cases} \mathbb{R}, & q = 0 \\ 0, & q > 0 \end{cases}$$

From this follows that the de Rham cohomology is a *global invariant of manifold*.

The proofs of the Poincare Lemma, as well as the homotopy invariance of the de Rham cohomology, use an algebraic homotopy which can be constructed in different ways (see, e. g., [6] and [14]).

The de Rham cohomology with compact support is not homotopically invariant. However, we have the isomorphism $H_c^*(M \times \mathbb{R}) \underset{\varepsilon^*}{\overset{\pi^*}{\longleftrightarrow}} H_c^{*-1}(M)$ where $\pi(x,t) = x$, $\varepsilon(x) = (x, 0)$, $x \in \mathbb{R}^n$, $t \in \mathbb{R}$. Hence follows

Lemma 2 (Poincaré Lemma for de Rham cohomology with compact support, [6])

$$H_c^q(\mathbb{R}^n) \cong \begin{cases} \mathbb{R}, & q = n \\ 0, & q \neq n \end{cases}$$

1.2.2 Sheaves and de Rham cohomology ([91],[94])

Let M be an n-dimensional smooth manifold. Let us denote by Ω_M^k the sheaf of k-forms on M. Then the exterior differential determines the sheaf morphism $d : \Omega_M^k \to \Omega_M^{k+1}$. Denote by $\mathbb{R}_M$ the sheaf of locally constant functions on M, then we have the sheaf sequence:

$$0 \to \mathbb{R}_M \xrightarrow{i} \Omega^0 \xrightarrow{d} \Omega^1 \xrightarrow{d} \dots \xrightarrow{d} \Omega^n \to 0. \tag{3}$$

From the Poincare lemma it follows that (3) is a fine resolution for the sheaf $\mathbb{R}_M$, hence, by the abstract de Rham theorem (see, e. g. [94]), the *de Rham cohomology $H(M)$ is isomorphic to the Cech cohomology $\check{H}(M; \mathbb{R}_M)$*. Note that this fact can be also proved using the corresponding double complex (see [6], [73]). Since the singular cohomology $H_{sing}(M)$ is isomorphic to the Čech cohomology $\check{H}(M)$ ([9]), the *de Rham cohomology algebra is isomorphic to the singular cohomology algebra*.

Since the de Rham cohomology is isomorphic to the singular cohomology, many constructions known for singular cohomology can be expressed in terms of de Rham cohomology.

1.2.3 Dimension of de Rham cohomology

Assume that an n-dimensional manifold M admits a finite open covering U_α such that each intersection $U_{\alpha_1} \cap U_{\alpha_2} \cap \dots U_{\alpha_k}$ is diffeomorphic to $\mathbb{R}^n$. Then each vector space $H^k(M)$ is finite-dimensional (see 1.2.2). This is true, in particular, for a compact manifold M.

In general, de Rham cohomology is infinite-dimensional. For example, let us take $M = \mathbb{R} \times \mathbb{R} \setminus \mathbb{Z} \times \mathbb{Z}$. Then $\dim H^1(M) = \infty$. In fact, if (x, y) are the coordinates on M, then, for $k, l \in \mathbb{Z}$, the forms $\omega_{k,l} = \frac{-(y-l)dx+(x-k)dy}{(x-k)^2+(y-l)^2}$ represent cohomology classes linearly independent in $H^1(M)$.

1.2.4 Connectedness and de Rham cohomology

Each connected component M_a of a manifold M determines its characteristic function χ_α, which is locally constant, and hence determines a cohomology class $[\chi_\alpha] \in H^0(M)$. In fact, $\{[\chi_\alpha]\}$ is a basis in $H^0(M)$. In particular, M is connected if and only if $H^0(M) \cong \mathbb{R}$.

1.2.5 Orientation and de Rham cohomology

Let M be an n-dimensional manifold. A form $\omega \in \Omega^n(M)$ nonvanishing at each $p \in M$ is called a *volume form*. If on M a volume form ω exists, then M is said to be *orientable*. Note that $d\omega = 0$ by dimension arguments, hence we have the cohomology class $[\omega]$.

Theorem 3 ([6]) *Let M be a smooth n-dimensional manifold.*

a) If M is compact, then M is orientable if and only if $H^n(M) \neq 0$. In this case $H^n(M) \cong \mathbb{R}$ and for each volume form $[\omega] \neq 0$.

b) If M is noncompact, then $H^n(M) = 0$. In particular, for each $\omega \in \Omega^n(M)$ one can find $\theta \in \Omega^{n-1}(M)$ such that $\omega = d\theta$.

1.2.6 Künneth formula

Let M and M' be smooth manifolds, and $\pi : M \times M' \to M$, $\pi' : M \times M' \to M'$ be the projections. Then we have the chain map

$$\Omega^k(M) \times \Omega^l(M') \to \Omega^{k+l}(M \times M'), (\omega, \omega') \to \pi^*\omega \wedge \pi'^*\omega'.$$

This map induces *the isomorphism*

$$H(M) \otimes H(M') \to H(M \times M'), [\omega] \otimes [\omega'] \to [\pi^*\omega \wedge \pi'^*\omega'].$$

1.2.7 Mayer-Vietoris sequence

Let U and V be open subsets in a manifold M such that $M = U \cup V$. Let $i_U : U \hookrightarrow M$, $i_V : V \hookrightarrow M$, $j_U : U \cap V \hookrightarrow U$, $j_V : U \cap V \hookrightarrow U$ be embeddings.

Theorem 4 ([6]) *The sequence of de Rham complexes*

$$0 \to \Omega(M) \xrightarrow{(i_U^*, i_V^*)} \Omega(U) \oplus \Omega(V) \xrightarrow{r} \Omega(U \cap V) \to 0,$$

where $r(\omega, \theta) = j_U^*\omega - j_V^*\omega$, is exact.

The corresponding exact cohomology sequence

$$\cdots \to H^q(M) \to H^q(U) \oplus H^q(V) \to H^q(U \cap V) \to H^{q+1}(M) \to \dots \tag{4}$$

is the Mayer-Vietoris sequence corresponding to $M = U \cup V$.

For the de Rham cohomology with compact support we also have the exact sequence

$$0 \to \Omega_c(U \cap V) \xrightarrow{\delta} \Omega_c(U) \oplus \Omega_c(V) \xrightarrow{s} \Omega_c(M) \to 0,$$

where $\delta(\omega) = (-j_U^*\omega, j_V^*\omega)$ and $s(\omega, \theta) = \omega + \theta$, hence follows the Mayer-Vietoris sequence for de Rham cohomology with compact support given by (4) with H replaced by H_c.

2 Integration and de Rham cohomology. De Rham currents. Harmonic forms [20]

2.1 Integration and de Rham cohomology

Connection between integration and de Rham cohomology is based on well-known

Theorem 5 (Stokes Theorem) *Let M be an n-dimensional oriented manifold with boundary ∂M (possibly $\partial M = \emptyset$). Then, for each $\omega \in \Omega_c^{n-1}(M)$,*

$$\int_M d\omega = \int_{\partial M} \omega$$

From the Stokes Theorem it follows that the following pairing is defined correctly:

$$H_c^k(M) \times H^{n-k}(M) \to \mathbb{R}, ([\omega],[\theta]) \to \int_M \omega \wedge \theta$$

For a compact manifold M, this pairing is nondegenerate and determines the Poincare duality isomorphism $H^k(M) \cong H^{n-k}(M)$.

2.2 De Rham currents [20]

Let M be an n-dimensional smooth manifold. A *current on M of degree k* is a continuous linear functional $T : \Omega^k(M) \to \mathbb{R}$ (we say that T is *continuous*, if for each sequence of forms φ_i, $i = 1, 2, \dots$ whose supports lie in a compact set covered by a coordinate system and any derivative of any φ_i uniformly converges to zero as $i \to \infty$, we have $T(\varphi_i) \to 0$). Let us denote the set of all currents of degree k on M by $\mathcal{E}^k(M)$.

Let U be an open subset in M. A current T is said to be zero in U if $T(\varphi) = 0$ for any form φ with compact support lying in U. If a current T vanishes at a neighborhood of each point of an open $U \subset M$, then $T = 0$ in U. Hence follows that each current T has a maximal open set where $T = 0$, and the complement of this set is called the *support* of T. We denote by $\mathcal{E}_c(M)$ the space of currents with compact support.

Example 2.1 a) Any $\alpha \in \Omega^k(M)$ defines the $(n-k)$-current: $\alpha(\varphi) = \int_M \alpha \wedge \varphi$, $n = \dim M$; b) Any k-chain c in M defines the k-current: $c(\varphi) = \int_c \varphi$; c) Any contravariant k-vector v at a point $p \in M$ defines the current $v(\varphi) = v^{i_1 \dots i_k} \varphi_{i_1 \dots i_k}$. Thus currents simultaneously describe forms and chains.

The external differential $d : \Omega^k(M) \to \Omega^{k+1}(M)$ determines the boundary operator $b : \mathcal{E}^{k+1}(M) \to \mathcal{E}^k(M)$ by the equality $bT(\varphi) = T(d\varphi)$. Thus we get a complex $(\mathcal{E}_k(M), b)$, and its homology is called the *de Rham homology of M*.

Theorem 6 ([20]) *Let M be a smooth manifold.*

a) Let $T \in \mathcal{E}^k(M)$, and $bT = 0$. Then a C^∞-form $\alpha \in \Omega^k(M)$ exists such that $T - \alpha = bT'$, $T' \in \mathcal{E}^{k+1}(M)$. If $T \in \mathcal{E}_c^k(M)$, then T' can also be taken in $\mathcal{E}_c^{k-1}(M)$.

b) If $\alpha \in \Omega^k(M)$ is such that $\alpha = bT$, $T \in \mathcal{E}^{k+1}(M)$, then $\alpha = d\beta$, where $\beta \in \Omega^{k-1}(M)$. And if the current T lies in $\mathcal{E}_c^{k+1}$, then one can take β also lying in $\Omega_c^{k-1}(M)$.

Using this theorem, G. de Rham obtained duality theorems for arbitrary (compact and noncompact) manifolds (see details in [20]).

2.3 Harmonic forms ([20], [91], [94])

Let g be a Riemannian metric on an oriented compact n-dimensional smooth manifold M, and ω be the corresponding volume form. The metric g determines the *Hodge operator*: $* : \Omega^k(M) \to \Omega^{n-k}(M)$ defined with respect to local coordinates by

$$(*\eta)_{i_1 \ldots i_{n-k}} = \omega_{i_1 \ldots i_{n-k}}{}^{j_1 \ldots j_k} \eta_{j_1 \ldots j_k}.$$

One can prove that $** = (-1)^{k(n-k)}$.

Now we set $\delta : \Omega^k(M) \to \Omega^{k-1}(M)$, $\delta = (-1)^{n(k+1)+1} * d*$ (on 0-forms δ is zero). The Laplace-Beltrami operator $\Delta : \Omega^k(M) \to \Omega^k(M)$ is defined as

$$\Delta = \delta d + d\delta.$$

It is clear that, if $M = \mathbb{R}^n$ and g is the standard Euclidean metric, then Δ is the standard Laplacian $(-1)^n \sum_{i=1}^{n} \partial^2/\partial^2 x^i$.

On $\Omega^k(M)$ consider the scalar product

$$(\alpha, \beta) = \int_M \alpha \wedge *\beta,$$

and denote by $||\alpha||$ the corresponding norm. The operator δ is adjoint of d on $\Omega(M)$ with respect to this scalar product: $(d\alpha, \beta) = (\alpha, d\beta)$, and the Laplace-Beltrami operator Δ is self-adjoint: $(\Delta\alpha, \beta) = (\alpha, \Delta\beta)$.

A form $\alpha \in \Omega^k(M)$ is called *harmonic* if $\Delta\alpha = 0$, and let us denote by $\mathcal{H}^k(M, g)$ the set of all harmonic k-forms. One can easily prove that $\alpha \in \Omega^k(M)$ is harmonic if and only if $d\alpha = 0$ and $\delta\alpha = 0$.

Theorem 7 (Hodge's decomposition theorem (see, e. g., [91])) *Let M be an oriented compact n-dimensional smooth manifold M, g be a Riemannian metric on M, and Δ be the corresponding Laplace-Beltrami operator.*

For each $0 \le k \le n$, the space $\mathcal{H}^k(M)$ is finite dimensional, and we have the following orthogonal direct sum decomposition:

$$\Omega^k(M) = d(\Omega^{k-1}(M)) \oplus \delta(\Omega^{k+1}(M)) \oplus \mathcal{H}^k(M).$$

Consequently, the equation $\Delta\alpha = \beta$ has a solution $\alpha \in \Omega^k(M)$ if and only if β is orthogonal to the space of harmonic k-forms.

Let us denote by $\mathcal{H}$ the projection of $\Omega^k(M)$ onto $\mathcal{H}^k(M)$. The *Green operator* is $G : \Omega^k(M) \to (\mathcal{H}^k(M))^\perp$, $G(\alpha)$ equals the unique solution of the equation $\Delta\omega = \alpha - \mathcal{H}(\alpha)$ in $(\mathcal{H}^k(M))^\perp$. For each $\alpha \in \Omega^k(M)$, we have

$$\alpha = d\delta G\alpha + \delta d G\alpha + \mathcal{H}\alpha.$$

From this expansion it follows

Theorem 8 *Each de Rham cohomology class of an oriented compact manifold M endowed by a Riemannian metric g contains a unique harmonic representative.*

The operator Δ is elliptic and from the theory of elliptic operators it follows that *on a compact manifold M endowed with a Riemannian metric g the space of harmonic forms is finite-dimensional.* This gives another way to prove that *the de Rham cohomology of a compact manifold is finite-dimensional*.

Remark 2.1 The results concerning the fact that the Laplacian is an elliptic operator and hence has kernel of finite dimension are generalized in the theory of elliptic complexes (see, e. g., [83]).

Remark 2.2 The fact that harmonic forms representing the de Rham cohomology classes of a manifold M are determined via a Riemannian metric g on M gives possibility to establish relation between the geometrical properties of g and the topological properties of M. This idea lies in the base of so-called "Bochner technique" (see the classical book by S.Bochner and K. Yano [96]), which is widely used in the modern Riemannian geometry (see e. g. J.P. Bourguignon's paper devoted to Weitzenböck formulas in [4], and S.E.Stepanov's recent survey on Bochner technique and its applications in differential geometry [82]).

3 Generalizations of the de Rham complex

The de Rham complex can be generalized in various ways. One can take analytical or locally integrable forms instead of smooth ones (see 3.1, 3.2 below), or to construct a complex in a purely algebraic way from the algebra of smooth functions (see 3.2.1) whose cohomology coincide with the de Rham cohomology. Another way is to construct a complex of de Rham type for a space generalizing smooth manifolds (see 3.3.1, 3.4).

3.1 Analytic de Rham complex

One may ask if we can take forms of another differentiability class to construct the de Rham complex. Certainly, if we take finite differentiable forms, the de Rham complex is not correctly defined since the exterior differential lows the differentiability class. However, if M is a real analytic manifold, then we can consider the analytic version of the de Rham complex. In [66] B. Malgrange remarked that the de Rham cohomology of a real-analytic manifold can be computed via real-analytic forms. In [3] L. Beretta re-proved this fact using results from [2] and the property that the sheaves of germs of real-analytic forms are locally free and coherent.

3.2 L_p-cohomology

Let M be an n-dimensional Riemannian manifold. For each measurable differential k-form ω whose module is locally integrable, let us define $d\omega$ by the equality

$$\int_M \omega \wedge du = (-1)^{k+1} \int_M d\omega \wedge u$$

for any smooth differential $(n-k-1)$-form u with compact support lying in an oriented domain of the interior of M.

Denote by $L_p^k(M)$ the set of all k-forms ω such that

$$||\omega||_{L_p^k(M)} = (\int_M |\omega(x)|^p dx)^{\frac{1}{p}} < \infty,$$

and by $W_p^k(M)$ the set of all $\omega \in L_p^k$ such that $d\omega \in L_p^{k+1}(M)$. The cohomology H_p^k of the complex (W_p^*, d) is called the *L_p-cohomology of Riemannian manifold M*.

The L_p-theory of differential forms plays an important role in the analysis on non-compact Riemannian manifolds as well as in the study of asymptotic invariants of such manifolds. In [52] V.I. Kuz'minov and I.A. Shvedov give a survey of papers on this subject, where they discussed the following problems: (1) approximation of L_p-forms by compactly-supported forms, and the same problem restricted to the class of closed forms; (2) computation of L_p-cohomology of warped products and, in particular, the Künneth formula and duality theorems; (3) normal solvability of the exterior differentiation operator.

The relation between properties of L_p-cohomology and the geometrical properties of the Riemannian metric g is intensively studied. For example, in [97] N. Yeganefar studied the L^p-cohomology of the complete manifolds of finite volume and pinched negative curvature and proved the following statements.

Theorem 9 *Let (M^n, g) be a complete n-dimensional manifold of finite volume and pinched negative curvature $-1 \leq K \leq -a^2 < 0$. Assume that $p \geq 1$ is a real number, and k an integer such that $k > (n+1+2(p-1)a)/(1+(p-1)a)$. Then we have the isomorphism $H_p^k(M) \simeq H_c^k(M)$, where $H_c^k(M)$ denotes the compact supported cohomology of M.*

Theorem 10 *Let (M^n, g) be a conformally compact n-dimensional manifold. Assume that $p \geq 1$ is a real number and k an integer such that $k < (n-10)/p$. Then we have the isomorphism $H_p^k(M) \simeq H_c^k(M)$.*

Note also the paper [12], where Gilles Carron, Thierry Coulhon and Rew Hassell studied L_p-cohomology for manifolds with Euclidean ends.

3.2.1 De Rham cohomology and Hochschild cohomology

Let $\mathcal{A}$ be a (possibly non-commutative) algebra over $\mathbb{C}$, and $\mathcal{M}$ be a bimodule over $\mathcal{A}$. Let $\mathcal{A}^e = \mathcal{A} \otimes \mathcal{A}^o$ be the tensor product of $\mathcal{A}$ and its opposite algebra. Then $\mathcal{M}$ becomes a left $\mathcal{A}^e$-module. Denote by $C^n(\mathcal{A}, \mathcal{M})$ the space of n-linear maps from $\mathcal{A}$ to $\mathcal{M}$, and let $b : C^n(\mathcal{A}, \mathcal{M}) \to C^{n+1}(\mathcal{A}, \mathcal{M})$, $T \to bT$, be given by

$$\begin{aligned}(bT)(a_1, \ldots, a_n, a_{n+1}) = a_1 T(a_2, \ldots, a_{n+1}) \\ + \sum_{i=1}^{n} (-1)^i T(a_1, \ldots, a_i a_{i+1}, \ldots, a_{n+1}) \\ + (-1)^{n+1} T(a_1, \ldots, a_n) a_{n+1}. \quad (5)\end{aligned}$$

We have $bb = 0$, and the complex $(C^*(\mathcal{A}, \mathcal{M}), b)$ is called the *Hochschild complex* of the algebra $\mathcal{A}$, and the cohomology of this complex the *Hochschild cohomology* of $\mathcal{A}$ with coefficients in $\mathcal{M}$.

The space $\mathcal{A}^*$ of all linear functionals on $\mathcal{A}$ is an $\mathcal{A}$-bimodule by $(a\varphi b)(c) = \varphi(bca)$. Therefore, we have the Hochschild complex $C^*(\mathcal{A}, \mathcal{A}^*)$ and the corresponding cohomology. Any element $T \in C^n(\mathcal{A}, \mathcal{A}^*)$ can be considered as an $n+1$-linear functional τ on $\mathcal{A}$:

$\tau(a_0, a_1, \ldots, a_n) = T(a_1, \ldots, a_n)(a_0)$. Let us denote by $C^n_\lambda(\mathcal{A})$ the space of cochains in $C^{n+1}(\mathcal{A}, \mathcal{A}^*)$ such that, for any *cyclic* permutation σ of $\{0, 1, \ldots, n\}$, $\sigma\tau = \mathrm{sgn}(\sigma)\tau$. The Hochschild coboundary b does not commute with the cyclic permutations, however it maps $C^n_\lambda(\mathcal{A})$ to $C^{n+1}_\lambda(\mathcal{A})$. Therefore, we obtain a subcomplex $(C^*_\lambda(\mathcal{A}), b)$ of the Hochschild complex, and its cohomology is called the *cyclic cohomology* of $\mathcal{A}$. Any algebra homomorphism $\rho : \mathcal{A} \to \mathcal{B}$ induces a cochain map $\rho^* : C^n_\lambda(\mathcal{B}) \to C^n_\lambda(\mathcal{A})$, $(\rho^*\varphi)(a_0, \ldots, a_n) = (\rho(\varphi(a_0), \ldots, \rho(a_n))$ and hence the cohomology map $H^n_\lambda(\mathcal{B}) \to H^n_\lambda(\mathcal{A})$.

The construction of Hochschild and cyclic cohomology can be transferred (with the corresponding changes) to the case of topological algebra with topology given by a system of seminorms (see [18] for details). Now let M be a compact manifold, and $\mathcal{A} = C^\infty(M)$ be the algebra of smooth functions endowed by the Frechet space topology defined by the family of seminorms $p_n(f) = \sup\limits_{\alpha \le n} |\partial_\alpha(f)|$ using local charts on M. In [18] A. Connes proves the following result:

Theorem 11 *a) The continuous Hochschild cohomology group $H^k(\mathcal{A}, \mathcal{A}^*)$ is isomorphic to the space of de Rham currents (see 2.2) of dimension k on M. To the $k+1$-linear functional φ is associated the current T such that*

$$T(f_0 df_1 \wedge df_2 \wedge \cdots \wedge df_k) = \sum_{\sigma \in \Sigma_{k+1}} \mathrm{sgn}(\sigma)\varphi(f_{\sigma(0)}, f_{\sigma(1)}, \ldots, f_{\sigma(k)}).$$

b) For each k, $H^k_\lambda(\mathcal{A})$ is isomorphic to the direct sum

$$\ker b(\subset \mathcal{D}^k) \oplus H_{k-2}(M; \mathbb{C}) \oplus H_{k-4}(M; \mathbb{C}) \oplus \ldots,$$

where $H_k(M : \mathbb{C})$ is the de Rham homology of M.

c) $H^*(\mathcal{A})$ is isomorphic to the de Rham homology $H(M; \mathbb{C})$.

This result demonstrates that de Rham complex on the spectrum of an algebra $\mathcal{A}$ can be expressed in terms of Hochschild cohomology of $\mathcal{A}$. This idea was widely used in the theory of deformation quantization (see, e.g., [8], [21], [93]). Also this result was generalized to other algebras, see, e. g. [39], where the authors construct an isomorphism between the de Rham cohomology of a manifold M and the relative Hochschild cohomology of the algebra of differential operators on M. For the methods of computing the cyclic cohomology for algebras of smooth functions on an orbifold see [92].

3.3 De Rham cohomology for generalized differential structures

The problem to define the de Rham cohomology for generalized spaces has been solved in many ways. One of the first papers devoted to this problem is [81]. Here we describe an approach to generalized differentiable structures, which was proposed by M.V. Losik in [59], [60].

3.3.1 M.V.Losik's R^n-sets. De Rham cohomology of R^n-sets

Let R_n be a category whose objects are open sets in $\mathbb{R}^n$. For objects U, V, a morphism between U and V is a diffeomorphism of U onto an open $W \subset V$. Let Z be a set. An R_n-chart on Z is a pair (U, k), where $U \in Obj(R_n)$ and $k : U \to Z$ is a map. A collection $\Phi = \{(U_\alpha, k_\alpha)\}$ of R_n-charts is called an R_n*-atlas* if $Z = \cup U_\alpha$. Each $\mathbb{R}_n$-atlas Φ determines a category A_Φ such that $Obj(A_\Phi) = \Phi$ and morphisms between objects (U_1, k_1) and (U_2, k_2) are the morphisms $m : U_1 \to U_2$ of R_n with property $k_1 = k_2 \circ m$. Then we have the covariant functor $I_\Phi : A_\Phi \to Sets$, $I_\Phi(U, k) = U$, $I_\Phi(m) = m$. An *atlas of* R_n*-set on a set* Z is an R_n-atlas on Z such that $\varinjlim I_\Phi = Z$. Any atlas of R_n-set determines a maximal atlas of R_n-set, and a set Z endowed with a maximal atlas is called an R_n*-set*. Let Z_1, Z_2 be R_n-sets endowed with maximal atlases Φ_1, Φ_2. A map $h : Z_1 \to Z_2$ is called a *morphism of* R_n*-sets* if for each R_n-chart $(U, k) \in \Phi_1$, the R_n-chart $(U, h \circ k)$ lies in Φ_2. Thus we obtain the category of R_n-sets.

Let us give several examples of R_n-sets. 1) A maximal atlas φ on a smooth manifold M determines an atlas of R_n-set, hence any manifold is an R_n-set. Any smooth map of manifolds is a morphism between R_n-sets, therefore the category of manifolds is a subcategory in the category of R_n-sets. 2) Any orbifold is an R_n-set. 3) Any manifold M endowed by a foliation $\mathcal{F}$ of codimension q admits a structure of R_q-set. Any foliated chart (U, x^i, x^α) (this means that, for each leaf L, the connected components of $U \cap L$ are given by equations $x^\alpha = 0$) determines a submersion $\varphi : U \to V \subset \mathbb{R}^q$, $p \to (x^1(p), \ldots, x^q(p))$. Take $s : V \to U$ such that $\phi \circ s = \mathbb{I}d_V$, then (V, s) is an R_n-chart and, taking an atlas of foliated charts on M, in this way we can construct an atlas of R_n-set on M. 4) Any quotient space M/Γ of an n-dimensional manifold M by an action of a pseudogroup Γ is an R_n-set.

Now let us consider a covariant functor $F : \mathcal{M}an \to \mathcal{S}ets$ from the category of manifolds to the category of sets. Let us set $F(Z) = \lim_{\leftrightarrow} F \circ I_\Phi$, thus we get extension of F to the category of R_n-sets.

Let $F = \Omega^p$ be the functor which maps each manifold M to the set $\Omega(M)$, and smooth map $f : M \to M'$ to the map $f^* : \Omega(M') \to \Omega(M)$. Then, for an R_n-set Z with atlas Φ, we obtain $\Omega(Z) = \varprojlim \Omega \circ I_\Phi$. Since $\Omega(M)$ is a vector space for any M, and f^* is linear, $\Omega(Z)$ is also a vector space. Moreover, the exterior product $\wedge$ and the exterior differential d determine the natural transformations of functors: $(\Omega^p, \Omega^q) \to \Omega^{p+q}$ and $\widetilde{d} : \Omega^p \to \Omega^{p+1}$. This gives the wedge product $\wedge : \Omega^p(Z) \times \Omega^q(Z) \to \Omega^{p+q}(Z)$, and the differential $d : \Omega^p(Z) \to \Omega^{p+1}(Z)$ which is compatible with the wedge product, and $d^2 = 0$. Thus we obtain the de Rham complex for the R_n-set Z. In the same way we can define the singular homology $H_*(Z)$ of an R_n-set Z and the paring $H^*(Z) \times H_*(Z) \to \mathbb{R}$.

This gives a possibility, in a unique way, to define the de Rham cohomology for orbifolds, spaces of leaves of foliations, diffeological spaces, quotients of manifold by action of discrete groups, and other spaces with generalized differentiable structures. Also M.V. Losik defined characteristic classes of $\mathbb{R}_n$-sets in terms of the de Rham cohomology.

3.4 Stratified de Rham cohomology

The problem of defining suitable complexes of differential forms on stratified spaces has been studied by many authors. We mention only several results along this lines. A. Verona [89], [90] proved a de Rham theorem for controlled forms in the context of an abstract stratification. In [30] M. Ferrarotti introduced a complex of infinitesimally controlled forms, containing the Verona's complex. In [23] C.-O. Ewald studies the de Rham cohomology of a class of spaces with singularities which are called stratifolds, and generalizes two classical results from the analysis of smooth manifolds to the class of stratifolds. The first one is Stokes theorem, the second one is the de Rham theorem which states that the de Rham cohomology of a stratifold is isomorphic to its singular cohomology with coefficients in R. He gives an explicit geometric construction of this isomorphism by integrating forms over stratifolds. The relation between the Hochschild homology and the de Rham cohomology for stratifolds (this relation for manifolds is described in 3.2.1) is studied by C.-O. Ewald in [24].

4 Equivariant de Rham cohomology

Let X be a topological space endowed by an action of a compact Lie group G. In algebraic topology the definition of equivariant cohomology group $H_G(X)$ is motivated by the principle that, for a free G-action, $H_G(X) = H(X/G)$. The standard way to define equivariant cohomology is to take a contractible E on which G acts freely, and to set $H_G(M) = H((X \times E)/G)$. This definition is correct, i. e. does not depend on the choice of E. Also, for each compact Lie group, one can find a contractible space E such that G acts freely on E. In this situation one cannot use de Rham cohomology, because a compact Lie group cannot act freely on a finite dimensional contractible manifold. Therefore H. Cartan found a special algebraic construction to deal with this problem. In [38] the reader can find the modern exposition of the theory of equivariant cohomology based on superalgebra language.

The main idea is to construct an "algebraic analog of the space $M \times E$". Let M be a manifold on which a compact Lie group G acts. Let $\mathfrak{g}$ be the Lie group of G, and $\sigma : \mathfrak{g} \to \mathfrak{X}(M)$, where $\sigma(v)(p) = \frac{d}{dt}\big|_{t=0} \exp(tv)(p)$, $p \in M$, be the Lie algebra homomorphism determined by the G-action on M [45]. Then $\mathfrak{g}$ acts on $\Omega(M)$ via the interior product $\iota_v : \Omega^k(M) \to \Omega^{k-1}(M)$, $\iota_v\omega = i_{\sigma(v)}\omega$. Also, we have the exterior differential $d : \Omega^k(M) \to \Omega^{k+1}(M)$, and the Lie derivative $L_v = \mathcal{L}_{\sigma(v)} : \Omega^k(M) \to \Omega^k(M)$, $v \in \mathfrak{g}$. These operations satisfy the *Weil equations*:

$$\begin{aligned} &\iota_v\iota_w + \iota_w\iota_v = 0, L_v\iota_w - \iota_w L_v = \iota_{[v,w]}, L_v L_w - L_w L_v = L_{[v,w]}, \\ &d\iota_v + \iota_v d = L_v, dL_v - L_v d = 0, d^2 = 0. \end{aligned} \tag{6}$$

In addition, the G-action on M induces the G-action ρ on $\Omega(M)$, and we have

$$\rho_g \circ L_v \circ \rho_g^{-1} = L_{Adgv}, \rho_g \circ i_v \circ \rho_g^{-1} = i_{Adgv}, \forall g \in G, \forall v \in \mathfrak{g}.$$

This data can be expressed in superalgebra terms. A *Z-graded superalgebra* is an Z-graded algebra $\mathfrak{g} = \oplus\mathfrak{g}_i$ with multiplication $[\ ,\ ]$ such that $[\mathfrak{g}_i, \mathfrak{g}_j] \subset \mathfrak{g}_{i+j}$ and

$$[v_i, v_j] + (-1)^{ij}[v_j, v_i] = 0, v_i \in \mathfrak{g}_i. \tag{7}$$

A *derivation D of a Z-graded superalgebra A* is a linear map $D : A \to A$ such that $D(ab) = D(a)b + (-1)^{mk}aD(b)$, where $a \in A_m$, and k is called the *degree* of D. Let $Der_k(A)$ be the set of derivations of degree k. For superalgebras A and B, the tensor product $A \otimes B$ is also a superalgebra with respect to the product law $(a_1 \otimes b_1) \cdot (a_2 \otimes b_2) = (-1)^{ij} a_1 a_2 \otimes b_1 b_2$, where $\deg a_2 = i$, $\deg b_1 = j$.

Any Lie algebra $\mathfrak{g}$ determines a Z-graded Lie superalgebra $\widetilde{\mathfrak{g}} = \mathfrak{g}_{-1} \oplus \mathfrak{g}_0 \oplus \mathfrak{g}_1$. Here the Lie algebra $\mathfrak{g}_0$ is isomorphic to the Lie algebra $\mathfrak{g}$, $\mathfrak{g}_{-1}$ is isomorphic to $\mathfrak{g}$ as vector space, $\mathfrak{g}_1$ is a one-dimensional space. Since $\widetilde{\mathfrak{g}}$ is a Z-graded algebra, we have $[\mathfrak{g}_{-1}, \mathfrak{g}_{-1}] = [\mathfrak{g}_1, \mathfrak{g}_1] = 0$. Denote elements of g_{-1} by ι_a, elements of $\mathfrak{g}_0$ by L_a, $a \in \mathfrak{g}$, and elements in $\mathfrak{g}_1$ by d. Then, we have $[L_v, L_w] = L_{[v,w]}$ and set $[L_v, \iota_w] = \iota_{[v,w]}$, $[L_v, \delta] = 0$.

An action of G on M determines a representation of $\widetilde{\mathfrak{g}}$ on the commutative superalgebra $\Omega(M)$ by derivations: $L_v \to \mathcal{L}_{\sigma(v)}$, $\iota_v \to i_{\sigma(v)}$, and a basic element $d \in \mathfrak{g}_1$ goes to the exterior differential d.

G^*-algebras and modules. A *G^*-algebra* is a commutative superalgebra endowed with a representation $\rho : G \to Aut(A)$ and a representation $\widetilde{\mathfrak{g}} = \mathfrak{g}_{-1} \oplus \mathfrak{g}_0 \oplus \mathfrak{g}_1 \to Der(A)$, where we denote the derivations corresponding to L_v, ι_v, and d by the same symbols, and

$$\begin{aligned} &\frac{d}{dt}\rho(\exp tv)\Big|_{t=0} = L_v \qquad \rho(g) \circ L(v) \circ \rho(g^{-1}) = L_{Adgv}, \\ &\rho(g) \circ i_v \circ \rho(g^{-1}) = i_{Adgv}, \qquad \rho(g) d\rho(g^{-1}) = d \end{aligned} \tag{8}$$

A *G^*-module* is a supervector space A together with a linear representation $\rho : G \to Aut(A)$ and a homomorphism $\widetilde{g} \to End(A)$ such that (8) hold. A G^* module A is said to satisfy *condition* (C) if there exists a linear map $\theta : \mathfrak{g}^* \to A_1$ such that $\theta(v^*)(\iota_w) = v^*(w)$ for all $v^* \in \mathfrak{g}^*$, $w \in \mathfrak{g}$, and $\theta(Ad(g)a) = \rho(g)\theta(v)$ for all $g \in G$, $v \in \mathfrak{g}$. Any G^* algebra A has derivation $d : A \to A$, the image of a nonzero element in $\mathfrak{g}_1$, and since $[\mathfrak{g}_1, \mathfrak{g}_1] = 0$, from the commutativity relation (7) we get $d^2 = 0$. Then, we set $H(A) = H(A, d)$. For a G^* module A, we denote by A_{bas} the set of G-invariant elements $a \in A$ such that $i_v a = 0$. Since, for each $a \in A_{bas}$, $L_v a = 0$, and hence $di_v = -i_v d$ (see (6)), we have $d(A_{bas}) \subset A_{bas}$. Thus we set $H_{bas}(A) = H(A_{bas}, d)$.

Equivariant cohomology. Let E be a G^* algebra which is acyclic ($H(A) = 0$) and satisfies condition (C). Then for any G^* algebra A, we set $H_G(A) = H((A \otimes E)_{bas}, d)$. This definition does not depend on a choice of E. For a G^* algebra morphism $\varphi : A \to B$, we have the morphism $\varphi \otimes \mathrm{id} : A \otimes E \to B \otimes E$, which induces the cohomology morphism $\varphi_G : H_G(A) \to H_G(B)$.

Now return to a manifold M endowed by an action of a Lie group G. In this case the superalgebra $\Omega(M)$ is a G^*-algebra.

Theorem 12 ([38]) *Let G be a compact Lie group acting on a smooth manifold M, then $H_G(M) = H_G(\Omega(M))$.*

It remains to construct, for any compact group G, an acyclic G^* algebra W. The algebra defined below is called the *Weil algebra*.

We set $W(\mathfrak{g}) = \Lambda(\mathfrak{g}) \otimes S(\mathfrak{g})$, and, for each $\theta \in \Lambda^k(\mathfrak{g})$, $\deg(\theta) = k$, and, for each $s \in S^m(\mathfrak{g})$, $\deg(s) = 2m$. So we set $W^{i,2j}(\mathfrak{g}) = \Lambda^i(\mathfrak{g}) \otimes S^{2j}(\mathfrak{g})$ Then, with respect to the natural multiplication

$$(\theta_1 \otimes s_1) \cdot (\theta_2 \otimes s_2) = (\theta_1 \wedge \theta_2) \otimes (s_1 \overset{s}{\otimes} s_2), \tag{9}$$

W is a superalgebra. The *Koszul differential* $d_k : W(\mathfrak{g}) \to W(\mathfrak{g})$ is uniquely defined by

$$d_K(x \otimes 1) = 1 \otimes x, d_K(1 \otimes x) = 0, \forall x \otimes 1, \tag{10}$$

and, evidently, $d_K^2 = 0$. Consider $Q : W(\mathfrak{g}) \to W(\mathfrak{g})$ determined by $Q(x \otimes 1) = 0$, $Q(1 \otimes x) = x \otimes 1$, then $[Q, d_K]t = (k + l)t$ for each $t \in W^{i,2j}(\mathfrak{g})$. Hence follows that (W, d) is acyclic.

The adjoint representation $ad : G \to GL(\mathfrak{g})$ extends to the action $\rho : G \to Aut(G)$, and d_K is clearly G-invariant. Now let us define an action of $\widetilde{\mathfrak{g}}$ on W. We set

$$L_v(\theta\otimes 1) = L_v\theta\otimes 1, L_v(1\otimes\theta) = 1\otimes L_v\theta, \iota_v(\theta\otimes 1) = \theta(v), \iota_v(1\otimes\theta) = L_v\theta\otimes 1. \quad (11)$$

where $L_v\theta(w) = -ad(v)\theta(w) = -\theta([v,w])$. Finally, the basic element of $\mathfrak{g}_1$ is represented by d_K.

Note that the G^* algebra W satisfies condition (C) since we have map $\mathfrak{g}^* \to W_1$, $\theta\otimes 1 \to \theta$. Thus, we have constructed the required acyclic G^* algebra which is used in the definition of equivariant cohomology.

Another definition of equivariant cohomology of a manifold endowed by an action of a Lie group was proposed by M. V. Losik (see [61], [62]).

For a manifold M endowed by an action of a group G, let $C^{p,q} = C^{p,q}(G,\Omega(M))$ be the space of smooth maps from G^p to $\Omega^q(M)$ for $p > 0$, and $C^{0,q} = \Omega^q(M)$. Define the maps $\delta' : C^{p,q} \to C^{p+1,q}$ and $\delta'' : C^{p,q} \to C^{p,q+1}$ as follows

$$\begin{aligned}(\delta' c)(g_1,\dots,g_{p+1}) &= g_1 c(g_2,\dots,g_{p+1})\\ &+ \sum_{i=1}^{p} c(g_1,\dots,g_i g_{i+1},\dots,g_{p+1}) + (-1)^p c(g_1,\dots,g_p),\\ (\delta'' c)(g_1,\dots,g_p) &= (-1)^p dc(g_1,\dots,g_p), \quad (12)\end{aligned}$$

where $c \in C^{p,q}$, $g_1,\dots,g_p \in G$, and d is the exterior derivative. The cohomology of the bicomplex $C(G,\Omega(M)$ endowed by the total coboundary operator $\delta' + \delta''$ is called the *equivariant cohomology of the G-manifold M*.

In [61] M.V.Losik used the spectral sequence for the bicomplex $(C(G,\Omega(M)),\delta',\delta'')$ in order to construct characteristic classes which give as partial cases the characteristic classes of A.G. Rejman, M.A. Semenov-Tyan-Shanskij, and L.D. Faddeev [25] for the automorphism group of a smooth principal fiber bundle and the characteristic classes of R. Bott [5] for the diffeomorphism group of a manifold. In [62], for a compact G, he obtains a generalization of the Cartan theorem on the cohomology of a homogeneous space.

5 Complexes of differential forms associated to differential geometric structures

The de Rham cohomology construction sets the pattern for constructions of complexes associated with various differential geometrical structures. In this section we give several examples of such complexes of de Rham type.

5.1 Dolbeaux cohomology ([36], [87], [94])

Let (M,J) be an n-dimensional complex manifold. The complexified cotangent bundle $T^*_{\mathbb{C}}M = T^*M\otimes\mathbb{C}$ splits into the direct sum of subbundles: $T^*_{\mathbb{C}}M = T^{1,0}M \oplus T^{0,1}M$, where $T^{1,0}M$ is locally spanned by dz^i, and $T^{0,1}M$ by $d\bar{z}^i$, $i = \overline{1,n}$. Let $\Omega_{\mathbb{C}}(M)$ be the algebra of complex-valued differential forms, and we say that $\alpha \in \Omega_{\mathbb{C}}M$ has type (p,q) if locally

$$\alpha = \alpha_{i_1\dots i_p j_1\dots j_q} dz^{i_1}\wedge\dots\wedge dz^{i_p}\wedge d\bar{z}^{j_1}\wedge\dots\wedge d\bar{z}^{j_q}$$

It is clear that $\Omega^k(M) = \oplus_{p+q=k}\Omega^{p,q}(M)$, where $\Omega^{p,q}(M)$ is the space of forms of type (p, q).

For each smooth function $f : M \to \mathbb{C}$, we have $df = \partial f + \overline{\partial} f$, where $\partial f = \frac{\partial}{\partial z^k} f dz^k$ is a section of $T^{1,0}M$, and $\overline{\partial} f = \frac{\partial}{\partial \overline{z}^k} f d\overline{z}^k$ is a section of $T^{0,1}M$. These differential operators extend to the algebra $\Omega_{\mathbb{C}}(M)$ and give the differential operators $\partial : \Omega^{p,q}(M) \to \Omega^{p+1,q}(M)$ and $\overline{\partial} : \Omega^{p,q}(M) \to \Omega^{p,q+1}(M)$. It is clear that $d = \partial + \overline{\partial}$, and

$$\partial^2 = 0, \qquad \overline{\partial}^2 = 0, \qquad \partial\overline{\partial} = \overline{\partial}\partial.$$

The complex $(\Omega^{p,*}(M), \overline{\partial})$ is called the *Dolbeaux complex* of the complex manifold (M, J), and its cohomology is denoted by $H^{p,q}(M)$.

Lemma 3 (Poincare $\overline{\partial}$-Lemma [36]) *For a polydisk $\Delta \subset \mathbb{C}^n$, we have $H^{p,q}(\Delta) = 0$ for all p and $q \geq 1$.*

From this lemma and the definition of $\overline{\partial}$ it follows that *the sheaf sequence*

$$0 \to \Omega^p_h \xrightarrow{i} \Omega^{p,0} \xrightarrow{\overline{\partial}} \Omega^{p,1} \ldots \xrightarrow{\overline{\partial}} \Omega^{p,q} \xrightarrow{\overline{\partial}} \Omega^{p,q+1} \xrightarrow{\overline{\partial}} \ldots$$

is a fine resolution for the sheaf of holomorphic p-forms Ω^p_h, hence $H^q(M; \Omega^p) \cong H^{p,q}(M)$.

The differential algebra $\Omega_{\mathbb{C}}(M)$ can be endowed by the structure of double complex $(\Omega^{p,q}(M), \partial, \overline{\partial})$, the corresponding spectral sequences are called the *Froelicher spectral sequences*. These spectral sequences relate the Dolbeaux cohomology of (M, J) and the de Rham cohomology of M.

5.1.1 Harmonic theory on compact complex manifolds

Let (M, J, h) be a compact Hermitian complex manifold, and $\dim_{\mathbb{C}} M = n$. The Hermitian metric h determines the Riemannian metric g, which, in turn, determines the Hodge operator $* : \Omega(M) \to \Omega(M)$ (see 2.3). Then, for any $k = \overline{0, 2n}$, on the space $\Omega^k_{\mathbb{C}}(M)$ we have the scalar product

$$(\varphi, \psi) = \int_M \varphi \wedge \bar{\psi} \tag{13}$$

This scalar product determines a positively definite Hermitian form on the complex vector space $\Omega_{\mathbb{C}}(M) = \oplus_{k=0}^{2n}\Omega^k(M)$ and the decomposition $\Omega^k_{\mathbb{C}}(M) = \bigoplus_{i+j=k} \Omega^{i,j}(M)$ is orthogonal with respect to (13). The operator $\overline{\partial} : \Omega^{p,q}(M) \to \Omega^{p,q+1}(M)$ has adjoint with respect to the scalar product (13) operator

$$\overline{\partial}^* = -\overline{*}\,\overline{\partial}\,\overline{*}$$

The $\overline{\partial}$-*Laplacian* is $\Delta_{\overline{\partial}} = \overline{\partial}\,\overline{\partial}^* + \overline{\partial}^*\,\overline{\partial}$ (sometimes this operator is denoted by $\overline{\Box}$). A form $\varphi \in \Omega^{p,q}(M)$ such that $\Delta_{\overline{\partial}}\varphi = 0$ is called a *harmonic form of type* (p, q). Denote the space of harmonic forms of type (p, q) by $\mathcal{H}^{p,q}(M)$.

Theorem 13 (Hodge decomposition theorem, see, e. g., [36]) *a)* $\dim \mathcal{H}^{p,q} < \infty$.

b) There is defined the orthogonal projection $\mathcal{H} : \Omega^{p,q}(M) \to \mathcal{H}^{p,q}(M)$ and there exists a unique operator, called the Green operator, $G : \Omega^{p,q}(M) \to \Omega^{p,q}(M)$, such *that* $G(\mathcal{H}^{p,q}(M)) = 0$, $\overline{\partial}G = G\overline{\partial}$, $\overline{\partial}^* G = G\overline{\partial}^*$, *and, for each* $\psi \in \Omega^{p,q}(M)$, $\psi = \mathcal{H}\psi + \overline{\partial}(\overline{\partial}^* G\psi) + \overline{\partial}^*(\overline{\partial}G\psi)$. *From this follows the Hodge decomposition*

$$\Omega^{p,q}(M) = \mathcal{H}^{p,q}(M) \oplus \overline{\partial}\Omega^{p,q-1}(M) \oplus \overline{\partial}^* \Omega^{p,q+1}(M).$$

Theorem 14 (Kodaira-Serre duality, see, e. g. [36]) *Let* M *be a compact complex manifold of complex dimension* n. *Then*

1) $H^m(M;\Omega^n) \cong \mathbb{C}$;

2) The paring

$$H^q(M;\Omega^p) \otimes H^{n-q}(M;\Omega^{n-p}) \to H^n(M;\Omega^n)$$

is nondegenerate, where Ω^k *is the sheaf of holomorphic forms.*

For a complex compact manifold the spaces of de Rham cohomology $H^r(M,\mathbb{C})$ and of Dolbeaux cohomology $H^{p,q}(M)$ are finite dimensional (see 2.3 and Theorem 13). In general, if a form φ is $\overline{\partial}$-closed, then φ need not to be d-closed, and if a k-form ψ is d-closed, and $\psi = \psi^{k,0} + \psi^{k-1,1} + \dots + \psi_{0,k}$, $\psi^{p,k-p} \in \Omega^{p,k-p}(M)$, then $\psi^{p,k-p}$ need not to be $\overline{\partial}$-closed.

Now let a complex compact manifold (M,J) admit a Kähler metric h (see, e. g. [36]). Then $\mathcal{H}^k(M) = \oplus_{p+q=k}\mathcal{H}^{p,q}(M)$, where $\mathcal{H}^k(M)$ is the space of Δ_d-harmonic forms, and $\mathcal{H}^{p,q}(M)$ is the space of $\Delta_{\overline{\partial}}$-harmonic forms of type (p,q). From this follows the *decomposition Hodge theorem for compact Kähler manifolds*:

Theorem 15 (see, e.g., [36], [94]) *Let* M *be a compact complex manifold admitting Kähler metric. Then*

$$H^k(M;\mathbb{C}) = \oplus_{p+q=k} H^{p,q}(M) \text{ and } \overline{H}^{p,q}(M) = H^{q,p}(M),$$

where $H^k(M;\mathbb{C})$ *is the de Rham cohomology of* M *with complex coefficients, and* $H^{p,q}(M)$ *is the Dolbeaux cohomology of* (M,J).

5.2 Vaisman cohomology

Let $\mathcal{F}$ be a foliation of codimension m on an $(m+r)$-dimensional manifold, and $T\mathcal{F}$ be the subbundle of TM tangent to $\mathcal{F}$. With respect to the coordinate system (x^a, x^α), $a = \overline{1,m}$, $\alpha = \overline{m+1, m+r}$, adapted to the foliation, the leaves are given by the equations $x^a = const$, and the integrable distribution $T\mathcal{F}$ is locally determined by $dx^a = 0$ (for the foliation theory we refer the reader to [42], [68], [87]).

A k-form $\theta \in \Omega^k(M)$ is said to be *basic* if with respect to the adapted coordinates

$$\theta = \theta_{a_1 \dots a_k}(x^a) dx^{a_1} \wedge \dots dx^{a_k}.$$

Evidently, the set of basic forms is a subcomplex (Ω_b, d_b) of the de Rham complex of M and its cohomology $H_b(M,\mathcal{F})$ is called the *basic cohomology of the foliated manifold* $(M,\mathcal{F})$. In general, $H_b(M,\mathcal{F})$ is infinite-dimensional even if M is compact. If the foliation $\mathcal{F}$ is simple (this means that the leaves of $\mathcal{F}$ are fibers of a submersion $\pi : M \to B$), then the basic cohomology is isomorphic to the de Rham cohomology of B.

Let Q be a distribution complementary to $T\mathcal{F}$: $TM = T\mathcal{F} \oplus Q$. Let $\eta^\alpha = 0$ locally determine Q, then (dx^a, η^α) is a local coframe field on M. Let us denote by $\Omega^{p,q}(M,\mathcal{F})$ the space of $(p+q)$-forms which are locally written as follows:

$$\theta = \theta_{a_1\ldots a_p\alpha_1\ldots\alpha_q} dx^{a_1} \wedge \ldots dx^{a_p} \wedge \eta^{\alpha_1} \wedge \ldots \eta^{\alpha_q}.$$

Then, we have natural decomposition $d = d_{1,0} + d_{0,1}$, where $d_{1,0} : \Omega^{p,q}(M,\mathcal{F}) \to \Omega^{p+1,q}(M,\mathcal{F})$, $d_{0,1} : \Omega^{p,q}(M,\mathcal{F}) \to \Omega^{p,q+1}(M,\mathcal{F})$, and, from $d^2 = 0$ it follows that

$$d_{1,0}^2 = 0, \quad d_{0,1}^2 = 0, \quad d_{1,0}d_{0,1} = d_{0,1}d_{1,0},$$

so, for each $p \geq 0$, we get the differential complex $(\Omega^{p,*}(M,\mathcal{F}), \partial)$, where $\partial = d_{0,1}$. The cohomology $H^{p,q}(M,\mathcal{F}) = H^q(\Omega^{p,*}(M,\mathcal{F}), \partial)$ of this complex is called *foliated cohomology* (or, sometimes, *leafwise cohomology*) of type (p,q). Foliated cohomology has many properties similar to the de Rham cohomology, for example, one can write the Mayer-Vietoris sequence for them [41].

For each $p \geq 0$, we have the exact sequence of sheaves over M:

$$0 \to \Omega_b^p \xrightarrow{i} \Omega^{p,0} \xrightarrow{\partial} \Omega^{p,1} \xrightarrow{\partial} \ldots \xrightarrow{\partial} \Omega^{p,k} \xrightarrow{\partial} \Omega^{p,k+1} \xrightarrow{\partial} \ldots$$

where i is the inclusion, which gives a fine resolution for the sheaf Ω_b^p of basic p-forms on $(M,\mathcal{F})$. Therefore, $H^q(M;\Omega_b^p) \cong H^{p,q}(M,\mathcal{F})$. The differential ring $(\Omega(M), d)$ admits the filtration by the differential ideals $F^k(M) = \oplus_{i\geq k}\Omega^{i,*}(M,\mathcal{F})$, and $(\Omega^{*,*}(M,\mathcal{F}), d_{1,0}, d_{0,1})$ is the corresponding double complex. Therefore, for the foliated manifold we get the spectral sequence with the first term $E_1^{p,q} = H^{p,q}(M,\mathcal{F})$ which converges to the de Rham cohomology of M.

The details of the above constructions the reader can find in I.Vaisman's book [87].

Let $(M,\mathcal{F})$ be a foliated manifold. A metric g on M is said to be *bundle-like* if, with respect to the adapted coordinates (x^a, x^α), $g = g_{ab}(x^c)dx^a \otimes dx^b + g_{\alpha\beta}(x^c, x^\gamma)\theta^\alpha \otimes \theta^\beta$, where $\theta^\alpha = dx^\alpha + t_a^\alpha dx^a$ is a coframe of $T\mathcal{F}^\perp$. If $\mathcal{F}$ admits a bundle-like metric, then $\mathcal{F}$ is called a *Riemannian foliation*. Let ∇ be the Levi-Civita connection of g, and for each $X \in T\mathcal{F}^\perp$, $W(X) : T\mathcal{F} \to T\mathcal{F}$ be the Weingarten map of the leaves. The *mean curvature* κ is a 1-form on M defined by $\kappa(X) = \mathrm{tr}(W(X))$ for $X \in T\mathcal{F}^\perp$, and $\kappa(X) = 0$ for $X \in T\mathcal{F}$. Now the usual global scalar product on forms restricts on $\Omega_b(M,\mathcal{F})$ to a scalar product. The adjoint δ_b of the operator d_b and the Laplacian $\Delta_b = \delta_b d_b + d_b\delta_b$ are therefore defined. In [43] a version of the Hodge theory is constructed for Riemannian foliations. In particular, the following result is proved:

Theorem 16 ([43]) *Let $\mathcal{F}$ be a transversally oriented Riemannian foliation on a compact oriented manifold M. Assume that there exists a bundle-like metric g whose mean curvature is a basic 1-form. Then*

$$\Omega_b = \mathrm{im}\, d_b \oplus \mathrm{im}\, \delta_b \oplus H_b,$$

where H_b is the kernel of Δ_b on Ω_b. This is a decomposition into mutually orthogonal subspaces with finite-dimensional H_b.

The foliated cohomology is used in construction of various versions of characteristic classes of foliations and obstructions to existence of basic geometric objects (see, e. g., [42],[68]).

Here we should also mention recent results by Crainic, Marius; Moerdijk, Ieke [19] who present a new Čech-De Rham model for the cohomology of the classifying space of a foliated manifold. In part, the Čech-De Rham model can be used to prove a version of Poincare duality for foliations.

5.3 Cohomology of manifolds over algebras

Let $\mathbb{A}$ be a finite-dimensional commutative associative algebra with unit over $\mathbb{R}$. Then $\mathbb{A}^n \cong \mathbb{R}^{nm}$ (as vector spaces over $\mathbb{R}$), where $m = \dim_{\mathbb{R}} \mathbb{A}$, this isomorphism defines a topology on $\mathbb{A}^n$. Let $U \subset \mathbb{A}^n$, $V \subset \mathbb{A}^k$ be open subsets. A map $F : U \to V$ is said to be $\mathbb{A}$-differentiable if, at each point $a \in U$, the differential $dF_a : T_a\mathbb{A}^n \cong \mathbb{A}^n \to T_{F(a)}\mathbb{A}^k \cong \mathbb{A}^k$ is $\mathbb{A}$-linear. Let Γ be the pseudogroup of all local $\mathbb{A}$-diffeomorphisms of $\mathbb{A}^n$. A maximal Γ-atlas on a topological space M is called a *structure of manifold over* $\mathbb{A}$ *on* M, and M is called *a manifold over* $\mathbb{A}$.

The algebra $\mathbb{A}$ is said to be *local* if $\mathbb{A} \cong \mathbb{R} \oplus \overset{\circ}{\mathbb{A}}$, where $\overset{\circ}{\mathbb{A}}$ is the radical (the set of nilpotent elements) of $\mathbb{A}$. The simplest example of such algebra is the algebra of dual numbers $\mathbb{R}(\varepsilon) = \{a + b\varepsilon \mid \varepsilon^2 = 0\}$. The examples of manifolds over algebras are provided by total spaces of jet bundles, in part, the total space of a tangent bundle is a manifold over $\mathbb{R}(\varepsilon)$. Any ideal $\mathbb{I} \subset \mathbb{A}$ determines a foliation on M whose tangent bundle is $\mathbb{I} \cdot TM$, in particular, the foliation determined by $\overset{\circ}{\mathbb{A}}$ is called the *canonical foliation* on M. For the general theory of manifolds over algebras we refer the reader to [76], [77] (see also references given there).

In [78] V.V.Shurygin constructed de Rham cohomology for manifolds over algebras. Let M be a manifold over $\mathbb{A}$, $\dim_{\mathbb{A}} M = n$, then each tangent space T_pM has natural structure of free module of rank n over $\mathbb{A}$. One can consider the complex of $\mathbb{A}$-valued differential form $(\Omega_{\mathbb{A}}(M) = \Omega(M) \otimes \mathbb{A}, d)$, where d is the exterior differential extended to $\Omega_{\mathbb{A}}(M)$ by linearity. A k-form $\varphi \in \Omega^k_{\mathbb{A}}(M)$ is said to be $\mathbb{A}$-differentiable if φ and $d\varphi$ are $\mathbb{A}$-linear with respect to the $\mathbb{A}$-module structure on the tangent spaces. In this case, if $Z^i : U \to \mathbb{A}$ are $\mathbb{A}$-valued coordinates on an open $U \subset M$, we have

$$\varphi|_U = \varphi_{i_1 \dots i_k}(Z^1, \dots, Z^k) dZ^{i_1} \wedge \cdots \wedge dZ^{i_k},$$

where $\varphi_{i_1 \dots i_k}(Z^1, \dots, Z^k)$ are $\mathbb{A}$-differentiable functions. Denote by $\Omega^k_{\mathbb{A}-diff}(M)$ the space of $\mathbb{A}$-differentiable forms, then we obtain the subcomplex $(\Omega_{\mathbb{A}-diff}(M), d)$ of the complex $(\Omega_{\mathbb{A}}(M), d)$ called the *de Rham complex of* $\mathbb{A}$*-differentiable forms on* M.

Theorem 17 (Poincare Lemma for $\mathbb{A}$-differentiable forms, [78]) *Let U be an open coordinate parallelepiped in $\mathbb{A}^n$, and let φ be an $\mathbb{A}$-differentiable k-form on U such that $d\varphi = 0$; then there exists an $\mathbb{A}$-smooth $(k-1)$-form ψ on U such that $\varphi = d\psi$.*

In [78] (see also [77]) V.V.Shurygin constructs complexes of differential forms determined by ideals of $\mathbb{A}$ and finds properties of their cohomology. Also, for manifolds over algebras, he constructs a bicomplex of differential forms, which, on one hand generalizes the Dolbeaux bicomplex of complex manifold (see 5.1), and on the other hand, the Vaisman bicomplex of foliated manifold (see 5.2).

For the properties of the de Rham complex of manifolds over algebras see also [31], [63], [67].

5.4 Poisson cohomology

Let M be a smooth manifold. Any $\Psi \in \Omega_2(M)$ determines a skewsymmetric bracket $\{\ ,\ \} : F(M) \times F(M) \to F(M)$ by $\{f, g\} = \Psi(df, dg)$. This bracket is said to be *Poisson* if it satisfies the Jacobi identity $\{\{f, g\}, h\} + \{\{h, f\}, g\} + \{\{g, h\}, f\} = 0$, and (M, Ψ) is called a *Poisson manifold*. Note that the Jacobi identity for $\{,\}$ is equivalent to the local equality $\Psi^{[is}\partial_s \Psi^{jk]} = 0$.

On a smooth manifold M we have the Schouten-Nijenhuis bracket $\Omega_k(M)\times\Omega_l(M) \to \Omega_{k+l-1}(M)$ uniquely defined by

$$[u_1\wedge u_2\wedge\cdots\wedge u_k, v_1\wedge v_2\wedge\cdots\wedge v_l] = \sum_{i,j}(-1)^{i+j}[u_i, v_j]\wedge u_1\wedge\ldots\widehat{u_i}\ldots u_k\wedge v_1\wedge\ldots\widehat{v_j}\ldots v_l. \tag{14}$$

For $A \in \Omega_k(M)$ and $B \in \Omega_l(M)$, we have $[A, B] \in \Omega_{k+l-1}(M)$, and $(\Omega^k(M), [\ ,\ ])$ is a Lie superalgebra:

$$[A, B] = (-1)^{\deg A \deg B}[B, A] \tag{15}$$

$$(-1)^{\deg A \deg C}[[A, B], C] + (-1)^{\deg C \deg B}[[C, A], B] + (-1)^{\deg B \deg A}[[B, C], A] = 0 \tag{16}$$

If $\Psi \in \Omega^2(M)$ is a Poisson structure on M, then

$$[\Psi, \Psi] = 0. \tag{17}$$

Now we have $D : \Omega_k(M) \to \Omega_{k+1}(M)$, $DA = [\psi, A]$, and from (15),(16), and (17) it follows that $D^2 = 0$. The cohomology $H_P(M, \Psi)$ of the complex (Ω_*, D) is called the *Lichnerowicz-Poisson cohomology* of (M, Ψ) [85].

Let $\widetilde{\Psi} : T^*M \to TM$, $\alpha_i \to v^i = \Psi^{ij}\alpha_j$, be the vector bundle morphism determined by a Poisson structure Ψ. Then, the operator $D : \Omega_k(M) \to \Omega_{k+1}(M)$ is given as follows:

$$Dw(\alpha^0, \ldots, \alpha^k) = \sum_{i=0}^{k} \widetilde{\Psi}\alpha_i(w(\alpha_0, \ldots, \widehat{\alpha_i}, \ldots, \alpha_k)) + \sum_{0=i<j}^{k} (-1)^{i+j} w(\{\alpha_i, \alpha_j\}, \alpha_0, \ldots, \widehat{\alpha_i}, \ldots, \widehat{\alpha_j}, \ldots, \alpha_k). \tag{18}$$

Recall that a vector field $V \in \mathfrak{X}(M)$ is called a *Poisson vector field* if $\mathcal{L}_V\Psi = 0$, and a *Hamiltonian vector field* if $V = V_f = \Psi^{ij}\partial_i f \partial_j$. Any Hamiltonian vector field is a Poisson vector field. From (18) it follows that, for a function $f \in \Omega_0(M)$, $Df = V_f$, and, for a vector field $W \in \Omega_1(M)$, $DW = -\mathcal{L}_W\Psi$. In these terms we can express Poisson cohomology in small dimensions. *$H^0_P(M, \Psi)$ is the set of Casimir functions of Ψ, that is of functions f such that $\Psi^{ij}\partial_j f = 0$. The cohomology group $H^1_P(M, \Psi)$ is the quotient space of the space of Poisson vector fields by the space of Hamiltonian vector fields. The space $H^2_P(M, \Psi)$ is the space of essential infinitesimal deformations of Ψ.*

In [70] N. Nakanishi calculated Poisson cohomology of quadratic Poisson structures on a plane. I.Vaisman constructed a double complex associated with Poisson structure (see [88], [86]). For various results on Poisson cohomology calculation, see, e. g. [32], [69], [71], [74], [95].

5.5 Koszul complex

For a Poisson manifold (M, ψ), J.L.Koszul defined the differential $\delta : \Omega^k(M) \to \Omega^{k+1}(M)$, $\delta = i(\Psi)d - di(\Psi)$, where $i(\Psi)$ is the contraction with the Poisson tensor Ψ, and proved that $\delta^2 = 0$ [47]. The corresponding complex

$$\dots \xrightarrow{\delta} \Omega^{k+1}(M) \xrightarrow{\delta} \Omega^k(M) \xrightarrow{\delta} \Omega^{k-1}(M) \xrightarrow{\delta} \dots$$

was called by J.-L. Brylinski [10] the *canonical complex*, and $H(\Omega^*(M), \delta) = H^*_{\text{can}}(M, \Psi)$ the *canonical homology* of Poisson manifold (M, Ψ).

The differential δ is given as follows [10]

$$\delta(f_0\, df_1 \wedge \dots \wedge df_k) = \sum_{1 \le i \le k} (-1)^{i+1} \{f_0, f_i\} df_1 \wedge \dots \wedge \widehat{df_i} \wedge \dots \wedge df_k$$
$$+ \sum_{1 \le i < j \le k} (-1)^{i+j} f_0 d\{f_i, f_j\} \wedge df_1 \wedge \dots \wedge \widehat{df_i} \wedge \dots \wedge \widehat{df_j} \wedge \dots \wedge df_k, \quad (19)$$

and $d \circ \delta + \delta \circ d = 0$.

Let (M, ω), $\dim M = 2m$, be a symplectic manifold. For $k \ge 0$, let $\wedge^k w : \wedge^k(T^*M) \times \wedge^k(T^*M) \to C^\infty(M)$ be the pairing defined by $\wedge^k w(\alpha, \beta) = w^{i_1 j_1} \dots w^{i_k j_k} \alpha_{i_1 \dots i_k} \beta_{j_1 \dots j_k}$. Also, on M we take the volume form $v_M = \omega^m / m!$.

J.-L.Brylinski introduced an analog of the Hodge operator, the operator $* : \Omega^k(M) \to \Omega^{2n-k}(M)$, $\beta \wedge (*\alpha) = \wedge^k w(\beta, \alpha) \cdot v_M$ for all $\alpha, \beta \in \Omega^k(M)$. This operator has the property $*(*\alpha) = \alpha$ and hence is an isomorphism. Also, for all $\alpha \in \Omega^k(M)$, $\delta\alpha = (-1)^{k+1} * d * \alpha$. Hence follows that *for a symplectic manifold* (M, ω), $\dim M = 2m$, *the operator* $*$ *gives an isomorphism* $H^k_{\text{can}}(M, w) \cong H^{2m-k}_{dR}(M)$.

Using the operator δ, one can construct another complex associated to a Poisson structure Ψ on a manifold M. Set $\Omega_0(M) = \oplus_{k \ge 0} \Omega^{2k}(M)$, $\Omega_1(M) = \oplus_{k \ge 0} \Omega^{2k+1}(M)$, and $D = d + \delta$. A form $\varphi \in \Omega_0(M)$ is said to be *even*, and $\varphi \in \Omega_1(M)$ is said to be *odd*. Then, the differentials d and D are odd. In [79] V.V.Shurygin, jr constructed a chain $\mathbb{Z}_2$-graded homomorphism $\varphi : (\Omega(M), d) \to (\Omega(M), D)$ which gives isomorphism in cohomology: $H_0(\Omega(M) \cong \oplus_{k \ge 0} \Omega^{2k}(M)$ and $H_1(\Omega(M) \cong \oplus_{k \ge 0} \Omega^{2k+1}(M)$.

This result can be applied to the *periodic double complex* $\mathcal{E}^{\text{per}}_{*,*}(M)$ defined by $\mathcal{E}^{\text{per}}_{p.q}(M) = \Lambda^{q-p}(M)$, for all $p, q \in \mathbb{Z}$, which has d for the horizontal differential and δ for the vertical differential, both of degree -1. From above it follows that its total cohomology $H^D_*(M)$ is isomorphic to the de Rham cohomology of M. For $\mathcal{E}^{\text{per}}_{*,*}(M)$, one has two spectral sequences $\{E^r(M)\}$ and $\{'E^r(M)\}$, both converging to the total homology. J.-L. Brylinski proved that, if M is a compact $2n$-dimensional symplectic manifold, the first spectral sequence $\{E^r(M)\}$ degenerates at $E^1(M)$ (i.e., $E^1(M) \cong E^\infty(M)$). In [28] M. Fernández, R. Ibáñez, and M. de Leon give an example of a 5-dimensional compact Poisson manifold M^5 for which $H^{\text{can}}_1(M^5) \not\cong H^4(M^5)$, and $E^1(M^5) \not\cong E^2(M^5)$.

Note also that in [10] J.L.Brylinski proved the following analog of Hodge's theorem (see 2.3):

Theorem 18 *For a compact manifold M with positive definite Kähler metric, any de Rham cohomology class $\xi \in H(M)$ has a symplectically harmonic representative: $\xi = [\alpha]$ and $d\alpha = 0$, $\delta\alpha = 0$.*

However, in general, the analog of Hodge's theorem fails, a counterexample is given in [27].

J.-L.Brylinski's complexes appear, for example, in H.-D. Cao-J. Zhou's deformation quantization of de Rham complex [11]. Using the above mentioned results, V.V.Shurygin, jr demonstrated that the quantum cohomology of Poisson manifold are obtained by deformation quantization of de Rham cohomology [79].

Properties of Poisson cohomology of the lift of a Poisson structure on a smooth manifold M to $T^A(M)$, where A is a Weil algebra, and T^A is the Weil functor, are studied in [80].

5.6 Coeffective cohomology

Another interesting complex of differential forms associated to a symplectic structure was introduced by T. Bouche in [7]. The *coeffective cohomology* of a symplectic manifold (M, ω) is the cohomology of the differential subcomplex $(A^*(M), d)$ of the de Rham complex consisting of the coeffective forms, i.e. those forms α such that $\alpha \wedge \omega = 0$. Bouche proved that, *if M is compact, the coeffective complex is elliptic for $p \geq n+1$, and hence its cohomology groups have finite dimension.* Moreover, Bouche proved that *for a compact Kähler $2n$-dimensional manifold the coeffective cohomology groups and the truncated de Rham cohomology groups by the de Rham class of the Kähler form are isomorphic for degree $p \geq n+1$.* This result does not hold for an arbitrary symplectic manifold, the counterexamples are constructed in [26]: the authors take the 6-dimensional nilmanifold $\mathbb{R}^6 = \Gamma \backslash G$, where G is the simply connected nilpotent Lie group of dimension 6, defined by left-invariant 1-forms $\{\alpha_i, 1 \leq i \leq 6\}$ such that

$$d\alpha_1 = d\alpha_2 = d\alpha_3 = 0,\ d\alpha_4 = \alpha_2 \wedge \alpha_1,\ d\alpha_5 = \alpha_3 \wedge \alpha_1,\ d\alpha_6 = \alpha_4 \wedge \alpha_1,$$

and Γ is a uniform subgroup of G, and the symplectic form is given by

$$\omega = \alpha_1 \wedge \alpha_5 + \alpha_1 \wedge \alpha_6 + \alpha_2 \wedge \alpha_5 + \alpha_3 \wedge \alpha_4 + \alpha_1 \wedge \alpha_3.$$

They calculate the coeffective cohomology of (M, ω) using direct calculation and the following Nomizu's type theorem

Theorem 19 ([26]) *Let G be a connected nilpotent Lie group endowed with an invariant symplectic form $\overline{\omega}$ and with a discrete subgroup Γ such that the space of right cosets $M = \Gamma \setminus G$ is compact. Then there is an isomorphism of cohomology groups $H_p(A(g)) = H_p(A(M))$ for all $p \geq n+1$, $\dim G = 2n$, where $H_p(A(g))$ is the coeffective cohomology with respect to $\overline{\omega}$ and $H_p(A(M))$ is the coeffective cohomology defined by the projected symplectic form ω on M.*

The relationship between the coeffective cohomology and the de Rham cohomology was studied in [29], where M. Fernández, R. Ibáñez, and M. de León obtained a bound for the coeffective numbers and proved that the lower bound is got for compact Kähler manifolds, and the upper one for non-compact exact symplectic manifolds. Also they studied the behavior of the coeffective cohomology under deformations.

Note that similar complexes can be constructed for other geometric structures which can be defined in terms of a closed differential form. For example, in [84] L. Ugarte studied the properties of the coeffective complex for $\mathrm{Spin}(7)$-manifolds.

5.7 General remarks on complexes of differential forms associated to differential geometric structures

Remark 5.1 The results on the canonical cohomology of Poisson manifold and the periodic double complex are generalized to the *Jacobi manifolds* in [17] and [58] (see also references there).

Remark 5.2 *Generalized complex structures* introduced by M. Hitchin [40] are now extensively studied. Generalized complex geometry is a new kind of geometrical structure which contains complex and symplectic geometry as its extremal special cases. In [16] the operators ∂ and $\overline{\partial}$ (see 5.1) and the corresponding cohomology were defined for the generalized complex structure. For this structure, there were constructed spectral sequences similar to the Frölicher spectral sequences and an analog of Serre duality theorem (Theorem 14) was proved (see [16], [37], and references there).

Remark 5.3 The construction of differential complexes for complex structure and its generalizations (like the manifolds over algebras) uses a decomposition of the tangent bundle determined by the corresponding structure. In [51] A.Kushner studies decomposition of the exterior differential on a manifold M endowed with almost product structure (TM is splitted in a direct sum of distributions), and applies the obtained tensor invariants to solving the problem of contact equivalence and the problem of contact linearization for Monge-Ampère equations.

Remark 5.4 One of the general ways to obtain a complex of differential forms associated to a differential geometric structure is to consider the Spencer complex for the Lie derivative [72]. A *differential geometric structure* on a smooth manifold M is a section s of a natural bundle $\xi : E_M \to M$. Then we can define the Lie derivative of s with respect to a vector field X on M, which determines a first order differential operator $\mathcal{L}s : TM \to s^*(VE)$, $X \to \mathcal{L}_X s$. For $\mathcal{L}s$ we have the Spencer P-complex [72], which, in many cases, gives the fine resolution of the sheaf of infinitesimal automorphisms of s. For example, in this way for the complex structure we obtain the Dolbeaux complex (see 5.1), for the foliation structure the Vaisman cohomology (see 5.2), etc. (see [64], [65], [72]).

Remark 5.5 The theory of partial differential equations and the corresponding geometry of jet bundles give rise to various complexes of differential forms, in particular, to so-called variational complex of a locally trivial bundle. The numerous interesting results in this field go far beyond the scope this paper, so we refer the reader to the book by I.S. Krasil'shchik, V.V. Lychagin, and A.M. Vinogradov [48], and to [49]. Also, in this connection, see the recent paper by D.Krupka [50] and references there.

Bibliography

[1] S. M. A. Balea: Cohomologies on the foliated Riemannian manifolds *Bull. Math. Soc. Sci. Math. Roumanie (N.S.)* **46(94)** ((2003), no. 1-2), (2004) 3–10

[2] R. Benedetti, A. Tognoli: Teoremi di approssimazione in topologia differenziale *I. Boll. Un. Mat. Ital. B. (5)* **14(3)** (1977) 866–887

[3] L. Beretta: An analytic version of the de Rham theorem *Ann. Univ. Ferrara Sez. VII (N.S.)* **40** (1994), (1996) 23–30

[4] L. Bérard Bergery (ed.), M. Berger (ed.), Ch. Houzel (ed.) (A. Besse): *Géométrie riemannienne en dimension 4* (Seminaire Arthur Besse 1978/79. (French) Textes Mathematiques, 3. Paris: CEDIC/Fernand Nathan. XVIII, 383 p., 1981)

[5] R. Bott: On the characteristic classes of groups of diffeomorphisms *Enseign. Math., II. Sér.* **23** (1977) 209–220

[6] R. Bott, L. W. Tu: *Differential forms in algebraic topology* (Graduate Texts in Mathematics, 82. New York - Heidelberg - Berlin: Springer-Verlag. XIV, xiv+331 p., 1982)

[7] T. Bouche: La cohomologie coeffective d'une variété symplectique *Bull. Sci. Math., II. Sér.* **114(2)** (1990) 115–122

[8] J. P. Brasselet, A. Legrand: Differential forms on singular varieties and cyclic homology *In: Singularity theory* (Bruce, Bill (ed.) et al., Proceedings of the European singularities conference, Liverpool, UK, August 18-24, 1996. Dedicated to C.T.C. Wall on the occasion of his 60th birthday. Cambridge: Cambridge University Press. Lond. Math. Soc. Lect. Note Ser. 263, 1999) 175–187

[9] G. E. Bredon: *Sheaf theory* (New York-St. Louis-San Francisco-Toronto-London-Sydney: McGraw-Hill Book Company. XI, 272 p., 1967)

[10] J.-L. Brylinski: A differential complex for Poisson manifolds *J. Differential Geom.* **28(1)** (1988) 93–114

[11] H.-D. Cao, J. Zhou: On Quantum de Rham Cohomology *Preprint* `arXiv:math.DG/9806157` (1998)

[12] G. Carron, T. Coulhon A. Hassell: Riesz transform and Lp-cohomology for manifolds with Euclidean ends *Duke Math. J.* **133(1)** (2006) 59–93

[13] H. Cartan: *Formes différentielles* (Hermann, Paris, 1967)

[14] H. Cartan: *Calcul différentiel* (Hermann, Paris, 1967)

[15] G. R. Cavalcanti, M. Gualtieri: Generalized complex structures on nilmanifolds *J. Symplectic Geom.* **2(3)** (2004) 393–410

[16] G. R. Cavalcanti: The decomposition of forms and cohomology of generalized complex manifolds *Geom. Phys.* **57** (to appear), (2007) 121–132

[17] D. Chinea, J. C. Marrero, M. de León: A canonical differential complex for Jacobi manifolds *Mich. Math. J.* **45(3)** (1998) 547–579

[18] A. Connes: Non-Commutative differential geometry *Publ. Math., Inst. Hautes Etud. Sci.* **62** (1985) 257–360

[19] M. Crainic, I. Moerdijk: Čech-De Rham theory for leaf spaces of foliations *Math. Ann.* **328(1-2)** (2004) 59–85

[20] G. de Rham: *Variétés différentiables. Formes, courants, formes harmoniques* (Actualités scientifiques et industrielles. 1222. Publications de l'institut mathématique de l'université de Nancago. III. Paris: Hermann & Cie. VII, 196 p., 1955)

[21] V. A. Dolgushev: Hochschild cohomology versus de Rham cohomology without formality theorems *Int. Math. Res. Not. 2005* **21** (English), (2005) 1277–1305

[22] B. A. Dubrovin, A. T. Fomenko, S. P. Novikov: *Modern geometry - methods and applications. Part III. Introduction to homology theory* (Graduate Texts in Mathematics, 124. New York etc.: Springer-Verlag. ix, 416 p., 1990)

[23] Ch.-O. Ewald: A de Rham isomorphism in singular cohomology and Stokes theorem for stratifolds *Int. J. Geom. Methods Mod. Phys.* **2(1)** (2005) 63–81

[24] Ch.-O. Ewald: *Hochschild homology and de Rham cohomology of stratifolds* (Heidelberg: Univ. Heidelberg, Naturwissenschaftlich-Mathematische Gesamtfakult, iii, 101 p., 2002)

[25] L. D. Faddeev, A. G. Reyman, M. A. Semenov-Tian-Shansky: Quantum anomalies and cocycles on gauge groups *Funct. Anal. Appl.* **18** (1984) 319–326; *translation from Funkts. Anal. Prilozh.* **18(4)** (1984) 64–72

[26] M. Fernández, R. Ibáñez, M. de León: A Nomizu's theorem for the coeffective cohomology *Math. Z.* **226(1)** (1997) 11–23

[27] M. Fernández, R. Ibáñez, M. de León: On a Brylinski conjecture for compact symplectic manifolds *In: Quaternionic structures in mathematics and physics* (Trieste, 1994, (electronic), Int. Sch. Adv. Stud. (SISSA), Trieste, 1998) 119–126

[28] M. Fernández, R. Ibáñez, M. de Leon: The canonical spectral sequences for Poisson manifolds *Isr. J. Math.* **106** (1998) 133–155

[29] M. Fernández, R. Ibáñez, M. de Leon: Coeffective and de Rham cohomologies of symplectic manifolds *J. Geom. Phys.* **27(3-4)** (1998) 281–296

[30] M. Ferrarotti: A complex of stratified forms satisfying de Rham's theorem *In: Stratifications, singularities and differential equations, II* (Marseille, 1990; Honolulu, HI, 1990, Travaux en Cours, 55, Hermann, Paris, 1997) 25–38

[31] T. I. Gaisin: To the question about the maximum principle for manifolds over local algebras *Sib. Mat. Zh.* **46** (2005) 79–89

[32] A. Gammella: An approach to the tangential Poisson cohomology based on examples in duals of Lie algebras *Pacific J. Math.* **203(2)** (2002) 283–319

[33] V. M. Gol'dshtejn, V. I. Kuz'minov, I. A. Shvedov: De Rham isomorphism of the Lp-cohomology of noncompact Riemannian manifolds *Sib. Math. J.* **29(2)** (1988) 190–197; *translation from Sib. Mat. Zh.* **29(2(168))** (1988) 34–44

[34] V. M. Gol'dshtein, V. I. Kuz'minov, I. A. Shvedov: On Cheeger's theorem: Extensions to Lp-cohomology of warped cylinders *Sib. Adv. Math.* **2(1)** (1992) 114–122; *translation from Tr. Inst. Mat. SO RAN* **21** (1992) 29–34

[35] W. Greub, S. Halperin, R. Vanstone: *Connections, curvature, and cohomology. Vol. I: De Rham cohomology of manifolds and vector bundles* (Pure and Applied Mathematics, 47. New York-London: Academic Press. XIX, 443 p., 1972)

[36] P. Griffiths, J. Harris: *Principles of algebraic geometry* (2nd ed. Wiley Classics Library. New York, NY: John Wiley & Sons Ltd. xii, 813 p., 1994)

[37] M. Gualtieri: Generalized complex geometry *Preprint* `arXiv:math.DG/0401221` (2004)

[38] V. W. Guillemin, S. Sternberg: *Supersymmetry and equivariant de Rham theory* (With reprint of two seminal notes by Henri Cartan. (English) Mathematics: Past and Present. Berlin: Springer. xxiii, 228 p., 1999)

[39] T. J. Harding, F. J. Bloore: Isomorphism of de Rham cohomology and relative Hochschild cohomology of differential operators *Ann. Inst. H. Poincarè Phys. Théor.* **58(4)** (1993) 433–452

[40] N. Hitchin: Generalized Calabi-Yau manifolds *Preprint* `arXiv:math.DG/0209099` (2002)

[41] A. El Kacimi-Alaoui, A. Tihami: Cohomologie bigraduée de certains feuilletages *Bull. Soc. Math. Belg., Sér. B* **38** (1986) 144–156

[42] F. W. Kamber, Ph. Tondeur: *Characteristic classes of foliated bundles* (Lect.Notes Math. V.494., 1975)

[43] F. W. Kamber, Ph. Tondeur: De Rham-Hodge theory for Riemannian foliations *Math. Ann.* **277** (1987) 415–431

[44] A. A. Kirillov: Local Lie algebras *Uspehi Mat. Nauk* **31(4(190))** (1976) 57–76; *Letter to the editors: Correction to "Local Lie algebras"* **31** (Uspehi Mat. Nauk (1976), no. 4(190), 57–76) Russian; *Uspehi Mat. Nauk* **32(1(193))** (1977) 268

[45] S. Kobayashi, K. Nomizu: *Foundations of differential geometry I* (New York-London: Interscience Publishers, a division of John Wiley & Sons. XI, 329 p., 1963)

[46] S. Kobayashi, K. Nomizu: *Foundations of differential geometry II* (New York-London-Sydney: Interscience Publishers a division of John Wiley and Sons, 1969)

[47] J.-L. Koszul: Crochet de Schouten-Nijenhuis et cohomologie *In: Élie Cartan et les mathématiques d'aujourd'hui, The mathematical heritage of Elie Cartan* (Semin. Lyon 1984, Astérisque, No.Hors Sér. 1985, 1985) 257–271

[48] I. S. Krasil'shchik, V. V. Lychagin, A. M. Vinogradov: *Geometry of jet spaces and nonlinear partial differential equations* (Advanced Studies in Contemporary Mathematics, 1. New York etc.: Gordon and Breach Science Publishers. xx, 441 p., 1986)

[49] I. S. Krasil'shchik (ed.), A. M. Vinogradov (ed.): *Symmetries and conservation laws for differential equations of mathematical physics* (Translations of Mathematical Monographs. 182. Providence, RI: American Mathematical Society (AMS). xiv, 333 p., 1999)

[50] D. Krupka: The total divergence equation *Lobachevskii Journal of Mathematics* **23** (2006) 71–93

[51] A. Kushner: Almost product structures and Monge-Ampere Equations *Lobachevskii Journal of Mathematics* **23** (2006) 151–181

[52] V. I. Kuz'minov, I. A. Shvedov: L_p-theory of differential forms on warped products *Sib. Adv. Math.* **8(3)** (1998) 74–89; *translation from Algebra, geometry, analysis and mathematical physics. 10th Siberian school, Novosibirsk* (1997) 110–120

[53] V. I. Kuz'minov, I. A. Shvedov: On compact solvability of the operator of exterior derivation *Sib. Math. J.* **38(3)** (1997) 492–506; *translation from Sib. Mat. Zh.* **38(3)** (1997) 573–590

[54] V. I. Kuz'minov, I. A. Shvedov: On normal solvability of the operator of exterior derivation of warped products *Sib. Math. J.* **37(2)** (1996) 276–287; *translation from Sib. Mat. Zh.* **37(2)** (1996) 324–337

[55] V. I. Kuz'minov, I. A. Shvedov: On the Künneth formula for reduced L_2-cohomologies *Sib. Math. J.* **36(1)** (1995) 92–99; *translation from Sib. Mat. Zh.* **36(1)** (1995) 102–110

[56] V. I. Kuz'minov, I. A. Shvedov: Addition formulas for reduced Lp-cohomologies *Sib. Math. J.* **35(2)** (1994) 344–350; *translation from Sib. Mat. Zh.* **35(2)** (1994) 380–388

[57] V. I. Kuz'minov, I. A. Shvedov: On normal solvability of the exterior differentiation on a warped cylinder *Sib. Math. J.* **34(1)** (1993) 73–82; *translation from Sib. Mat. Zh.* **34(1)** (1993) 85–95

[58] M. de León, J. C. Marrero, E. Padrón: Lichnerowicz-Jacobi cohomology of Jacobi manifolds *C. R. Acad. Sci. Paris Sér. I Math.* **324(1)** (1997) 71–76

[59] M. V. Losik: Characteristic classes of structures on manifolds *Funct. Anal. Appl.* **21** (1987) 206–216; *translation from Funkts. Anal. Prilozh.* **21(3)** (1987) 38–52

[60] M. V. Losik: A generalization of manifold and its characteristic classes *Funct. Anal. Appl.* **24(1)** (Russian, English), (1990) 26–32; *translation from Funkts. Anal. Prilozh.* **24(1)** (1990) 29–37

[61] M. V. Losik: Characteristic classes of transformation groups *Differ. Geom. Appl.* **3(3)** (1993) 205–218

[62] M. V. Losik: The cohomology of the complex of G-invariant forms on G-manifolds *Ann. Global Anal. Geom.* **13(4)** (1995) 323–338

[63] M. A. Malakhal'tsev: An analog of Dolbeault cohomologies for manifolds over the algebra of dual numbers *Izv. Vyssh. Uchebn. Zaved., Mat. 1990* **(11(342))** (1990) 82–84

[64] M. A. Malakhal'tsev: The Lie derivative and cohomology of G-structures *Lobachevskii Journal of Mathematics* **3** (1999) 215–220

[65] M. A. Malakhal'tsev: Differential complex associated to closed differential forms of nonconstant rank *Lobachevskii Journal of Mathematics* **23** (2006) 183–192

[66] B. Malgrange: Plongement des variétés analytiques-réelles *Bull. Soc. Math. France* **85** (1957) 101–112

[67] M. A. Mikenberg: Manifolds over the algebra of dual numbers whose canonical foliation has an everywhere dense leaf *Russ. Math.* **47(3)** (2003) 28–30; *translation from Izv. Vyssh. Uchebn. Zaved., Mat. 2003* **(3)** (2003) 31–33

[68] P. Molino: *Riemannian foliations* (Birkhäuser, 1988)

[69] P. Monnier: Poisson cohomology in dimension two *Israel J. Math.* **129** (2002) 189–207

[70] N. Nakanishi: Poisson cohomology of plane quadratic Poisson structures *Publ. RIMS, Kyoto Univ.* **33** (1997) 73–89

[71] A. Pichereau: Cohomologie de Poisson en dimension trois *C. R. Acad. Paris, Ser. I* **340** (2005) 151–154

[72] J. F. Pommaret: *Systems of partial differential equations and Lie pseudogroups* (Math. and Appl. 14, 1978)

[73] M. M. Postnikov: *Lectures on geometry. 3rd semester. Smooth manifolds* (Textbook. Moskva: Izdatel'stvo "Nauka". Glavnaya Redaktsiya Fiziko- Matematicheskoj Literatury. 480 p., 1987)

[74] C. Roger, P. Vanhaecke: Poisson cohomology of the affine plane *J. Algebra* **251** (2002) 448–460

[75] A. P. Shirokov: Geometry of tangent bundles and spaces over algebras *J. Sov. Math.* **21** (1983) 151–177

[76] A. P. Shirokov, V. V. Shurygin, V. V. Vishnevskij: *Spaces over algebras* ((Prostranstva nad algebrami), (Russian), Kazan: Izdatel'stvo Kazanskogo Universiteta. 263 p., 1985)

[77] V. V. Shurygin: Smooth manifolds over local algebras and Weil bundles *J. Math. Sci., New York* **108(2)** (2002) 249–294; *translation from Itogi Nauki Tekh., Ser. Sovrem Mat. Prilozh., Temat. Obz., VINITI, Moscow* **73** (2002) 162–236

[78] V. V. Shurygin: On the cohomology of manifolds over local algebras *Russ. Math.* **40(9)** (1996) 67–81; *translation from Izv. Vyssh. Uchebn. Zaved., Mat. 1996* **(412)** (1996) 71–85

[79] V. V. jun. Shurygin: Cohomology of the Brylinski double complex of Poisson manifolds and quandum de Rham cohomology *Russ. Math.* **48(10)** (2004) 70–76; *translation from Izv. Vyssh. Uchebn. Zaved., Mat. 2004* **(10)** (2004) 75–81
[80] V. V. jun. Shurygin: Poisson structures on Weil bundles *Lobachevskii J. Math.* **17** (electronic only), (2005) 229–256
[81] J. W. Smith: The de Rham theorem for general spaces *Tôhoku Math. J., II. Ser.* **18** (1966) 115–137
[82] S. E. Stepanov: Vanishing theorems in affine, Riemannian, and Lorentz geometries *Fundam. Prikl. Mat.* **11(1)** (2005) 35–84
[83] N. N. Tarkhanov: *The parametrix method in the theory of differential complexes* (Ed. by M. A. Shubin. (Metod parametriksa v teorii differentsial'nykh kompleksov.) Novosibirsk: Nauka. 248 p., 1990)
[84] L. Ugarte: Coeffective numbers of Riemannian 8-manifolds with holonomy in Spin(7) *Ann. Global Anal. Geom.* **19(1)** (2001) 35–53
[85] I. Vaisman: d_f -cohomologies of Lagrangian foliations *Manatschefte fur Mat.* **106** (1988) 221–244
[86] I. Vaisman: Remarks on the Lichnerowicz-Poisson cohomology *Ann. Inst. Fourier, Grenoble* **40(4)** (1990) 951–963
[87] I. Vaisman: *Cohomology and differential forms* (Translation editor: Samuel I. Goldberg. Pure and Applied Mathematics, 21. Marcel Dekker, Inc., New York, vii+284 pp., 1973)
[88] I. Vaisman: *Lectures on the Geometry of Poisson Manifolds* (Progress in Math. 118, Birkhäuser, Basel, 1994)
[89] A. Verona: Sur la cohomologie de de Rham des prestratifications de Whitney *C. R. Acad. Sci., Paris, Sér. A* **274** (1972) 1340–1343
[90] A. Verona: Le théorème de Rham pour les préstratifications abstraites *C. R. Acad. Sci., Paris, Sér. A* **273** (1971) 886–889
[91] F. W. Warner: *Foundations of differentiable manifolds and Lie groups* (Glenview, Illinois-London: Scott, Foresman & Comp. 270 p., 1983)
[92] A. Wassermann: Cyclic cohomology of algebras of smooth functions on orbifolds *In: Operator algebras and application. Vol. 1: Structure theory; K-theory, geometry and topology* (Pap. UK-US Jt. Semin., Warwick/UK 1987, Lond. Math. Soc. Lect. Note Ser. 135, 1988) 229–244
[93] A. Weinstein, P. Xu: Hochschild cohomology and characteristic classes for star-products *In: Geometry of differential equations* (Khovanskij, A. (ed.) et al., Dedicated to V. I. Arnold on the occasion of his 60th birthday. Providence, RI: American Mathematical Society. Transl., Ser. 2, Am. Math. Soc. 186(39), 1998) 177–194
[94] R. O. Wells: *Differential analysis on complex manifolds* (Prentice-Hall, Inc. Englewood Cliffs, N.J., 1973)
[95] P. Xu: Poisson cohomology of regular Poisson manifolds *Ann. Inst. Fourier Grenoble* **42** (1992) 967–988
[96] K. Yano, S. Bochner: *Curvature and Betti numbers* (Annals of Mathematics Studies 32. Princeton: N. J.: Princeton University Press IX, 190 p., 1953)

[97] N. Yeganefar: L_p-cohomology of negatively curved manifolds *Ark. Mat.* **43(2)** (2005) 427–434

M. A. Malakhaltsev

Kazan State University, Faculty of Mechanics and Mathematics,
Kremlevskaya, 18, Kazan:420008, Russia

E-mail: `mikhail.malakhaltsev@ksu.ru`

This page intentionally left blank

Handbook of Global Analysis
Demeter Krupka and David Saunders

Topology of manifolds with corners

J. Margalef-Roig and E. Outerelo Domínguez

Contents

1 Introduction

The study of manifolds with corners was originally developed by J. Cerf [4] and A. Douady [7] as a natural generalization of the concept of finite-dimensional manifold with smooth boundary. When one considers two manifolds with smooth boundary, its product is not a manifold with smooth boundary, but it has an ordinary structure of manifold with corners, and every finite product of manifolds with corners is again a manifold with corners.

Stability under finite products of manifolds with corners is of a great importance and one can consider enough by itself to justify the detailed study of this type of manifolds. For example, the important Schwartz Kernel Theorem (*The linear operators from smooth functions to distributions on a manifold with smooth boundary can be identified with distributions, namely their Schwartz kernels, on the product of the manifold with itself*) uses this result of stability of products in an essential way (See [22]).

Applications of the manifolds with corners in differential topology arise immediately after its definition, thus K. Jänich [15] uses this type of manifolds to the problem of classifying transformation group actions on smooth manifolds, and in the integration theory on manifolds a general formulation of the Stokes theorem is given in the setting of manifolds with corners (see J. M. Lee [19]).

On the other hand the ideas on calculus of variations developed by M. Morse in [25] lead to a change of direction in the study of manifolds. J. Eells in [8] proves that for a large variety of spaces $\mathcal{F}$ of functions from a compact topological space S into a Riemannian manifold M, the Riemannian structure on M determines a differentiable structure (necessarily infinite-dimensional) on $\mathcal{F}$. This is one of the first steps for the study of infinite dimensional manifolds and for the use of new methods in classical analysis, that historically can be trace back to the work of B. Riemann [26]. Thus a very natural task is to extend the results of finite-dimensional manifolds with corners to infinite dimension. The authors of the present chapter realized a quite systematic study of infinite-dimensional Banach manifolds with corners in [21].

Here, we survey the main features of the manifolds with corners modeled on Banach spaces or on larger categories of spaces as can be the normed spaces, the locally convex vector spaces and the convenient vector spaces, that have arisen as important, in the last years, in Global Analysis.

2 Quadrants

The local models to construct the manifolds with corners are open subsets of quadrants of topological vector spaces. Thus in this paragraph we introduce the quadrants in vector spaces and topological vector spaces and we survey the main properties of them.

All the vector spaces will be real vector spaces.However,we remark that the results established here can be extended to complex vector spaces by considering its real restrictions and carrying out adecuate adjustments of the notions considered.

Quadrants in vector spaces

2.1 Let E be a real vector space (rvs). A subset Q of E is called quadrant of E if there exist a basis $B = \{u_i\}_{i\in I}$ of E and a subset K of I such that

$$Q = L\{u_k\,|k \in K\} + \left\{ \sum_{i\in I-K} a_i u_i \,\middle|\, \sum_{i\in I-K} a_i u_i \textit{ has positive finite support} \right\}$$

(positive finite support means: There exists a finite subset F of $I - K$ such that $a_i > 0$ for all $i \in F$ and $a_i = 0$ for all $i \in (I - K) - F$). In this case, we say that the pair (B, K) is adapted to the quadrant Q.

Note that B is a subset of Q and

$$L\{u_k\,|k \in K\} \cap \left\{ \sum_{i\in I-K} a_i u_i \,\middle|\, \sum_{i\in I-K} a_i u_i \textit{ has positive finite support} \right\} = \{\overline{0}\}$$

($L\{u_k|k \in K\}$ is the vector subspace of E generated by $\{u_k|k \in K\}$).

Note that if E is non-trivial, then $\{\overline{0}\}$ is not a quadrant of E. If $E = \{\overline{0}\}$, then $Q = E$ is a quadrant of E.

2.2 Let Q be a quadrant of a rvs E and let

$$\left(B = \{u_i\}_{i\in I}, K \subset I\right) \quad \text{and} \quad \left(B' = \{u'_m\}_{m\in M}, N \subset M\right)$$

be pairs adapted to Q. Then:

(*i*) $L\{u_k|k \in K\} = L\{u'_n|n \in N\}$.

(*ii*) $card\,(I) = card\,(M)$, $card\,(K) = card\,(N)$, $card\,(I - K) = card\,(M - N)$.

(*iii*) There exists a bijective map $\sigma : I - K \longrightarrow M - N$ such that $u'_{\sigma(i)} = r_i u_i + x_i$ for all $i \in I - K$, where $x_i \in L\{u_k|k \in K\}$ and r_i is a positive real number.

(*iv*) If $K = \varnothing$, there exists a bijective map $\sigma : I \longrightarrow M$ such that $u'_{\sigma(i)} = r_i u_i$ for all $i \in I$, where r_i is a positive real number.

This result proves the consistence of the following definitions.

2.3 Let Q be a quadrant of a rvs E and $(B = \{u_i | i \in I\}, K \subset I)$ a pair adapted to Q. Then:

(*i*) $L\{u_k | k \in K\}$ is called the kernel of Q and is denoted by Q^0.

(*ii*) $card\,(I - K)$ is called the index of Q and is denoted by $index\,(Q)$. Finally $card\,(K)$ (i.e. the dimension of Q^0) is called coindex of Q and is denoted by $coindex\,(Q)$.

2.4 Let Q and Q' be quadrants of a rvs E such that $index\,(Q) = index\,(Q')$ and $coindex\,(Q) = coindex\,(Q')$. Then there exists a linear isomorphism $\alpha : E \longrightarrow E$ such that $\alpha(Q) = Q'$ and $\alpha\left(Q^0\right) = Q'^0$.

2.5

(*i*) Let Q be a quadrant of a rvs E and $(B = \{u_i | i \in I\}, K \subset I)$ a pair adapted to Q. Then there exists a linear isomorphism

$$\delta : E \longrightarrow Q^0 \times \mathbb{R}^{(I-K)}$$

such that $\delta(Q) = Q^0 \times \left(\mathbb{R}^{(I-K)}\right)^+$, where $\mathbb{R}^{(I-K)} =$

$$\left\{x \in \mathbb{R}^{I-K} \,|\, \textit{there is } F_x \textit{ finite subset of } I - K \textit{ with } x_i = 0 \textit{ for all } i \notin F_x \right\}$$

and

$$\left(\mathbb{R}^{(I-K)}\right)^+ = \left\{x \in \mathbb{R}^{(I-K)} | x_i \geqslant 0 \textit{ for all } i\right\},$$

and $\delta\left(Q^0\right) = Q^0 \times \{\bar{0}\}$.

(*ii*) Let Q and Q' be quadrants with finite indexes of a rvs E. Suppose that $index\,(Q) = index\,(Q')$. Then $coindex\,(Q) = coindex\,(Q')$, ($Q^0$ and Q'^0 have the same finite codimension).

2.6 New description of quadrants using linear maps

Let Q be a quadrant of a rvs E. Then there exists $\Lambda = \{\lambda_m | m \in M\}$, a linearly independent system of elements of $E^* = L(E, \mathbb{R})$, such that

$$Q = E^+_\Lambda = \{x \in E | \lambda_m(x) \geqslant 0 \textit{ for all } m \in M\},$$
$$Q^0 = E^0_\Lambda = \{x \in E | \lambda_m(x) = 0 \textit{ for all } m \in M\}$$

and $card\,(M) = index\,(Q)$. Therefore Q is a wedge set of E and, in particular, Q is a convex set of E.

Indeed, let $(B = \{u_i | i \in I\}, K \subset I)$ be a pair adapted to Q. For every $j \in I - K$, let λ_j be the element of E^* defined by: $\lambda_j(u_k) = 0$ for all $k \in I - \{j\}$ and $\lambda_j(u_j) = 1$. We take $M = I - K$ and all the statements are easily proved.

2.7

(*a*) Let E be a vector space, F a vector subspace of E^* that separates points of E and $\{x_1, ..., x_p\}$ a finite subset of E. Then $\{x_1, ..., x_p\}$ is a linearly independent system of elements of E if and only if there exists a finite subset $\{\lambda_1, ..., \lambda_p\}$ of F such that $\lambda_i(x_j) = \delta_{ij}$

(*b*) Let E be a vector space and $\{\lambda_1, ..., \lambda_p\}$ a finite subset of E^*. Then $\{\lambda_1, ..., \lambda_p\}$ is a linearly independent system of elements of E^* if and only if there exists a finite subset $\{x_1, ..., x_p\}$ of E such that $\lambda_i(x_j) = \delta_{ij}$.

(*c*) Let E be a rvs and $\Lambda = \{\lambda_1, ..., \lambda_p\}$ a finite linearly independent system of elements of E^*. Then E^+_Λ is a quadrant of E such that $index\left(E^+_\Lambda\right) = card\,(\Lambda) = p$ and $\left(E^+_\Lambda\right)^0 = E^0_\Lambda$. Indeed, by (b), we have $E = E^0_\Lambda \oplus L\{x_1, ..., x_p\}$.

(*d*) There exist a rvs E and a linearly independent system Λ of elements of E^* such that E^+_Λ is not a quadrant of E (see the example of **2.8**).

In fact we have: Let E be a rvs of infinite dimension. Then there exists a linearly independent system Λ of elements of E^* such that $E^+_\Lambda = \{\overline{0}\}$ (The example of **2.8** below, may be adapted).

2.8 Let Q be a quadrant of finite index of a rvs E. Then, if $\Lambda = \{\lambda_m | m \in M\}$ and $\Lambda' = \{\mu_p | p \in P\}$ are linearly independent systems of elements of E^* such that $Q = E^+_\Lambda = E^+_{\Lambda'}$, one verifies:

(*i*) $Q^0 = E^0_\Lambda = E^0_{\Lambda'}$.

(*ii*) $card(M) = card(P) = index(Q)$.

(*iii*) There exists a bijective map $\sigma : M \longrightarrow P$ such that $\mu_{\sigma(m)} = r_m \lambda_m$, where r_m is a positive real number, for all $m \in M$.

Remark In the preceding result, the statement "$index\,(Q)$ is finite" is an essential condition, as proves the following:

Example Let E be a rvs with $dim\,E = \aleph_0 = card\,(\mathbb{N})$. Then E is isomorphic to $\mathbb{R}^{(\mathbb{N})}$, E^* is isomorphic to $\mathbb{R}^{\mathbb{N}}$, and $dim\,E^* = 2^{\aleph_0} > \aleph_0$. Let $B = \{e_n | n \in \mathbb{N}\}$ be a basis of E. Then $Q = \{x \in E | x = \sum a_n e_n$, with $a_n \geqslant 0$ for all $n \in \mathbb{N}\}$ is a quadrant of E and $(B, \varnothing \subset \mathbb{N})$ is a pair adapted to Q, $Q^0 = \{\overline{0}\}$ and $index\,(Q) = \aleph_0$. Let us consider the set of projections, $P = \{p_n | n \in \mathbb{N}\}$ $(p_n(e_m) = \delta_{nm})$. Then P is a linearly independent system of elements of E^*, $Q = E^+_P$, $Q^0 = E^0_P$, and $index\,(Q) = \aleph_0 = card\,(P)$. For all $(m, n) \in \mathbb{N} \times \mathbb{N}$, let us consider $p^n_m \in E^*$ defined by: $p^n_m(e_n) = 1$, $p^n_m(e_{n+m}) = 1$, $p^n_m(e_{n+2m}) = 1$, $p^n_m(e_{n+3m}) = 1, \ldots$, and $p^n_m(e_i) = 0$for all $i \notin \{n, n+m, n+2m, n+3m, ...\}$. We know that the product mapping $\theta : E^* \rightarrow \mathbb{R}^{\mathbb{N}} = \prod_{n \in \mathbb{N}} \mathbb{R}$, $x^* \longmapsto (x^*(e_n))$,is a linear isomorphism. Hence $M = \{p^n_m | (m, n) \in \mathbb{N} \times \mathbb{N}\}$

is a linearly independent system of elements of E^*. Moreover $E_M^+ = Q$, $E_M^0 = Q^0$, $index\,(Q) = \aleph_0 = card\,(M)$ and the preceding property (*iii*) is not verified. Finally $P \cup \{\theta^{-1}(-1,-1,-1,\ldots)\}$ $(= \Lambda)$ is a linearly independent system of elements of E^* and $E_\Lambda^+ = \{\overline{0}\}$is not a quadrant of E, $(\theta^{-1}(-1,-1,-1,\ldots)(e_n) = -1$ for all $n \in \mathbb{N}$).

2.9

(*a*) Let Q be a quadrant of a rvs E and $(\{u_i | i \in I\}, K \subset I)$ a pair adapted to Q. Then:

(*i*) We know that $Q^0 = L\{u_k | k \in K\}$ is an intrinsic subset of Q (**2.2).** If $K = I$, then $Q^0 = E = Q$ and if $K = \varnothing$, then $Q^0 = \{\overline{0}\}$.

(*ii*) For all $n \in \mathbb{N}$, let Q^n be the set $\{x \in Q | \textit{there exist} x_0 \in Q^0, i_1, \ldots, i_n \in I - K \textit{ and } a_{i_1}, \ldots, a_{i_n} \in \mathbb{R} \textit{ such that } a_{i_1} > 0, \ldots, a_{i_n} > 0, \textit{ and } x = x_0 + a_{i_1}u_{i_1} + \cdots + a_{i_n}u_{i_n}\}$. It is clear that Q^n can be empty, but if $Q^n \neq \varnothing$ then $Q^m \neq \varnothing$ for all $m \in \mathbb{N}$ with $m < n$.

This definition of Q^n does not depend of the pair adapted to Q considered, (**2**.**2**).

(*b*) Let Q be a quadrant of a rvs E. Then:

(*i*) If $index\,(Q)$ is infinite, then $\{Q^n \,| n \in \mathbb{N} \cup \{0\}\}$ is a partition of the quadrant Q.

(*ii*) If $index\,(Q) = n$, $n \in \mathbb{N} \cup \{0\}$, then $Q^m = \varnothing$ for all $m > n$ and $\{Q^0, Q^1, \ldots, Q^n\}$ is a partition of Q.

(*c*) Let Q be a quadrant of a rvs E. Then:

(*i*) If $index\,(Q)$ is infinite and $x \in Q^n$, $(n \in \mathbb{N} \cup \{0\})$, we say that x has coindex n $(coindex\,(x) = n)$.

(*ii*) If $index\,(Q)$ is finite $(index\,(Q) = n)$ and $x \in Q^m$, $0 \leqslant m \leqslant n$, we say that x has *coindex* m and *index* $n - m$ $(index(x) = n - m)$. In this case (of $index\,(Q) = n$), $\{x \in Q \mid x \textit{ has index } n\} = Q^0$.

(*d*) Let Q be a quadrant with finite index n of a rvs E. Let $\Lambda = \{\nu_1, \ldots, \nu_n\}$ be a linearly independent system of elements of E^* such that $Q = E_\Lambda^+$, (consequently $Q^0 = E_\Lambda^0$ (**2.8**)). Then for $x \in Q$ we have: "*coindex* $(x) = m$ *if and only if* $card\,\{i | \nu_i(x) \neq 0\} = m$" and "*index* $(x) = p$ *if and only if* $card\,\{i | \nu_i(x) = 0\} = p$".

2.10 Product of quadrants

(*a*) Let Q be a quadrant of a rvs E and Q' a quadrant of a rvs E'. Then:

(*i*) $Q \times Q'$ is a quadrant of $E \times E'$, $(Q \times Q')^0 = Q^0 \times Q'^0$ and $(Q \times Q')^n = \bigcup\limits_{p+q=n} Q^p \times Q'^q$ for all $n \in \mathbb{N}$. Indeed if $(\{u_i \mid i \in I\}, K \subset I)$ is a pair adapted to Q and $(\{v_j \mid j \in J\}, L \subset J)$ is a pair adapted to Q', then

$$(\{(u_i, 0) \mid i \in I\} \cup \{(0, v_j) \mid j \in J\}, K \uplus L \subset I \uplus J)$$

is a pair adapted to $Q \times Q'$ ($\uplus$ is the disjoint union).

(*ii*) If $index\,(Q) = n$, $index\,(Q') = m$, $\Lambda = \{\lambda_1, ..., \lambda_n\}$ is a linearly independent system of elements of E^* with $Q = E^+_\Lambda$, and $M = \{\mu_1, ..., \mu_m\}$ is a linearly independent system of elements of E'^* with $Q' = E'^+_M$, (consequently $Q^0 = E^0_\Lambda$ and $Q'^0 = E'^0_M$), then $\{\lambda_i \cdot p_1 | i = 1, 2, ..., n\} \cup \{\mu_j \cdot p_2 | j = 1, 2, ..., m\}$ $(= \Lambda\, p_1 \cup M\, p_2)$ is a linearly independent system of elements of $(E \times E')^*$ and $(E \times E')^+_{\Lambda\, p_1 \cup M\, p_2} = E^+_\Lambda \times E'^+_M$, $(E \times E')^0_{\Lambda\, p_1 \cup M\, p_2} = E^0_\Lambda \times E'^0_M$.

(*b*) For all $i \in \{1, ..., n\}$, let Q_i be a quadrant of a rvs E_i. Then $\prod_{i=1}^{n} Q_i$ is a quadrant of the product real vector space $\prod_{i=1}^{n} E_i$, $\left(\prod_{i=1}^{n} Q_i\right)^0 = \prod_{i=1}^{n} Q_i^0$ and $\left(\prod_{i=1}^{n} Q_i\right)^p = \bigcup_{p_1 + ... + p_n = p} (Q_1^{p_1} \times ... \times Q_n^{p_n})$ for all $p \in \mathbb{N}$.

Of course (ii) of (*a*) is also generalized.

However an infinite product of quadrants is not, in general, a quadrant.

Example (1) $[0, \longrightarrow) = \{\, r \in \mathbb{R} \mid r \geqslant 0\}$ is a quadrant of $\mathbb{R}$.

(2) $\prod_{n \in \mathbb{N}} [0, \longrightarrow)_n (=P)$, where $[0, \longrightarrow)_n = [0, \longrightarrow)$ for all natural number n, is the set $\left\{\, x \in \mathbb{R}^{\mathbb{N}} \mid x_n \geqslant 0 \text{ for all } n \in \mathbb{N}\right\}$ and P is not a quadrant of $\mathbb{R}^{\mathbb{N}} (= E)$.

Remark Let Q be a quadrant of a rvs E. Then, if Λ, Λ' are linearly independent systems of elements of E^* such that $Q = E^+_\Lambda = E^+_{\Lambda'}$, one verifies that $Q^0 = E^0_\Lambda = E^0_{\Lambda'}$ (see **2.8**).

Dealing with the example, suppose that P is a quadrant of E. Let $\Lambda = \{p_n | n \in \mathbb{N}\}$ be the set of the projections (Λ is a linearly independent system of elements of E^*). Then $E^+_\Lambda = P$ and $E^0_\Lambda = \{\overline{0}\} = P^0$ (see the preceding remark).

Let $(\{u_i | i \in I\}, K \subset I)$ be a pair adapted to P. By **2.6**, there exists $\overline{\Lambda} = \{\lambda_m | m \in M\}$ a linearly independent system of elements of E^* such that $P = E^+_{\overline{\Lambda}}$, $P^0 = E^0_{\overline{\Lambda}}$ and $card\,(M) = index\,(P)$. In fact $M = I - K$ and for all $j \in M$, $\lambda_j\,(u_j) = 1$ and $\lambda_j\,(u_k) = 0$ for all $k \in I$ with $k \neq j$.

Then $E^0_{\overline{\Lambda}} = \{\overline{0}\}$, $K = \varnothing$, $card\,(I) = index\,(P)$, $u_i \in E^+_\Lambda$ for all $i \in I$ and, finally, there exists $i_0 \in I$ such that u_{i_0} has infinite positive coordinates. Let us consider $v = \left(x_1 - \frac{1}{2}x_1, x_2 - \frac{1}{3}x_2, \ldots, x_n - \frac{1}{n+1}x_n, \ldots\right)$, where $u_{i_0} = (x_1, x_2, \ldots x_n, \ldots)$. Then v, $u_{i_0} \in P$ and $v = a_1 u_{i_1} + \cdots + a_m u_{i_m}$, with $a_1 > 0, \ldots, a_m > 0$, and $u_{i_0} - v = u_{i_0} - a_1 u_{i_1} - \cdots - a_m u_{i_m} \in P$ and, consequently, $m = 1$, $u_{i_0} = u_{i_1}$, $a_1 < 1$ and $v = a_1 u_{i_0}$, which is a contradiction.

2.11 Let Q be a quadrant of a rvs E, $x \in Q$ and $v \in E$. Then we have one and only one of the following statements:

(*i*) There is $\varepsilon > 0$ such that $x + tv \in Q$ for all $t \in (-\varepsilon, \varepsilon)$.

(*ii*) There is $\varepsilon > 0$ such that $x + tv \in Q$ for all $t \in [0, \varepsilon)$ and $x + tv \notin Q$ for all $t < 0$.

(*iii*) There is $\varepsilon > 0$ such that $x + tv \in Q$ for all $t \in (-\varepsilon, 0]$ and $x + tv \notin Q$ for all $t > 0$.

(*iv*) For all $t \neq 0$, $x + tv \notin Q$.

Indeed, we take a pair $(B, K \subset I)$ adapted to Q, span x and v with respect to B and compare these expansions.

Quadrants in topological vector spaces

2.12 Characterization of quadrants with non-void interior

(i) Let Q be a quadrant of a real topological vector space (rtvs) E. Then the following statements are equivalent:

- (a) $int(Q) \neq \varnothing$ ($int(Q)$ is the interior of Q in E).
- (b) $index(Q)$ is finite and Q^0 is closed in E (we remark that if $index(Q) = 0$, then $Q^0 = Q = E$).

The step "$(a) \Longrightarrow (b)$" follows from the lemma:

Let Q be a quadrant of a rtvs E. Then:

- (1) *if $index(Q)$ is infinite, $int(Q) = \varnothing$, and*
- (2) *if $index(Q)$ is finite and Q^0 is not closed, $int(Q) = \varnothing$.*

Step "$(b) \Longrightarrow (a)$". Let $(\{e_i \mid i \in I\}, \; K \subset I)$ be a pair adapted to Q. Then $I - K = \{i_1, ... i_n\}$, where $index(Q) = n$, and $\Lambda = \{\nu_1, ..., \nu_n\}$ is a linearly independent system of elements of E^* with $Q = E_\Lambda^+$ and $Q^0 = E_\Lambda^0$, where $\nu_k(e_{i_k}) = 1$ and $\nu_k(e_j) = 0$ for all $j \in I$ with $j \neq i_k$, $k = 1, 2, ..., n$.

Moreover $Q^0 \times L\{e_{i_1}, ..., e_{i_n}\} \xrightarrow{+} E$ is a linear homeomorphism, since we apply the general result:

Let E be a topological vector space and F a closed vector subspace of finite codimension. Then every algebraic supplement G of F in E is a topological supplement of F in E. Consequently $\Lambda \subset \mathcal{L}(E, \mathbb{R})$ and

$$int(Q) = \left\{x_0 + \lambda_1 e_{i_1} + \cdots + \lambda_n e_{i_n} | x_0 \in Q^0, \; \lambda_1 > 0, ..., \lambda_n > 0\right\}.$$

(ii) Let Q be a quadrant of a rtvs E and $n \in \mathbb{N}$. Then the following statements are equivalent:

- (a) $index\,(Q) = n$ and Q^0 is closed in E.
- (b) There exists a linearly independent system $\Lambda = \{\lambda_1, \ldots \lambda_n\}$ of elements of $\mathcal{L}(E, \mathbb{R})$ such that $Q = E_\Lambda^+$. (See **2.7**)

Moreover if (b) is fulfilled, $Q^0 = E_\Lambda^0$.

Note that if (b) holds, then there exists a finite subset $\{x_1, ..., x_n\}$ of E such that $\lambda_i(x_j) = \delta_{ij}$ and consequently $E = E_\Lambda^0 \oplus L\{x_1, ..., x_n\}$. Let $B^0 = \left\{x_k^0 | k \in K\right\}$ be a basis of E_Λ^0 and consider the set $I = K \cup \{1, ..., n\}$. Then $B = B^0 \cup \{x_1, ..., x_n\}$ is a basis of E,$Q = L\left\{x_k^0 | k \in K\right\} + \left\{\sum_{i=1}^{n} a_i x_i | a_1 \geqslant 0, ..., a_n \geqslant 0\right\}$, $(B, K \subset I)$ is adapted to Q, $index\,(Q) = n$ and $Q^0 = E_\Lambda^0$.

The preceding result obviously holds for $n = 0$.

(*iii*) We remark that there exists a Hausdorff rtvs (Hrtvs) E such that $E \neq 0$ and $\mathcal{L}(E, \mathbb{R}) = 0$, [17]. Note that if Q is a quadrant in E with $int\,(Q) \neq \varnothing$, then $Q = E$. Finally, if Q is a quadrant in E with $index\,(Q) = n \in \mathbb{N}$, then $int\,(Q) = \varnothing$.

We see that in order to deal with quadrants the class of Hausdorff real topological vector spaces is a too much large class.

(*iv*) Let Q be a quadrant of a rtvs E. Then:

(*a*) $int\,(Q) \neq \varnothing$ and A non-void open subset of Q imply $int\,(A) \neq \varnothing$ (By (i), step "$(b) \implies (a)$", we have that $Q^0 \times L\{e_{i_1}, ..., e_{i_n}\} \xrightarrow{+} E$ is a linear homeomorphism).

(*b*) For every $y \in int\,(Q)$ and every $x \in Q$, $\{tx + (1-t)\,y | 0 \leqslant t < 1\} = [y, x) \subset int\,(Q)$.

(*v*) Let E be a rtvs and Q a quadrant of E of finite index, $n \in \mathbb{N}$, and closed kernel and let $(\{e_i | i \in I\}, K \subset I)$ be a pair adapted to Q $(I - K = \{i_1, ..., i_n\})$. Then the map

$$Q^0 \times L\{e_{i_1}, ..., e_{i_n}\} \xrightarrow{+} E$$

is a linear homeomorphism and

$$\left\{x_0 + \lambda_1 e_{i_1} + \cdots + \lambda_n e_{i_n} | x_0 \in Q^0, \lambda_1 > 0, ..., \lambda_n > 0\right\} = int\,(Q)\,.$$

This result and the topological characterization of Fréchet spaces obtained by H. Toruńczyk (*Every Fréchet space is homeomorphic to a Hilbert space*, see [27]), give the following:

Theorem *Every quadrant Q with non-void interior of a Fréchet space (in particular, a Banach space) is homeomorphic to a quadrant of a Hilbert space.*

Note that Q^0 is again a Fréchet space.

(*vi*) Let E be a rtvs and $n \in \mathbb{N}$. Let us consider the set $\{\{\lambda_i\}_{i \in I} \mid card(I) = n$ *and* $\{\lambda_i\}_{i \in I}$ *is a linearly independent system of elements of* $\mathcal{L}(E, \mathbb{R})\}$(=$A(E,\ n)$). Let us consider the binary relation on $A(E, n)$:

$\{\lambda_i\}_{i \in I} \sim \{\mu_j\}_{j \in J}$ if and only if there is a bijective map $\sigma : I \to J$ such that $\mu_{\sigma(i)} = \rho_i \lambda_i$, where ρ_i is a positive real number, for all $i \in I$.

Then $\sim$ is an equivalence relation. Moreover, the map $A(E, n)/\sim \longrightarrow \{Q \mid Q$ *is a quadrant in E of index n and closed kernel,*$[\{\lambda_i\}_{i \in I}] \mapsto \{x \in E | \lambda_i(x) \geqslant 0$ $for\ all\ i \in I\}$, is bijective (the inverse map is given by the preceding (ii)).

(*vii*) Let E be a rtvs, $n \in \mathbb{N}$ and Q a subset of E. Then Q is a quadrant in E with $index(Q) = n$ and Q^0 closed if and only if there is $\{\lambda_1, ..., \lambda_n\}$, a linearly independent system of elements of $\mathcal{L}(E, \mathbb{R})$, such that $Q = E^+_{\{\lambda_1, ..., \lambda_n\}}$.

(*viii*) Let E be a rtvs and $n \in \mathbb{N}$. Let us consider the set $\{(\{u_i\}_{i \in I}, K \subset I) \mid \{u_i\}_{i \in I}$ *is a basis of E,* $card(I - K) = n$ *and* $L\{u_k | k \in K\}$ *is closed in* $E\}$(=$B(E, n)$). Let us consider the binary relation on $B(E, n)$:

$(\{u_i\}_{i\in I}, K\subset I)\sim(\{v_j\}_{j\in J}, M\subset J)$ if and only if $L\{u_k|k\in K\}=L\{v_m|m\in M\}$ and there is a bijective map $\sigma : I-K\to J-M$ such that $v_{\sigma(i)}=r_iu_i+x_i$, $r_i>0$, $x_i\in L\{u_k|k\in K\}$, for all $i\in I-K$.

Then $\sim$ is an equivalence relation. Moreover, the map $B(E,n)/\sim\longrightarrow\{Q\mid Q$ *is a quadrant in* E *of index* n *and closed kernel*$\}$, $[(\{u_i\}_{i\in I}, K\subset I)]\mapsto L\{u_k|k\in K\}+\{\sum_{i\in I-K}a_iu_i|\sum_{i\in I-K}a_iu_i$ *has positive finite support*$\}$, is bijective (the inverse map is given by the preceding (*ii*)).

2.13

(*i*) Let E be a rtvs and Q a quadrant of E such that $\mathit{index}\,(Q)=n\in\mathbb{N}$ and Q^0 is closed in E. Then by **2.12** (*ii*) there exists a linearly independent system $\Lambda=\{\lambda_1,...,\lambda_n\}$ of elements of $\mathcal{L}(E,\mathbb{R})$ such that $Q=E^+_\Lambda$ and $Q^0=E^0_\Lambda$. Now by **2.7** (*b*), there exists a finite subset $\{x_1,...,x_n\}$ of E such that $\lambda_i\,(x_j)=\delta_{ij}$. Consequently $E^0_\Lambda\times L\,\{x_1,...,x_n\}\xrightarrow{+}E$ is a linear homeomorphism, (that is $E=E^0_\Lambda\oplus_T L\,\{x_1,...,x_n\}$), the map $\alpha : E^0_\Lambda\times\mathbb{R}^n\longrightarrow E$, $\alpha\,(x_0,r_1,...,r_n)=x_0+r_1x_1+\cdots+r_nx_n$, is a linear homeomorphism, $\alpha\left(E^0_\Lambda\times(\mathbb{R}^+\cup\{0\})^n\right)=E^+_\Lambda$, where $\mathbb{R}^+=\{x\in\mathbb{R}|x>0\}$, Q is closed in E, $\alpha\left(E^0_\Lambda\times(\mathbb{R}^+)^n\right)=\mathit{int}\,(Q)=\{x\in Q|\,\mathit{card}\,\{i|\lambda_i\,(x)>0\}=n\}=\{x\in Q|\,\mathit{index}\,(x)=0\}$,$\mathcal{F}_r\,(Q)=Q-\mathit{int}\,(Q)=\{x\in Q|\,\mathit{card}\,\{i|\lambda_i\,(x)=0\}\geq 1\}=\{x\in Q|\,\mathit{index}\,(x)\geqslant 1\}$, and E^+_Λ is homeomorphic to E^+_λ for all $\lambda\in\Lambda$ (In fact, $E^+_\Lambda\approx E^0_\Lambda\times(\mathbb{R}^n)^+_{\{p_1,...,p_n\}}\approx E^0_\Lambda\times(\mathbb{R}^n)^+_{p_1}\approx E^+_{\lambda_1}$, where $\approx$ means homeomorphism. Analogously for λ_2,...,λ_n).

(*ii*) Let E be a rtvs and $\Lambda=\{\lambda_1,\ldots\lambda_n\}$, $M=\{\mu_1,...,\mu_m\}$ finite linearly independent systems of elements of $\mathcal{L}\,(E,\mathbb{R})$ such that $E^+_\Lambda=E^+_M$. Then:

(1) E^+_Λ is a quadrant Q of E, such that $\mathit{index}\,(Q)=n$ and $Q^0=E^0_\Lambda$ and, consequently $n=m$, $E^0_\Lambda=E^0_M$ (**2.7** (*c*)), and Q^0 is closed in E.

(2) There exists a bijective map $\tau : \{1,...,n\}\longrightarrow\{1,...,n\}$ such that for every $i\in\{1,...,n\}$ there exists a positive real number r_i satisfying $\lambda_i=r_i\mu_{\tau(i)}$, (**2.8**).

(*iii*) Let Q and P be quadrants of a rtvs E such that Q^0 and P^0 are closed, and $\mathit{index}\,(Q)=\mathit{index}\,(P)=n\in\mathbb{N}$. Then, there exists a linear homeomorphism $\alpha : E\longrightarrow E$ such that $\alpha\,(Q)=P$ and $\alpha\left(Q^0\right)=P^0$ (We apply the following lemma : *Let E be a rtvs and F, G closed vector subspaces of E with the same finite codimension. Then F and G are linearly homeomorphic*).

Remark Let H be a real Hilbert space, $\lambda : H\longrightarrow\mathbb{R}$ a linear discontinuous map and $\mu : H\longrightarrow\mathbb{R}$ a linear continuous map with $\mu\neq 0$. Then by **2.7** (*c*), we have that H^+_λ and H^+_μ are quadrants of H of index 1 and $\left(H^+_\lambda\right)^0=\ker(\lambda)$, $(H^+_\mu)^0=\ker(\mu)$, but there is not a homeomorphism $\alpha : H\longrightarrow H$ such that $\alpha\,(\ker(\lambda))=\ker(\mu)$. However there exists a linear isomorphism $\beta : H\longrightarrow H$ such that $\beta\left(H^+_\lambda\right)=H^+_\mu$ and $\beta\left(H^0_\lambda\right)=H^0_\mu$, ($\ker(\lambda)$ is not closed in H).

(*iv*) Let Q be a quadrant of a rtvs E such that $index\,(Q) = n \in \mathbb{N}$ and Q^0 is closed in E. Let F be a rtvs and $\alpha : Q \longrightarrow F$ a continuous map such that $\alpha\,(x+y) = \alpha\,(x) + \alpha\,(y)$ for all $x, y \in Q$ and $\alpha\,(rx) = r\alpha\,(x)$ for all $x \in Q$ and for all non-negative real number r. Then, α is linear on Q^0 and there exists a unique continuous linear map $\overline{\alpha} : E \longrightarrow F$ such that $\overline{\alpha}|_Q = \alpha$.

(By the preceding (*i*), $Q^0 \times L\,\{x_1, ..., x_n\} \xrightarrow{+} E$ is a linear homeomorphism)

2.14 Let Q be a quadrant of a rtvs E, x an element of an open subset U of Q and v an element of E. Then we have one and only one of the following statements:

(*i*) There is $\varepsilon > 0$ such that $x + tv \in U$ for all $t \in (-\varepsilon, \varepsilon)$.

(*ii*) There is $\varepsilon > 0$ such that $x + tv \in U$ for all $t \in [0, \varepsilon)$ and $x + tv \notin Q$ for all $t < 0$.

(*iii*) There is $\varepsilon > 0$ such that $x + tv \in U$ for all $t \in (-\varepsilon, 0]$ and $x + tv \notin Q$ for all $t > 0$.

(*iv*) For all $t \in \mathbb{R}$ with $t \neq 0$, $x + tv \notin Q$ (see **2.11**)

Note that (*iv*) is not possible if $v \in Q \cup \{-Q\}$.

Let x be an element of an open subset U of a quadrant Q of a rtvs E. We introduce the following definitions:

(1) $A\,(Q, U, x) = \{v \in E \mid \textit{there is } \varepsilon > 0 \textit{ with } x + (-\varepsilon, \varepsilon) \cdot v \subset int\,(U)\}$.

(2) $B\,(Q, U, x) = \{v \in E \mid \textit{there is } \varepsilon > 0 \textit{ with } x + (-\varepsilon, \varepsilon) \cdot v \subset U\}$.

(3) $B^+\,(Q, U, x) = \{v \in E \mid \textit{there is } \varepsilon > 0 \textit{ with } x + [0, \varepsilon) \cdot v \subset U \textit{ and } x + tv \notin Q \textit{ for all } t < 0\}$.

(4) $B^-\,(Q, U, x) = \{v \in E \mid \textit{there is } \varepsilon > 0 \textit{ with } x + (-\varepsilon, 0] \cdot v \subset U \textit{ and } x + tv \notin Q \textit{ for all } t > 0\}$.

(5) $B^0\,(Q, U, x) = \{v \in E \mid x + tv \notin Q \textit{ for all } t \in \mathbb{R} \textit{ with } t \neq 0\}$.

Note that one verifies:

(*a*) $A\,(Q, U, x) \subset B\,(Q, U, x)$; if $A\,(Q, U, x) \neq \varnothing$, then $x \in int\,(Q)$; if $B\,(Q, U, x) = E$ and Q^0 is closed in E ,then $index\,(Q)$ is finite and $x \in int\,(Q)$.

(*b*) If $x \in int\,(Q)$, then $index\,(Q)$ is finite, Q^0 is closed in E, $x \in int\,(U) = U \cap int\,(Q)$, $A\,(Q, U, x) = B\,(Q, U, x) = E$ and $B^+\,(Q, U, x) = B^-\,(Q, U, x) = B^0\,(Q, U, x) = \varnothing$.

(*c*) If $x \notin int\,(Q)$ and Q^0 is closed in E, then $x \in \mathcal{F}_r\,(Q)$, $x \notin int\,(U)$, $A\,(Q, U, x) = \varnothing$, $Q^0 \subset B\,(Q, U, x)$, $B^+\,(Q, U, x) \neq \varnothing$, $B^-\,(Q, U, x) \neq \varnothing$, $Q \subset B\,(Q, U, x) \cup B^+\,(Q, U, x)$ and $-Q \subset B\,(Q, U, x) \cup B^-\,(Q, U, x)$.

(*d*) If $x \notin int(Q)$, Q^0 is closed in E and $coindex\,(x) + 2 \leqslant index\,(Q)$, then $B^0\,(Q, U, x) \neq \varnothing$.

(*e*) If $x \notin int\,(Q)$, Q^0 is closed in E and $index\,(Q) = n \in \mathbb{N}$ and $n - coindex\,(x) = 1$, then $B^0\,(Q, U, x) = \varnothing$ and $int\,(Q) \neq \varnothing$.

(*f*) If $x \notin int\,(Q)$ and Q^0 is not closed in E, then $x \in \mathcal{F}_r\,(Q)$, $x \notin int\,(U)$, $A\,(Q, U, x) = \varnothing$, $Q^0 \subset B\,(Q, U, x)$, $Q \subset B\,(Q, U, x) \cup B^+\,(Q, U, x)$ and $-Q \subset B\,(Q, U, x) \cup B^-\,(Q, U, x)$.

(*g*) If $x \notin int\,(Q)$, Q^0 is not closed in E, $index\,(Q) = n \in \mathbb{N}$ and $coindex\,(x) = n$, then $int\,(Q) = \varnothing$ and $B\,(Q, U, x) = E$.

(*h*) If $x \notin int\,(Q)$, Q^0 is not closed in E and $coindex\,(x) + 2 \leqslant index\,(Q)$, then $B^0\,(Q, U, x) \neq \varnothing$.

(*i*) If $x \notin int\,(Q)$, Q^0 is not closed in E, $index\,(Q) = n \in \mathbb{N}$ and $n - coindex\,(x) = 1$, then $B^0\,(Q, U, x) = \varnothing$.

2.15 Let U be an open subset of a quadrant Q of a rtvs E. Suppose that $int\,(Q) \neq \varnothing$, (**2.12** (*i*)). For all k with $0 \leqslant k \leqslant index\,(Q)$, one defines:

(*i*) $B_k(U) = \{x \in U \mid index\,(x) = k\}$, (**2.9**), (If $index\,(Q) = 0$, $B_0\,(U) = U$). It is clear that $B_0\,(U) = int\,(U) = U \cap int\,(Q)$ and $int\,(U)$ is dense in U. Finally $\{B_k(U) \mid 0 \leqslant k \leqslant index\,(Q)\}$ is a partition of U.

(*ii*) $\partial^k(U) = \{x \in U \mid index\,(x) \geqslant k\}$, (If $index\,(Q) = 0$, $\partial^0(U) = U$), and this set is called the k-boundary of U. It is clear that $\partial^k(U) = \bigcup_{j \geqslant k} B_j(U)$ is closed in U and $\partial^0(U) = U$. The set $\partial^1(U)$ is called the boundary of U and is denoted by $\partial(U)$.

Note that $B_k(U) = U \cap B_k(Q)$ and $\partial^k(U) = U \cap \partial^k(Q)$ for all k with $0 \leqslant k \leqslant index\,(Q)$.

Quadrants in locally convex real topological vector spaces

2.16 Let (E, T) be a Hausdorff locally convex rtvs (Hlcrtvs) and $\mathcal{V}\,(0) = \{V \subset E | V$ *is a bornivorous (i.e., absorbs the bounded sets) absolutely convex subset of* $(E, T)\}$. Then:

(*i*) $\mathcal{V}\,(0)$ is a 0-neighbourhood basis of a (unique) Hausdorff locally convex topology on E, denoted by T_{born} and called bornologification of (E, T). This topology, T_{born}, is finer than T ($T \subset T_{born}$).

(*ii*) (E, T) and (E, T_{born}) have the same collection of bounded sets, and (E, T_{born}) is the finest Hausdorff locally convex topology on E with this property. ([17], [18]).

Let (E, T) be a Hlcrtvs. Then:

(1) (E, T_{born}) is a bornological Hlcrtvs (A Hlcrtvs (E', T') is said to be a bornological space, if every bornivorous absolutely convex subset in (E', T') is a neighbourhood of 0 in (E', T')).

(2) (E, T) is bornological if and only if $T = T_{born}$.

(3) For any Hlcrtvs F, $\{h : (E, T_{born}) \to F \mid h \text{ is a linear continuous map}\} = \{f : (E,T) \to F \mid f \text{ is a bounded linear map}\}$.

(4) Let A be an absolutely convex set in E. Then A is a 0-neighbourhood in (E, T_{born}) if and only if A is a bornivorous subset.

(5) The continuous seminorms on (E, T_{born}) are exactly the bounded seminorms on (E, T).

(6) If F is a vector subspace of E of finite codimension, then $(T|_F)_{born} = (T_{born})|_F$.

Note that the topological product of at most countably many bornological spaces is again bornological.

Let E be a Hlcrtvs.Then: If E is metrizable, E is bornological. In particular, if E is a Fréchet (or normable) space, then E is bornological.

2.17 Let E be a Hlcrtvs and let $c : \mathbb{R} \longrightarrow E$ be a map (a curve). Then:

(*i*) c is called differentiable if the derivative, $c'(t) = \lim_{s \to 0} \frac{c(t+s)-c(t)}{s}$ at t, exists for all $t \in \mathbb{R}$, (consequently c is continuous).

(*ii*) We say that c is C^0 if c is continuous on $\mathbb{R}$. We say that c is C^1 if $c'(t)$ exists for all $t \in \mathbb{R}$ and c' is continuous on $\mathbb{R}$. We say that c is C^2 if c' and $(c')'(= c'')$ exist and c'' is continuous. In general, we say that c is C^n if $c', c'', \ldots, c^{(n}$ exist and $c^{(n}$ is continuous.

(*iii*) c is called smooth or C^∞ if all the iterated derivatives exist.

The set of C^p curves in E will be denoted by $C^p(\mathbb{R}, E)$, $(p \in \{0\} \cup \mathbb{N} \cup \{\infty\})$. We remark that if (E,T) is a Hlcrtvs,then (E,T) and (E, T_{born}) have the same smooth curves, [18].

Let (E,T) be a Hlcrtvs. We denote by $C^\infty T$, (which will be called C^∞-topology), the final topology in E induced by the family

$$C^\infty(\mathbb{R}, E) = \{c : \mathbb{R} \longrightarrow E \mid c \text{ is a smooth curve in } (E,T)\}$$

The topological space $(E, C^\infty T)$ will be also denoted by $C^\infty E$. One has that $T \subset T_{born} \subset C^\infty T$, even more $C^\infty(T_{born}) = C^\infty T$. Note that $(E, C^\infty T)$ is not, in general, a rtvs ($C^\infty \mathbb{R}^I$, where $card(I) \geqslant 2^{\aleph_0}$, is not completely regular (see [18])).

Let (E,T) be a metrizable Hlcrtvs. Then, $T = T_{born} = C^\infty T$. In particular, this takes place if (E,T) is a Fréchet space.

Note that $(\mathbb{R}^n, T_u^n)$ is a Hlcrtvs and $C^\infty T_u^n = T_u^n$. In general, the C^∞-topology of a product of two Hlcrtv spaces is not the product of the C^∞-topologies of the factor spaces. However the C^∞-topology of $E \times \mathbb{R}^n$, (where (E,T) is a Hlcrtvs), is the product topology of $C^\infty T$ by T_u^n (the usual topology of $\mathbb{R}^n$).

Let (E,T) be a Hlcrtvs and F a closed vector subspace of E. Then we have that the topology $C^\infty T$ induces in F the topology $C^\infty(T|_F)$. If F is not closed this result is not true in general.

Let (E,T) be a Hlcrtvs and X a subset of E. Let us consider the set

$$\mathcal{F} = \{c : \mathbb{R} \longrightarrow E \mid c \textit{ is a smooth curve in } (E,T) \textit{ with } im(c) \subset X\} .$$

The final topology in X generated by the family $\mathcal{F}$ will be called $C^\infty-$topology of X and the corresponding topological space will be denoted by $C^\infty X$. This topology contains the topology $(C^\infty T)|_X$, but, in general, they do not coincide. If X is open of $C^\infty T$ or X is convex and locally closed in $(E, C^\infty T)$, then the $C^\infty-$topology of X is equal to $(C^\infty T)\,|_X$.

Moreover, if U is a convex subset of a bornological Hlcrtvs E, then U is C^∞-open if and only if U is open in E (use the following result: *Let V be an absolutely convex subset of a bornological Hlcrtvs (E,T). Then V is a 0-neighbourhood of (E,T) if and only if V is 0-neighbourhood in $C^\infty E$*).

2.18 Let (E,T) be a Hlcrtvs and K a convex set of E with non-void $C^\infty-$interior, that is, $int_{C^\infty T}\,(K) \neq \varnothing$. Then, we have that the segment $(x,y] = \{x + t(y-x) | 0 < t \leqslant 1\} \subset int_{C^\infty T}\,(K)$ for all $x \in K$ and all element y of $int_{C^\infty T}\,(K)$. The $C^\infty-$interior of K is convex and open even in (E, T_{born}), (weaker than $(E, C^\infty T)$), and K is closed in (E, T_{born}) if and only if it is closed in $(E, C^\infty T)$, [18].

2.19

(*i*) Let Q be a quadrant in a Hlcrtvs (E,T). Then the following statements are equivalent:

(*a*) The $C^\infty-$interior of Q is non-void, (that is $int_{C^\infty T}\,(Q) \neq \varnothing$).

(*b*) $index\,(Q)$ is finite and Q^o is closed in (E, T_{born}) (**2.12, 2.18**).

Note that if (*a*) is fulfilled, $int_{T_{born}}\,(Q) \neq \varnothing$.

(*ii*) Let Q be a quadrant in a Hlcrtvs (E,T) and $n \in \mathbb{N}$. Then the following statements are equivalent:

(1) $index\,(Q) = n$ and Q^o is closed in (E, T_{born}).

(2) There exists a linearly independent system $\Lambda = \{\lambda_1, ..., \lambda_n\}$ of elements of $LB_T(E, \mathbb{R})$ (the space of bounded linear maps from $(E,T\ )$ to $\mathbb{R}$) such that $Q = E^+_\Lambda$.

Moreover if (2) holds, $Q^0 = E^0_\Lambda$. (Apply **2.12** (*ii*) and **2.16**).

Finally if (2) holds, then:

(*a*) Q and Q^0 are closed in (E, T_{born}) and therefore in $(E, C^\infty T)$.

(*b*) $int_{C^\infty T}\,(Q) = int_{T_{born}}\,(Q)$, (see **2.18**).

(*c*) The $C^\infty-$topology on Q (**2.17**), coincides with $(C^\infty T)|_Q$ and the C^∞-topology on Q^0 coincides with $(C^\infty T)\,|_{Q^0}$.

(*d*) Let $\{x_1, ..., x_n\}$ be a subset of E such that $\lambda_i(x_j) = \delta_{ij}$. Then, by (*a*), $Q^0 \times L\{x_1, ..., x_n\} \to E, (x_0, a_1x_1 + ... + a_nx_n) \mapsto x_0 + a_1x_1 + ... + a_nx_n$, is a linear homeomorphism with the topology T_{born}. Consequently $\alpha : Q^0 \times \mathbb{R}^n \to E, (x_0, a_1, ..., a_n) \mapsto x_0 + a_1x_1 + ... + a_nx_n$, is a linear homeomorphism with the topology T_{born} and $\alpha : Q^0 \times \mathbb{R}^n \to E$ is also a linear homeomorphism with the topology $C^\infty T$, (again by (*a*)), and $\alpha\left(Q^0 \times (\mathbb{R}^n)^+_{\{p_1,...,p_n\}}\right) = Q$.

Let us prove the last linear homeomorphism: Let U be an open set of $C^\infty T$. We need to prove that $\alpha^{-1}(U)$ is open in $Q^0 \times \mathbb{R}^n$ with the topology $(C^\infty T)|_{Q^0} \times T^n_u$, but we know that this topology coincides with $C^\infty\left(T|_{Q^0} \times T^n_u\right)$ which in its turn coincides with $C^\infty\left(T_{born}|_{Q^0} \times T^n_u\right)$. Also we know that $(C^\infty T)|_{Q^0} = C^\infty\left(T|_{Q^0}\right)$ which in its turn is the $C^\infty-$topology of Q^0 generated by the family

$$\left\{c : \mathbb{R} \longrightarrow E \mid c \textit{ is a smooth curve in } (E, T) \textit{ with } im(c) \subset Q^0\right\}.$$

Finally, recall that $\left(T|_{Q^0} \times T^n_u\right)_{born} = T_{born}|_{Q^0} \times T^n_u$, and the spaces $Q^0 \times \mathbb{R}^n$ with the topologies $T|_{Q^0} \times T^n_u$ and $T_{born}|_{Q^0} \times T^n_u$ have the same smooth curves. Let $c : \mathbb{R} \longrightarrow Q^0 \times \mathbb{R}^n$ be a smooth curve respect the topology $T|_{Q^0} \times T^n_u$. Then $\alpha \circ c : \mathbb{R} \longrightarrow E$ is a smooth curve respect the topology T_{born} in E (also respect to T). Consequenly $(\alpha \circ c)^{-1}(U) = c^{-1}\alpha^{-1}(U)$ is an open set of $\mathbb{R}$, which ends the proof.

If we consider α^{-1}, the same argument proves its continuity.

(*e*) Let U be a non-void element of $(C^\infty T)|_Q$. Then $int_{C^\infty T}(U)$ is also non-void.

3 Differentiation theories

The second basic tool (after the quadrants) to reach the notion of differentiable manifold with corners, is a differentiation theory on open subsets of quadrants in topological vector spaces.

In this field there are many options, for infinite dimensional topological vector spaces, that are generalizations of the classical calculus in euclidean spaces. We only consider four of them, and the choice is based upon in a larger use in analysis global today.

Differentiation theories in normable spaces

3.1

(*i*) Let x be an element of an open subset A of a quadrant Q in a rtvs E. Then x is cluster point of $A - \{x\}$.

(*ii*) Let U be an open subset of a quadrant Q in a normable (with the norm $\|\ \|$) rtvs E, F a Hlcrtvs defined by the collection of seminorms$\{\|\ \|_i \,|i \in I\}$, $f : U \longrightarrow F$ a map, $x \in U$ and $u \in \mathcal{L}(E, F)$. Then, we say that u is tangent to f at x if $\lim_{y\to x}(f(y) - f(x) - u(y - x)) \parallel y - x \parallel^{-1}= 0 \in F$ (which is equivalent to $\lim_{y\to x} \parallel f(y) - f(x) - u(y - x) \parallel_i \parallel y - x \parallel^{-1}= 0 \in \mathbb{R}$, for all $i \in I$).

Note that to be tangent to f at x is independent of the equivalent norms considered in E. Moreover, if u, v are tangent to f at x, then $u = v$ (that is, if the tangent to f at x exists, it is unique).

(*iii*) Let U be an open subset of a quadrant Q in a normable rtvs E, F a Hlcrtvs, $f : U \longrightarrow F$ a map, and $x \in U$. If there exists $u \in \mathcal{L}(E, F)$ such that u is tangent to f at x, we say that f is differentiable (or Fréchet differentiable) at x, we denote u, (the tangent to f at x), by $Df(x)$ and, finally, we say that $Df(x)$ is the differential of f at x. If f is differentiable at every $x \in U$, f is said to be differentiable on U and in this case we have the map $Df : U \longrightarrow \mathcal{L}(E, F)$ defined by $x \mapsto Df(x)$. Obviously if $Q = E$, then U is an open set of E and the preceding notion coincide with the similar classical notions given in [2]. This coincidence also occurs when the open subset U of the quadrant Q is an open subset of E. The above definitions do not depend of the quadrant Q in E that contains U and such that U is open in Q, that is, if Q and P are quadrants in E and U is open in Q and in P, then the preceding definitions coincide for both Q and P.

The preceding notions are independent of the equivalent norms considered in E.

(*iv*) If f is differentiable at x ((*iii*)), then f is continuous at x.

3.2 Let f be differentiable at x, (**3.1** (*iii*)). Then:

(*i*) For all $u \in B(Q, U, x)$, (**2.14**), (recall that $x \in int(U)$ implies $B(Q, U, x) = E$),

$$\lim_{t \to 0} \frac{f(x + tu) - f(x)}{t} = Df(x)(u).$$

(*ii*) Let u be an element of $B^+(Q, U, x)$,(in this case it must be $x \notin int(Q)$). Then

$$\lim_{t \to 0^+} \frac{f(x + tu) - f(x)}{t} = Df(x)(u), \ (\mathbf{2.14}).$$

(*iii*) Let u be an element of $B^-(Q, U, x)$,(in this case it must be $x \notin int(Q)$). Then

$$\lim_{t \to 0^-} \frac{f(x + tu) - f(x)}{t} = Df(x)(u), \ (\mathbf{2.14}).$$

Note Let us consider E, Q, U, F, $f : U \longrightarrow F$, $x \in U$ as in (*iii*) of **3.1**. Suppose that $E = \mathbb{R}$, (of course Q may be $(\longleftarrow, 0]$ or $[0, \longrightarrow)$ or $\mathbb{R}$, and suppose that $f'(x)$ exists, (lateral derivatives when it is demanded in the cases $x = 0$. Then f is differentiable at x and $Df(x)(r) = rf'(x)$ for all $r \in \mathbb{R}$.

3.3 Let U be an open subset of a quadrant Q of a normable rtvs (with the norm $\| \ \|$) E, F a Hlcrtvs (defined by the seminorms $\{\| \ \|_i \, | i \in I\}$), and $f : U \longrightarrow F$ a map. We know that $\mathcal{L}(E, F)$ is a Hlcrtvs defined by the family of seminorms $\left\{\| \ \|_i^* \, | i \in I\right\}$, where $\|\lambda\|_i^* = Supreme \{\|\lambda(x)\|_i \mid \|x\| \leqslant 1\}$.

Then we give the following definitions:

(1) f is of class 0, $\left(C^0\right)$, or a C^0-map if f is continuous on U.

(2) f is of class 1, $\left(C^1\right)$, or a C^1-map if f is differentiable on U and $Df : U \longrightarrow \mathcal{L}(E,F)$ is C^0.

(3) f is of class 2, $\left(C^2\right)$, or a C^2-map if f is differentiable on U and $Df : U \longrightarrow \mathcal{L}(E,F)$ is C^1 (which is equivalent to: f is differentiable on U, Df is differentiable on U and $D(Df) : U \longrightarrow \mathcal{L}(E, \mathcal{L}(E,F))$ is continuous).

Note We know that $\mathcal{L}^2(E,F)$ is a Hlcrtvs defined by the collection of seminorms $\left\{\| \|_i^{**} | i \in I\right\}$, where $\|\lambda\|_i^{**} = Supreme\left\{\|\lambda(x,y)\|_i \,|\, \|x\| \leqslant 1, \|y\| \leqslant 1\right\}$.

Moreover $\theta : \mathcal{L}(E,\mathcal{L}(E,F)) \longrightarrow \mathcal{L}^2(E,F)$, $\theta(g)(x,y) = g(x)(y)$, is a linear homeomorphism and $\theta \circ D(Df) : U \longrightarrow \mathcal{L}^2(E,F)$ will be denoted $D^2 f$.

(4) If r is a natural number with $r \geqslant 1$, f is of class r, (C^r), or a C^r-map if f, Df, $D^2 f, ..., D^{r-1} f$ are diferentiable maps on U and $D^r f : U \longrightarrow \mathcal{L}^r(E,F)$ is continuous, (inductive process).

(5) f is of class ∞, (C^∞), or a C^∞-map if f is a C^r-map for all $r \in \mathbb{N} \cup \{0\}$.

One has the following properties:

(*i*) If F is normable, $int(Q) \neq \varnothing$, $D^{r-1} f$ exists on U ($r \geqslant 2$) and $D^{r-1} f$ is differentiable at $x \in U$, then $D^r f(x)$ is a symmetric r-linear continuous map of E^r into F.

(*ii*) Suppose that f is C^r on U, then for all $p \in \mathbb{N} \cup \{0\}$ with $p \leqslant r$, f is C^p on U and $D^p f$ is C^{r-p} on U.

(*iii*) Suppose that $f = \lambda|_U$, where $\lambda \in \mathcal{L}(E,F)$. Then f is C^∞ on U and $Df(x) = \lambda$, $D^r f(x) = 0$ for all element x of U and all natural number r with $r > 1$.

(*iv*) If U is an open subset of a quadrant Q in a normable space E, which is product of the normable spaces E_1 and E_2, and $f = \lambda|_U$, where $\lambda : E_1 \times E_2 \to F$ is a continuous bilinear map, then f is a C^∞-map on U, and $Df(x_1,x_2)(v_1,v_2) = \lambda(x_1,v_2) + \lambda(v_1,x_2)$ for all $(x_1,x_2) \in U$ and all $(v_1,v_2) \in E_1 \times E_2$. Moreover, $D^2 f(x_1,x_2)\left\{(v_1,v_2),(u_1,u_2)\right\} = \lambda(v_1,u_2) + \lambda(u_1,v_2)$ and $D^r f(x_1,x_2) = 0$ for all $r > 2$. This result can be extended to n-linear continuous maps.

3.4 In order to build manifolds with corners modeled on normable spaces, we explain four basic properties of the C^r-maps.

(*i*) **Restrictions to open sets** If $f : U \longrightarrow F$ is a C^r-map, (**3.3**), and V is an open set of U, then $f|_V : V \longrightarrow F$ is a C^r-map and, (if $r \geqslant 1$) $D(f|_V)(x) = Df(x)$ for all $x \in V$.

(*ii*) **Open covering property** If $\{V_j | j \in J\}$ is an open covering of U (open subset of Q), then $f : U \longrightarrow F$ is a C^r-map if and only if $f|_{V_j} : V_j \longrightarrow F$ is a C^r-map for all $j \in J$.

(*iii*) **Chain rule** Let us consider E_1, Q_1, U_1 as in (iii) of **3.1** and E_2, Q_2, U_2 as in (iii) of **3.1**. Let $f : U_1 \to U_2$ and $g : U_2 \to G$ be maps, where G is a Hlcrtvs.Then if f is differentiable at $x_0 \in U_1$ and g is differentiable at $f(x_0)$, one verifies that $g \circ f$ is differentiable at x_0 and

$$D(g \circ f)(x_0) = Dg(f(x_0)) \circ Df(x_0).$$

Moreover if f and g are C^r, $r \in \{0\} \cup \mathbb{N} \cup \{\infty\}$, then $g \circ f$ is C^r.

(*iv*) **Restrictions to vector subspaces** Let $f : U \longrightarrow F$ be a C^r-map (**3.3**). Let G be a vector subspace of E and Q_G a quadrant of G. Suppose that $U \cap G$ is an open subset of Q_G. Then $f|_{U \cap G} : U \cap G \longrightarrow F$ is a C^r-map and, (if $r \geqslant 1$), $D(f|_{U \cap G})(x) = Df(x)|_G$ for all $x \in U \cap G$.

(*i*) and (*iv*) are consequence of (*iii*).

Note In the study of properties of manifolds with corners are also useful the following results:

(1) If $f : U \to F$ is a map (**3.3**) and F_1 is a closed vector subspace of F such that $f(U) \subset F_1$, then $f : U \to F_1$ is a C^r-map if and only if $f : U \to F$ is a C^r-map (**3.3**). (See **3.2**). Of course $\mathcal{L}(E, F_1)$ is a closed vector subspace of $\mathcal{L}(E, F)$.

(2) Let $\{f_i : U \to F_i\}_{i \in I}$ be a collection of maps, where U is an open subset of a quadrant Q of a normable space E and $\{F_i\}_{i \in I}$ is a family of Hlcrtv spaces. Then the map $(f_i)_{i \in I}(= f) : U \to \prod_{i \in I} F_i$ is differentiable at $x \in U$ if and only if f_i is differentiable at x for all $i \in I$ (**3.1**). In this case $Df(x)(v) = (Df_i(x)(v))_{i \in I}$ for all $v \in E$. Finally, f is C^r on U if and only if f_i is C^r on U for all $i \in I$.

3.5 Let U be an open subset of a quadrant Q of a normable rtvs E, and V an open subset of a quadrant P of a normable rtvs G. We say that a map $f : U \to V$ is a C^r-diffeomorphism if f is a bijective map and f, $f^{-1}$are C^r-maps (**3.3**). In this case, it is clear that $f^{-1} : V \to U$ is also a C^r-diffeomorphism.

Of course the composition of C^r-diffeomorphisms is a C^r-diffeomorphism, and the identity map is a C^r-diffeomorphism. The chain rule (**3.4** (*iii*)) implies that: If $f : U \to V$ is a C^r-diffeomorphism ($r \geqslant 1$), then $Df(x) : E \to G$ is a linear homeomorphism and $(Df(x))^{-1} = D(f^{-1})(f(x))$ for all $x \in U$.

"The boundary is preserved by diffeomorphisms". This statement is explained in the sequel.

3.6 Invariance of the boundary for C^r-diffeomorphisms

Let U be an open subset of a normable rtvs E, F a Hlcrtvs, $f : U \to F$ a map and $\lambda \in \mathcal{L}(F, \mathbb{R})$ with $\lambda \neq 0$. Suppose that $f(U) \subset F_\lambda^+$, f is differentiable at $x \in U$ and $f(x) \in F_\lambda^0$. Then $Df(x)(E) \subset F_\lambda^0$. (See [21]).

Note If we put "U is an open subset of E_Λ^+, $\Lambda = \{\lambda_1, ..., \lambda_n\}$" instead of "$U$ is an open subset of E", then we obtain $Df(x)(E_\Lambda^+) \subset F_\lambda^+$.

Theorem *Let $f : U \to V$ be a C^r-diffeomorphism ($r \geqslant 1$), (**3.5**), where $int(Q) \neq \varnothing$, $int(P) \neq \varnothing$. Then we have:*

(a) *$index(x) = index(f(x))$ for all $x \in U$* (**2.9**).

(b) *$\partial U \neq \varnothing$ if and only if $\partial V \neq \varnothing$* (**2.15**). *Moreover, $f(\partial^k U) = \partial^k V$ for all k with $0 \leqslant k \leqslant index(Q)$.*

(c) *$int(U) \neq \varnothing$ if and only if $int(V) \neq \varnothing$. Moreover, $f(B_k U) = B_k V$ for all k with $0 \leqslant k \leqslant index(Q)$* (**2.15**, **2.13** (*i*)).

(d) *$f|_{int(U)} : int(U) \to int(V)$ is a C^r-diffeomorphism and $D(f|_{int(U)})(x) = Df(x)$, for all $x \in int(U)$.*

A non-void open subset of a quadrant Q of a normable rtvs E with $int(Q) \neq \varnothing$ is the simplest example of C^∞-manifold with corners modeled on normable rtv spaces and this type of manifolds are the local models.

Note If in the preceding theorem, $0 \in U$ and $0 \in V$, then $index(Q) = index(P)$.

Corollary *Let Q and P be quadrants of a normable rtvs E with $int(Q) \neq \varnothing$ and $int(P) \neq \varnothing$. Then $index(Q) = index(P)$ if and only if P, Q are C^∞-diffeomorphic.*

It follows from the preceding theorem and **2.13** (*iii*).

3.7 Inverse mapping theorem for quadrants in Banach spaces

Let U be an open subset of a quadrant Q of a banachable rtvs E ([21]) with $int(Q) \neq \varnothing$, P a quadrant of a banachable rtvs G with $int(P) \neq \varnothing$, $f : U \to P$ a C^r-map ($r \geqslant 1$) and $x \in U$ (**3.3**). Suppose that there exists an open neighbourhood V of x in U such that $f(V \cap \partial U) \subset \partial P$ (**2.15**), and that $Df(x) : E \to G$ is a linear homeomorphism (by **2.12** (*ii*), $Q = E^+_\Lambda$ and $P = G^+_M$). Then there exist an open neighbourhood U_1 of x in U and an open neighbourhood U' of $f(x)$ in P such that $f(U_1) = U'$ and $f|_{U_1} : U_1 \to U'$ is a C^r-diffeomorphism (**3.5**), [21].

Note that $\partial V = V \cap \partial U$.

3.8

(1) Let us consider E, Q, U, F, f as in **3.3.** Suppose that $E = \mathbb{R}^2$ and $Q = \{(x, y) \in \mathbb{R}^2 | x \geqslant 0, y \geqslant 0\}$ (or $Q = \{(x, y) \in \mathbb{R}^2 | x \geqslant 0\}$ or $Q = \{(x, y) \in \mathbb{R}^2 | y \geqslant 0\}$ or $Q = \mathbb{R}^2$). Then the following statements are equivalent:

  (*i*) $f : U \to F$ is a C^1-map on U (**3.3**).

  (*ii*) There exist the maps $D_1 f : U \to F$, $D_2 f : U \to F$ (partial derivatives) and they are continuous.

Moreover, if (*ii*) is true, then $Df(x, y)(s, t) = sD_1 f(x, y) + tD_2 f(x, y)$ for all $(x, y) \in U$ and all $(s, t) \in \mathbb{R}^2$.

(2) Let us consider E, Q, U, F, f as in the preceding (1). Then the following statements are equivalent:

  (*i*) f is C^r on U ($r \geqslant 1$), (**3.3**).

  (*ii*) There exist the maps $\frac{\partial^{p+q} f}{\partial^p t \partial^q s}$, $p \geqslant 0$, $q \geqslant 0$, $1 \leqslant p + q \leqslant r$ from U to F, and these maps are continuous.

(3) Let us consider E, Q, U, F, f as in the preceding (1). Suppose that f is C^1 on U and $Df : U \to \mathcal{L}(\mathbb{R}^2, F)$ is differentiable at $x_0 \in U$ (**3.1**). Then $D_1 f$, $D_2 f : U \to F$ are differentiable at x_0 and $D_1(D_2 f)(x_0) = D_2(D_1 f)(x_0)$. Moreover for any fixed $(s, t) \in \mathbb{R}^2$, the map, $(x, y) \mapsto Df(x, y)(s, t)$, from U to F, is differentiable at x_0 and $D[Df(\cdot)(s, t)](x_0)(s', t') = D^2 f(x_0)((s', t'), (s, t))$.

This results can be generalized to finite products of normed spaces as it is explained in the next item.

3.9 Let $(E_1, \| \|_1), ..., (E_n, \| \|_n)$ be normed spaces, and $Q_1, ..., Q_n$ quadrants of $E_1, ..., E_n$, respectively, such that $int(Q_i) \neq \varnothing$ for all $i \in \{1, ..., n\}$. Let U be an open subset of the quadrant $Q = Q_1 \times ... \times Q_n$ (**2.10**) of the normed space $(E, \| \|)$ product of the given normed spaces, $x = (x_1, ..., x_n)$ an element of U and $f : U \to F$ a map, where F is a Hlcrtvs. One defines the partial derivatives of f at x, $D_1 f(x), ..., D_n f(x)$, as usually (whenever they exist).

Then, if f is differentiable at x (**3.1** (*iii*)), the partial derivatives of f at x exist and $Df(x)(v_1, ..., v_n) = D_1 f(x)(v_1) + ... + D_n f(x)(v_n)$ for all element $(v_1, ..., v_n)$ of E (of course $D_i f(x) \in \mathcal{L}(E_i, F)$).

By induction one defines the partial derivatives

$D_{i_m} ... D_{i_1} f(x) \in \mathcal{L}^m(E_{i_1}, ..., E_{i_m}; F)$, for all $i_m, ..., i_1 \in \{1, ..., n\}$. (If $D^{m-1} f$ exists on U and is differentiable at $x \in U$, then we can change the order of partial derivatives).

Finally we have: *f is C^r on U if and only if f has continuous partial derivatives $D_{i_m} ... D_{i_1} f$ on U for all $\{i_1, ..., i_m\} \subset \{1, ..., n\}$ and all $m \leqslant r$.*

The preceding results can be formulated for arbitrary quadrants, with non-void interior, in a finite product of normable rtv spaces. In this case we use the linear homeomorphism that transforms the given quadrant in a quadrant that is product of a finite number of quadrants (**2.13** (*iii*)).

3.10

Lemma (Seeley [21]) *There are two sequences of real numbers $\{a_n\}$ and $\{b_n\}$ such that:*

(i) $\{b_n\}$ *is a strictly decreasing sequence of negative real numbers that converge to* $-\infty$.

(ii) $\sum_{n=0}^{\infty} |a_n||b_n|^p < \infty$ *for all* $p \in \mathbb{N} \cup \{0\}$.

(iii) $\sum_{n=0}^{\infty} a_n (b_n)^p = 1$ *for all* $p \in \mathbb{N} \cup \{0\}$.

Using this Lemma and the preceding characterization of C^r-maps, one proves the following result:

Proposition *Let U be an open subset of a quadrant Q with non-void interior of a normable rtv space E and $f : U \to F$ a map, where F is a normable rtvs. Then for all natural number p, the following statements are equivalent:*

(a) *f is a C^p-map* (**3.3**).

(b) *For every element x of U, there are an open neighbourhood V^x of x in E and a C^p-map $f^* : V^x \to F$* (**3.3**) *(in the sense of ordinary differential calculus, which in this case coincides with* **3.3**) *such that $f^*|_{V^x \cap U} = f|_{V^x \cap U}$. (See* [21]*)*

Remark If at every point of U there is an open neighbourhood in U on which all the derivatives of f are bounded by the same constant K, then the preceding proposition is also true for C^∞-maps (**3.3**).

Strong differentiability

See mainly [2] and [13].

3.11 A subset A of a topological space X is said to be admissible if $A \subset \overline{int(A)}$.

Lemma *Let A be an admissible subset of an admissible subset B of a topological space X. Then A is admissible in X. In particular $G \cap B$, where G is open in X, is admissible in X.*

Let Q be a quadrant of a rtvs E with $int(Q) \neq \varnothing$. Then Q is admissible in E and every open subset U of Q is admissible in E ($U \subset \overline{int(U)}$).

Indeed, recall that if $(\{e_i | i \in I\}, K \subset I)$ is a pair adapted to Q, $index(Q) = n$ and $I - K = \{i_1, ..., i_n\}$, then the map $+ : Q^0 \times L\{e_{i_1}, ..., e_{i_n}\} \to E$ is a linear homeomorphism. Consequently $\overline{int(Q)} = Q$.

Let A be an admissible subset of a topological space X.Then $A - int(A)$(=∂A) will be called boundary of A. The elements of ∂A will be called boundary points of A.

When U is an open subset of a quadrant Q of a rtvs E, with $int(Q) \neq \varnothing$, we have seen that U is admissible in E and we have defined $\partial U = U - int(U)$. But in **2.15** we have also defined $\partial U = \{x \in U | \, index(x) \geqslant 1\}$ (which becomes closed in U). Of course these two definitions coincide.

Definition Let a be an element of an admissible subset A of a normable (with norm $\| \, \|$) rtvs E and $f : A \to F$ a map, where F is a Hlcrtvs defined by the collection of seminorms $\{\| \, \|_i \, | i \in I\}$. We say that f is strongly differentiable at a if there is $u \in \mathcal{L}(E, F)$ such that $\lim_{(x,y)\to(a,a)} (f(y) - f(x) - u(y - x)) \|y - x\|^{-1} = 0 \in F$ (which limit is equivalent to:

$$\lim_{(x,y)\to(a,a)} \|f(y) - f(x) - u(y - x)\|_i \, \|y - x\|^{-1} = 0 \in \mathbb{R}, \text{ for all } i \in I).$$

If f is strongly differentiable at every point of A, f is said to be strongly differentiable on A.

If f is strongly differentiable at x, then the continuous linear map u, that fulfils the above limit, is unique and will be denoted by $df(a)$, and will be called the strong differential of f at a.

Note The preceding notions are independent of the equivalent norms considered in E.

3.12

(1) Let us consider E, A, F, f, a as in the preceding definition. Then:

(a) If f is strongly differentiable at a, then f is continuous at a.

(b) If F is normable and f is strongly differentiable at a, then for all real number $c > \|df(a)\|$, there exists $\delta > 0$ such that $\|f(y) - f(x)\| \leqslant c\,\|y - x\|$ for all $x, y \in B_\delta^{\|\,\|}(a) \cap A$ (which implies that f is uniformly continuous on $B_\delta^{\|\,\|}(a) \cap A$).

(c) Suppose that f is strongly differentiable on A. Then the map $df : A \to \mathcal{L}(E, F)$, $y \mapsto df(y)$, is continuous, where the topology on $\mathcal{L}(E, F)$ is the locally convex topology defined in **3.3**.

(2) Let us consider E, Q, U, F, f, $a \in U$ as in **3.1** (*iii*). Suppose that $int(Q) \neq \varnothing$ (so U is admissible in E). Then if f is strongly differentiable at a, f is differentiable at a (**3.1** (*iii*)) and $Df(a) = df(a)$. The converse of this result is not true, in general.

Moreover, if f is strongly differentiable on U, then f is a C^1-map on U (**3.3** (2)).

3.13 Mean value theorems

By the Hahn-Banach Theorem and the classical Mean value Theorem ([5], p. 153) one has the following:

Lemma *Let F be a Hlcrtvs, $c : [a, b] \to F$ a continuous map, where $a, b \in \mathbb{R}$ and $a < b$, let D be a countable subset of $[a, b]$ and $h : [a, b] \to \mathbb{R}$ a continuous monotone function. Suppose that c and h are differentiable maps at every $x \in [a, b] - D$. Let C be a convex closed subset of F such that $c'(t) \in h'(t) \cdot C$ for all $t \in [a, b] - D$. Then $c(b) - c(a) \in (h(b) - h(a)) \cdot C$ (see* [18], *p.10).*

Let us consider E, Q, U, F, f as in **3.3.** Let x, y be elements of U such that $[x, y] \subset U$ ($[x, y] = \{(1 - t)x + ty | 0 \leqslant t \leqslant 1\}$). As a consequence of the preceding Lemma we have the following properties:

(1) Suppose that f is differentiable at every $z \in [x, y]$ (**3.1**). Then:

(a) Let C be a closed convex subset of F such that $Df(z)(y - x) \in C$ for all $z \in [x, y]$. Then $f(y) - f(x) \in C$. In particular $f(y) - f(x)$ belongs to the closed convex hull of $\{Df(z)(y - x) | z \in [x, y]\}$.

(b) Let i be an element of I such that there is $M > 0$ with $\|Df(z)\|_i^* \leqslant M$ for all $z \in [x, y]$ (**3.3**). Then: $\|f(y) - f(x)\|_i \leqslant M\,\|y - x\|$.

(2) Suppose that f is differentiable on U (**3.1** (*iii*)) and let x_0 be a point of U. Suppose that for $i \in I$ there there is $\rho > 0$ such that $\|Df(z) - Df(x_0)\|_i^* \leqslant \rho$ for all $z \in [x, y]$. Then $\|f(y) - f(x) - Df(x_0)(y - x)\|_i \leqslant \rho\,\|y - x\|$.

This property (2) permits us to prove that if f is a C^1-map on U (**3.3** (2)) and $int(Q) \neq \varnothing$, then f is strongly differentiable on U (**3.11**).Thus, by **3.12** (2), we have:

Proposition *Let us consider E, Q, U,F, f as in **3.3**, and suppose that $int(Q) \neq \varnothing$. Then f is a C^1-map on U if and only if f is strongly differentiable on U (in this case $Df(a) = df(a)$ for all $a \in U$).*

3.14 Let us consider E, A, F, f as in **3.11** (Definition of strong differentiability). Then we give the following definitions:

(1) f is a C^1_S-map on A if f is strongly differentiable on A (in this case we have $df(a)$ for all $a \in A$ and $df : A \to \mathcal{L}(E, F)$ is continuous).

(2) f is a C^2_S-map on A if f is a C^1_S-map on A and $df : A \to \mathcal{L}(E, F)$ is a C^1_S-map on A (which is equivalent to: f and df are strongly differentiable on A). In this case $d(df) : A \to \mathcal{L}(E, \mathcal{L}(E, F))$ is continuous and we put $\theta \circ d(df) = d^2 f : A \to \mathcal{L}^2(E, F)$ (see **3.3** (3)).

(3) Let r be a natural number with $r \geqslant 1$. Then f is a C^r_S-map on A if f, df, $d^2 f, \ldots, d^{r-1} f$ are strongly differentiable on A. In this case $d^r f : A \to \mathcal{L}^r(E, F)$ is continuous (inductive process).

(4) f is a C^∞_S-map on A, if f is a C^k_S-map on A for each positive integer k.

3.15

(1) Let us consider $E,\ Q,\ U,\ F,\ f$ as in **3.3** and suppose that $int(Q) \neq \varnothing$. Then:

   (*a*) f is C^1 on U if and only if f is C^1_S on U (in this case $Df(a) = df(a)$ for all point a of U). Moreover, by induction, f is C^k on U if and only if f is C^k_S on U ($k \geqslant 1$). Finally if f is C^k on U, then $D^k f = d^k f$ on U.

   (*b*) Suppose that f is C^k_S on U. Then $f|_{int(U)} : int(U) \to F$ is C^k_S on $int(U)$ and therefore C^k on $int(U)$.

(2) Let us consider E, A, F, f as in **3.11.** Then:

   (*a*) If f is C^k_S on A, we have that $f|_{int(A)}$ is C^k_S on $int(A)$ and therefore C^k on $int(A)$.

   (*b*) If f is C^r_S on A, then for all $p \in \mathbb{N} \cup \{0\}$ with $p \leqslant r$, f is C^p_S on A and $d^p f$ is C^{r-p}_S on A.

   (*c*) If $\lambda \in \mathcal{L}(E, F)$ and $f = \lambda|_A$, then f is a C^∞_S-map on A and $df(x) = \lambda$, $d^r f(x) = 0$ for all point x of A and all natural number $r > 1$ (**3.14**).

3.16 In order to build manifolds with generalized boundary modeled on normable spaces, we need essentially the four properties of C^r_S-maps that follow (see **3.4**):

(*i*) **Restrictions to open sets** If $f : A \to F$ is a C^r_S-map on A (**3.14**) and V is an open subset of A, then $f|_V : V \to F$ is a C^r_S-map on V and $d(f|_V)(x) = df(x)$ for all point x of V (note that V is admissible in E).

(*ii*) **Open covering property** Let E, A, F, f be as in **3.14,** and $\{U_j | j \in J\}$ an open covering of A. Then f is C^r_S on A if and only if $f|_{U_j}$ is C^r_S on U_j for all $j \in J$.

(*iii*) **Chain rule** Let E_1, A_1, F_1, f_1 be as in **3.14** and let E_2, A_2, F_2, f_2 be as in **3.14**. Suppose that $F_1 = E_2$ and $f_1(A_1) \subset A_2$. Then:

(*a*) If f_1 is strongly differentiable at $x_0 \in A_1$ and f_2 is strongly differentiable at $f_1(x_0)$, then $f_2 \circ f_1$ is strongly differentiable at x_0 and $d(f_2 \circ f_1)(x_0) = df_2(f_1(x_0)) \circ df_1(x_0)$.

(*b*) If f_1 and f_2 are C^r_S ($r \in \mathbb{N} \cup \{\infty\}$) on A_1 and A_2, respectively, then $f_2 \circ f_1$ is C^r_S on A_1(**3.14**).

(*iv*) **Restrictions to vector subspaces** Let E, A, F, f as in **3.14.** Let G be a vector subspace of E, and suppose that $A \cap G$ is admissible in G and $f : A \to F$ is C^r_S on A (**3.14**). Then $f|_{A\cap G} : A \cap G \to F$ is C^r_S on $A \cap G$ and $d(f|_{A\cap G})(x) = df(x)|_G$ for all point x of $A \cap G$.

((*i*) and (*iv*) follow from (*iii*)).

Note In the study of properties of manifolds with generalized boundary is also useful the following result: if $f : A \to F$ is a map (**3.11**) and F_1 is a closed vector subspace of F such that $f(A) \subset F_1$, then $f : A \to F_1$ is a C^r_S-map if and only if $f : A \to F$ is a C^r_S-map (**3.14**). Of course $\mathcal{L}(E, F_1)$ is a closed vector subspace of $\mathcal{L}(E, F)$.

3.17 Let E, A, F, f be as in **3.11**, and B an admissible subset in F. Suppose that $f(A) \subset B$ and that F is normable. We say that $f : A \to B$ is a C^r_S-diffeomorphism if $f : A \to B$ is a bijective map and f, f^{-1} are C^r_S-maps (**3.14** (3)). In this case f is a homeomorphism.

It is clear that if $f : A \to B$ is a C^r_S-diffeomorphism, then $f^{-1} : B \to A$ is also a C^r_S-diffeomorphism. Moreover, the composition of C^r_S-diffeomorphisms is a C^r_S-diffeomorphism (apply the chain rule), and the identity map is a C^r_S-diffeomorphism.

Finally, by the chain rule (**3.16**), one has that if $f : A \to B$ is a C^r_S-diffeomorphism, then $df(x) : E \to F$ is a linear homeomorphism for all $x \in A$.

3.18 Invariance of the boundary for C^r_S-diffeomorphisms

We recall (**3.7**) that in Banach spaces we have the "nverse mapping theorem". This theorem implies the following important result: *Let $f : A \to B$ be a C^r_S-diffeomorphism* (**3.17**)*, and suppose that E, F are banachable spaces. Then $f(\partial A) = \partial B$.* (**3.11** and **3.15** (1) (*b*)).

This theorem allows us to define manifolds with generalized boundary, modeled on banachable spaces.

Differentiation theory in locally convex spaces

We generalize some ideas contained in [11], [14] and [24], and construct a differential calculus on open subsets of quadrants in Hlcrtvs. See also the article by József Szilasi and Rezső L. Lovas.

3.19 Let U be an open subset of a quadrant Q of a Hlcrtvs E, F a Hlcrtvs and $f : U \to F$ a map. Then f is said to be a weakly C^1-map on U (or C^1_W-map on U) if:

(*i*) $f : U \to F$ is a continuous map.

(*ii*) For all element x of U, we have:

(1) For all $v \in B(Q, U, x)$ (**2.14**),there exists $\lim\limits_{t\to 0} \frac{f(x+tv)-f(x)}{t}$ (=$d_v f(x)$).

(2) For all $v \in B^+(Q,U,x)$ (**2.14**),there exists $\lim\limits_{t\to 0^+} \frac{f(x+tv)-f(x)}{t}$(=$d_v f(x)$ or $d_v^+ f(x)$).

(3) For all $v \in B^-(Q,U,x)$ (**2.14**),there exists $\lim\limits_{t\to 0^-} \frac{f(x+tv)-f(x)}{t}$(=$d_v f(x)$ or $d_v^- f(x)$).

(*iii*) If $A_U = \{(x,v) \in U \times E | v \in B(Q,U,x) \cup B^+(Q,U,x) \cup B^-(Q,U,x)\}$, then the map $d.f(\cdot) : A_U \to F,\ (x,v) \longmapsto d_v f(x)$, is continuous.

Remarks (*a*) If x is an element of U and v a vector of E, then $-v \in B^+(Q,U,x)$ if and only if $v \in B^-(Q,U,x)$. Consequently (*ii*) (2) implies (*ii*) (3).

(*b*) If E and F are Fréchet spaces and $Q = E$, then $A_U = U \times E$ and the preceding definition is the one used in [14].

(*c*) If $Q = E$, then $A_U = U \times E$ and the preceding definition is the one used in [11].

(*d*) Consider $x \in U$, $v \in B(Q,U,x)$ and $u \in Q$. Then there exists $\delta > 0$ such that for all $t \in [0,\delta)$, $x+tu \in U$ and $v \in B(Q,U,x+tu)$ (consequently $(x+tu,v) \in A_U$).

(*e*) Consider $x \in U$, $v \in B^+(Q,U,x)$ and $u \in Q$. Then there exists $\delta > 0$ such that for all $t,s \in [0,\delta)$, $x+tu+sv \in U$ and $x+sv \notin Q$ for all $s < 0$ (then $v \notin -Q$). Suppose, moreover, that $v \in Q$ and $coindex(u+x) < coindex(v)$. Then for all $t \in [0,\delta)$, $v \in B^+(Q,U,x+tu)$.

(*f*) Consider $x \in U$, $v \in B^-(Q,U,x)$ and $u \in Q$. Then there exists $\delta > 0$ such that for all $(t,s) \in [0,\delta) \times (-\delta,0]$, $x+tu+sv \in U$ and $x+tv \notin Q$ for all $t > 0$ (then $v \notin Q$). Suppose, moreover, that $v \in -Q$ and $coindex(u+x) < coindex(-v)$. Then for all $t \in [0,\delta)$, $v \in B^-(Q,U,x+tu)$.

(*g*) $int(U) \times E \subset A_U$ and $U \times (Q \cup (-Q)) \subset A_U$ (**2.14**).

3.20 Suppose that $f : U \to F$ is a C^1_W-map on U (**3.19**). Then:

(1) For all $(x,v) \in A_U$ and all real number r, one has $(x,rv) \in A_U$ and $rd_v f(x) = d_{rv} f(x)$.

(2) If $x \in int(U)$ and $u,v \in E$, then, applying **3.8**, one deduces that $d_{u+v} f(x) = d_u f(x) + d_v f(x)$ (see (g) of the preceding remark).

(3) If $x \in U$ and $u,v \in Q$ (or $u,v \in -Q$), then $d_{u+v} f(x) = d_u f(x) + d_v f(x)$.

Theorem *Let $f : U \to F$ be a C^1_W-map on U (**3.19**). Suppose that $int(Q) \neq \varnothing$. Then there exists a unique continuous map $df : U \times E \to F$ such that $df|_{A_U} = d.f(\cdot) : A_U \to F$ and for all $x \in U$, $df(x,\cdot) : E \to F$, $u \mapsto df(x,u)$, is a linear continuous map (called the (Gâteaux) differential of f at x).*

Indeed,by **2.12** (*ii*) there exists a linearly independent system $\Lambda = \{\lambda_1, ..., \lambda_n\}$ of elements of $\mathcal{L}(E,\mathbb{R})$ such that $Q = E^+_\Lambda$ and $Q^0 = E^0_\Lambda$ ($n = Q$)). By **2.7** (*b*), there exists a finite subset $\{x_1, ..., x_n\}$ of E such that $\lambda_i(x_j) = \delta_{ij}$. Consequently the map $\alpha : E^0_\Lambda \times \mathbb{R}^n \to E$, $(u, r_1, ..., r_n) \mapsto u + r_1 x_1 + ... + r_n x_n$, is a linear homeomorphism and

$\alpha(E^0_\Lambda \times (\mathbb{R}^n)^+_{\{p_1,\dots,p_n\}}) = E^+_\Lambda = Q$. Then one defines $df(x, u_0 + r_1 x_1 + \dots + r_n x_n) = d_{u_0} f(x) + r_1 d_{x_1} f(x) + \dots + r_n d_{x_n} f(x)$, where $u_0 \in Q^0,\ r_1, \dots, r_n \in \mathbb{R}$. To end the proof use the preceding (1), (2) and (3).

3.21 Let E, F be Hlcrtv spaces and let U be an admissible subset of E (**3.11**). C. Wockel in [28] says that a map $f : U \to F$ is a C^1-map on U if:

(*i*) f is a continuous map.

(*ii*) $f|_{int(U)} : int(U) \to F$ is a C^1_W-map on $int(U)$ (**3.19**) (i.e., for all $x \in int(U)$ and all $v \in B(E, int(U), x) = E$ there exists $\lim_{t\to 0} \frac{f(x+tv)-f(x)}{t} = d_v(f|_{int(U)})(x)$ and $d.(f|_{int(U)})(\cdot) : A_{int(U)} = int(U) \times E \to F$ is continuous).

(*iii*) There exists a continuous map $df : U \times E \to F$ (called the differential (Wockel) of f) such that $df|_{int(U)\times E} = d.(f|_{int(U)})(\cdot)$.

Remarks (*a*) Of course $d.(f|_{int(U)})(\cdot) = d(f|_{int(U)})$ (Theorem of **3.20**).

(*b*) $int(U) \times E$ is admissible in $E \times E$.

(*c*) For all $x \in int(U)$, $d.(f|_{int(U)})(x) : E \to F$ is a continuous linear map, and therefore $df(x, \cdot) : E \to F$ is also a continuous linear map for all $x \in U$ (use a net on $int(U)$ which converges to x).

(*d*) From the definition, df is unique.

Theorem *Let U be an open subset of a quadrant Q of a Hlcrtvs E with $int(U) \neq \varnothing$, and $f : U \to F$ a map, where F is a Hlcrtvs. Then f is a C^1_W-map on U* (**3.19**) *if and only if f is a C^1-map on U according to the above definition given by Wockel.*

By **3.19** and the Theorem of **3.20,** we have the step "Definition **3.19** implies Definition **3.21**". For the converse step use **2.12** (*iv*) (*b*).

Corollary *(a) Let E be a Hlcrtvs and let $c : [0, \delta) \to E$ be a continuous curve such that $c|_{(0,\delta)} : (0, \delta) \to E$ is a C^1-map. Assume that the derivative $c' : (0, \delta) \to E$ has a continuous extension to $[0, \delta)$. Then c is differentiable at* 0 (**3.1** (*iii*)) *and $c'(0) = \lim_{t\to 0^+} c'(t)$ ($c'(0) = Dc(0)(1)$).*

(b) Let E be a Hlcrtvs and let $c : \mathbb{R} \to E$ be a continuous curve which is a C^1-map on $\mathbb{R} - \{0\}$. Assume that the derivative $c' : \mathbb{R} - \{0\} \to E$ has a continuous extension to $\mathbb{R}$. Then c is differentiable at 0 (**3.1** (*iii*)) *and $c'(0) = \lim_{t\to 0} c'(t)$.*

Use the Lemma of **3.13.**

3.22 The following results are necessary to study the relations between C^r_W-maps and C^r-maps.

Proposition (The evaluation map) *Let E be a normable rtvs and let F be a Hlcrtvs defined by a collection of seminorms $\{\| \|_i \,|i \in I\}$. Then the evaluation map $e : E \times \mathcal{L}(E, F) \to F$, $(x, \lambda) \mapsto \lambda(x)$, is continuous, where in $\mathcal{L}(E, F)$ one considers the topology of its structure of Hlcrtvs defined in* **3.3**.

In the preceding result, the condition "E normable"" cannot be changed by "E Hausdorff locally convex". Indeed: Let E be a Hlcrtvs. Consider $\mathcal{L}(E,\mathbb{R})$ equipped with a topology of a topological vector space (for example, the locally convex topology of the uniform convergence [16]), and let $e : E \times \mathcal{L}(E,\mathbb{R}) \to \mathbb{R}$ be the evaluation map. Assume that e is continuous. Then there are a convex neighbourhood V^0 of 0 in E and a neighbourhood U^0 of 0 in $\mathcal{L}(E,\mathbb{R})$ such that $e(V^0 \times U^0) \subset [-1,1]$. Hence V^0 is bounded in E (Recall the general result: *Let E be a Hlcrtvs and let A be a subset of E. Then A is bounded in E if and only if $\lambda(A)$ is bounded for every $\lambda \in \mathcal{L}(E,\mathbb{R})$* [16].) On the other hand, for all $\lambda \in \mathcal{L}(E,\mathbb{R})$ there is a natural number n such that $\frac{1}{n}\cdot\lambda \in U^0$. Consequently E is normable (apply the general result: *Let E be a tvs. Then:* (i) *E is seminormable if and only if there exists a bounded convex neighbourhood of 0, and* (ii) *E is normable if and only if it is seminormable and Hausdorff.*)

Corollary *Let X be a topological space, E a normable rtvs, F a Hlcrtvs and $\alpha : X \to \mathcal{L}(E,F)$ a continuous map (with the topology of the structure of Hlcrtvs of $\mathcal{L}(E,F)$ defined in* **3.3***). Then $\overset{\smile}{\alpha} : X \times E \to F$, $(x,y) \mapsto \alpha(x)(y)$, is a continuous map.*

As a consequence we have: *If E, Q, U, F, f are as in* **3.3** *and f is a C^1-map on U, then f is a C^1_W-map on U, and $Df(x)(v) = d_v f(x)$ for all $(x,v) \in A_U$ (analogously for C^r-maps, whenever that $int(Q) \neq \varnothing$).*

One has the following converses of the preceding Corollary:

(1) Let X be a topological space, F a Hlcrtvs and $\beta : X \times \mathbb{R}^n \to F$ a continuous map such that $\beta(x,\cdot) : \mathbb{R}^n \to F$, $y \mapsto \beta(x,y)$ is a linear map for all point x of X. Then $\overset{\wedge}{\beta} : X \to \mathcal{L}(\mathbb{R}^n,F)$, $x \mapsto \beta(x,\cdot)$, is a continuous map.

Note that the map $\psi : \mathcal{L}(\mathbb{R},F) \times ... \times \mathcal{L}(\mathbb{R},F) \to \mathcal{L}(\mathbb{R}^n,F)$, $(\lambda_1,...\lambda_n) \mapsto \lambda_1 p_1 + ... + \lambda_n p_n$, is a linear homeomorphism. It is clear that $\|\lambda_j p_j\|_i^* = \|\lambda_j\|_i^*$, $j = 1,...,n, i \in I$, and $\psi^{-1}(\mu) = (\mu \circ j_1,...,\mu \circ j_n)$. Finally the family of seminorms $\| \|_i^* \cdot \bar{p}_j$,$i \in I$, $j = 1,...,n$, describes the topology of $\mathcal{L}(\mathbb{R},F) \times ... \times \mathcal{L}(\mathbb{R},F)$,[17].

As a consequence of this statement and the Theorem of **3.20,** we have: *If $f : U \to F$ is a C^1_W-map on U and $E = \mathbb{R}^n$, then f is a C^1-map on U* (**3.3**).

(2) Let U be an open subset of a quadrant Q, with $int(Q) \neq \varnothing$, of a normable rtvs E, G a normable rtvs, F a Hlcrtvs and $\beta : U \times G \to F$ a C^p-map ($p \in \mathbb{N} \cup \{\infty\}$) such that $\beta(x,\cdot) : G \to F$ is a linear continuous map for all point x of U. Then the map $\overset{\wedge}{\beta} : U \to \mathcal{L}(G,F)$ is of class C^{p-1}(of course if $p = \infty$, $p-1 = \infty$) (**3.3**).

It is sufficient to remark that $D_2\beta : U \times G \to \mathcal{L}(G,F)$ is a C^{p-1}-map and $\overset{\wedge}{\beta}(x) = D_2\beta(x,y)$ for all $(x,y) \in U \times G$.

Theorem *Let E, Q, U, F, f be as in* **3.19***, and suppose that $E = \mathbb{R}^n$ (hence, $int(Q) \neq \varnothing$). Then the following statements are equivalent:*

(i) f is strongly differentiable on U (**3.11**).

(ii) f is a C^1-map on U (**3.3**).

(iii) f is a C^1_W-map on U (**3.19**).

The equivalence between (*i*) and (*ii*) has been established in the Proposition of **3.13.** The proof of the step "(*iii*) implies (*ii*)" follows from the preceding (1) and Theorem of **3.20**. Finally, the proof of the step "(*ii*) implies (*iii*)" follows from the consequence of the preceding Corollary.

3.23

(1) Let $f : U \to F$ be a C^1_W-map on U (**3.19**), and let V be an open subset of U (hence an open subset of Q). Then $f|_V : V \to F$ is a C^1_W-map on V, $A_V \subset A_U$ and $d.f(\cdot)|_{A_V} = d.(f|_V)(\cdot)$. Moreover if $int(Q) \neq \varnothing$, then $df|_{V\times E} = d(f|_V)$ (Theorem of **3.20**).

(2) Let E, Q, U, F, f be as in **3.19.** Let $\{V_j\}_{j\in J}$ be an open covering of U. Then $f : U \to F$ is a C^1_W-map on U if and only if $f|_{V_j} : V_j \to F$ is a C^1_W-map on V_j for all $j \in J$. In this case $d.f(\cdot)|_{A_{V_j}} = d.(f|_{V_j})(\cdot)$, for all $j \in J$.

(3) Let E, Q, U, F, f be as in **3.19** and let E_1, Q_1, U_1, F_1, f_1 be as in **3.19.** Suppose that $F = E_1$, $f(U) \subset U_1$, f is a C^1_W-map on U, f_1 is a C^1_W-map on U_1 and $int(Q_1) \neq \varnothing$. Then $f_1 \circ f : U \to F_1$ is a C^1_W-map on U and $d_v(f_1 \circ f)(x) = df_1(f(x), d_v f(x))$ for all $(x, v) \in A_U \subset U \times E$ (Theorem of **3.20**).

Finally if $int(Q) \neq \varnothing$, then $d(f_1 \circ f)(x, v) = df_1(f(x), df(x, v))$ for all $(x, v) \in U \times E$.

(4) Let $f : U \to F$ be a C^1_W-map on U (**3.19**), where $int(U) \neq \varnothing$, and let H be a linear subspace of E (therefore H is a Hlcrtvs). Suppose that $U \cap H$ is an open subset of a quadrant Q_H of H, quadrant that has finite index and closed kernel**.** Then $f|_{U\cap H} : U \cap H \to F$ is a C^1_W-map on $U \cap H$ and $df(x, \cdot)|_H = d(f|_{U\cap H})(x, \cdot) : H \to F$, for all $x \in U \cap H$.

3.24 Invariance of the boundary for C^1_W-diffeomorphisms

Lemma *Let U be an open subset of a Hlcrtvs E, F a Hlcrtvs, $\lambda \in \mathcal{L}(F, \mathbb{R})$ with $\lambda \neq 0$, $f : U \to F$ a C^1_W-map on U with $f(U) \subset F^+_\lambda$ and $x \in U$ (***3.19***). Suppose that $f(x) \in F^0_\lambda$. Then $df(x, u) \in F^0_\lambda$ for all element u of E (***3.20***), (see ***3.6***).*

Definition Let E, Q, U, F, f be as **3.19**. Let V be an open subset of a quadrant P of F and suppose that $f(U) \subset V$. We say that $f : U \to V$ is a C^1_W-diffeomorphism if $f : U \to V$ is a bijective map and $f : U \to V, f^{-1} : V \to U$ are C^1_W-maps (**3.19**). In this case f is a homeomorphism.

If $f : U \to V$ is a C^1_W-diffeomorphism and $int(Q) \neq \varnothing$ and $int(P) \neq \varnothing$, then $df(x, \cdot) : E \to F$ is a linear homeomorphismo for all $x \in U$, (**3.20**), and $df(x, \cdot)$, $df^{-1}(f(x), \cdot)$ are inverse maps.

Theorem *Let $f : U \to V$ be a C^1_W-diffeomorphism and suppose that $int(Q) \neq \varnothing$ and $int(P) \neq \varnothing$. Then we have:*

*(a) For all $x \in U \subset Q$, $index(x) = index(f(x))$ (**2.9** (c), **2.15**).*

(b) $\partial U \neq \varnothing$ if and only if $\partial V \neq \varnothing$. Moreover, $f(\partial U) = \partial V$.

(c) *$Int(U) \neq \varnothing$ if and only if $int(V) \neq \varnothing$. Moreover, $f(int(U)) = int(V)$.*

(d) *$f|_{int(U)} : int(U) \to int(V)$ is a C^1_W-diffeomorphism. Moreover:*

$$df|_{int(U)\times E} = d(f|_{int(U)}) : int(U) \times E \to F(\mathbf{3.20}).$$

(e) *$f(\partial^k U) = \partial^k V,\ f(B_k U) = B_k V,\ 0 \leqslant k \leqslant index(Q)$.*

Remark Let E, Q, U, f, F as in **3.19**. Suppose that U is open in E. Then the definition of "f is a C^1_W-map on U" does not depend of the quadrant Q in E that contains U.

3.25

(1) Let $f : U \to F$ be a C^1_W-map on U (**3.19**) and suppose that $int(Q) \neq \varnothing$. Then we have the map $df : U \times E \to F$ (**3.20**). On the other hand $E \times E$ is a Hlcrtvs, $Q \times E$ is a quadrant of $E \times E$ with $int(Q \times E) \neq \varnothing$ and $U \times E$ is an open subset of $Q \times E$. Suppose that df is a C^1_W-map on $U \times E$. Then f is said to be a C^2_W-map on U. In this case we have $d(df)(= d^2 f) : U \times E \times E \times E \to F$ (**3.20**). Recall that $d^2 f|_{A_{U\times E}} = d.(df)(\cdot)$ and $d^2 f(x, v, \cdot, \cdot) : E \times E \to F$ is a linear continuous map for all $(x, v) \in U \times E$ and $A_{U\times E} = B_2 \times E$, where $B_2 = \{(x, v, u) \in U \times E \times E | (x, u) \in A_U\}$.

(2) Let $f : U \to F$ be a C^2_W-map on U and suppose that $d^2 f : U \times E \times E \times E \to F$ is a C^1_W-map on $U \times E \times E \times E$ (**3.19**) (Note that $Q \times E^3$ is a quadrant of E^4 and $int(Q \times E^3) \neq \varnothing$). Then we say that $f : U \to F$ is a C^3_W-map on U. In this case we have $d(d^2 f)(= d^3 f) : U \times E \times E \times E \times E^4 \to F$ (**3.20**). Recall that $d^3 f(x, u_1, u_2, u_3, \cdot, \cdot, \cdot, \cdot) : E^4 \to F$ is a linear continuous map for all $(x, u_1, u_2, u_3) \in U \times E \times E \times E$, $(d^3 f : U \times E^{2^3-1} \to F)$.

(3) Let U be an open subset of a quadrant Q, with $int(Q) \neq \varnothing$, of a Hlcrtvs E, and F a Hlcrtvs. Then $U \times E^{2^k-1}$, k a natural number, is an open subset of $Q \times E^{2^k-1}$, which is a quadrant of E^{2^k} with finite index and closed kernel (**2.10**), (in fact, $index(Q \times E^{2^k-1}) = index(Q)$). Now if $\varphi : U \times E^{2^k-1} \to F$ is a C^1_W-map, then $d\varphi : U \times E^{2^{k+1}-1} \to F$ (**3.20**).

(4) Inductively we have the following definition:

Let $f : U \to F$ be a C^1_W-map on U (**3.19**) and suppose that $int(Q) \neq \varnothing$. We say that f is a C^{k+1}_W-map on U, k a natural number, if:

$df(= d^1 f) : U \times E \to F$ is a C^1_W-map on $U \times E$, $d(df) = d^2 f : U \times E^{2^2-1} \to F$ is a C^1_W-map on $U \times E^{2^2-1}$, $d(d^2 f) = d^3 f : U \times E^{2^3-1} \to F$ is a C^1_W-map on $U \times E^{2^3-1}$,..., $d(d^{k-1} f) = d^k f : U \times E^{2^k-1} \to F$ is a C^1_W-map on $U \times E^{2^k-1}$. In this case we have the extension $d(d^k f)(= d^{k+1} f) : U \times E^{2^{k+1}-1} \to F$ (**3.20**).

(5) Let $f : U \to F$ be as the preceding (4). We say that f is a C^∞_W-map on U if f is a C^k_W-map on U for all $k \in \mathbb{N}$.

Note that if $f = \lambda|_U$, where λ is a continuous linear map of E into F, then $f : U \to F$ is a C^∞_W-map on U.

3.26 The results **3.23** (1),(restrictions to open sets), **3.23**(2), (open covering property), **3.23** (3), (chain rule), and **3.23** (4), (restrictions to vector subspaces), remain true for C^r_W-maps, $r \in \mathbb{N} \cup \{\infty\}, r > 1$.

The definition of **3.24** is obviously generalized to obtain the notion of C^r_W-diffeomorphism, $r > 1$. It is clear that if $f : U \to V$ is a C^r_W-diffeomorphism, then $f^{-1} : V \to U$ is a C^r_W-diffeomorphism. On the other hand the composition of C^r_W-diffeomorphisms is a C^r_W-diffeomorphism, and the identity map is a C^r_W-diffeomorphism.

If in Theorem of **3.24** we put "$f : U \to V$ is a C^r_W-diffeomorphism" instead of "$f : U \to V$ is a C^1_W-diffeomorphism", then $f|_{int(U)}$ of (d) is a C^r_W-diffeomorphism.

Remark Let Q and Q' be quadrants, with non-void interiors, of a Hlcrtvs E. Then Q and Q' are C^∞_W-diffeomorphic if and only if $index(Q) = index(Q')$ (**2.13**).

These results are the essential tools to construct manifolds with corners modeled on Hlcrtv spaces.

3.27 Relation between C^r_W-maps and C^r-maps

Let E, Q, U, f, F be as in **3.3** and suppose that $int(Q) \neq \varnothing$. Then:

(*i*) If f is a C^r-map on U (**3.3**), then f is a C^r_W-map on U, for all $r \in \mathbb{N}$ (Use Corollary of **3.22** and **3.28** below).

(*ii*) If f is a C^{r+1}_W-map on U, then f is a C^r-map on U (**3.3**), for all $r \in \mathbb{N}$ (Use **3.22** (2) and **3.28** below).

(*iii*) f is a C^∞-map on U (**3.3**) if and only if f is a C^∞_W-map on U (Use **3.22** (2)).

3.28 An alternative definition of C^r_W-maps

(*i*) Let $f : U \to F$ be a C^1_W-map on U (**3.19**), with $int(Q) \neq \varnothing$. Let us consider the extension $df : U \times E \to F$ (Theorem of **3.20**). Suppose that:

(1) For all vector v of E, $df(\cdot, v) : U \to F$ is a C^1_W-map on U (**3.19**), (then we have the extension $d(df(\cdot, v)) : U \times E \to F$).

(2) The map $\{(x, v, u) \in U \times E \times E | (x, u) \in A_U\}(= B) \to F$, $(x, v, u) \mapsto d_u(df(\cdot, v))(x) = d(df(\cdot, v))(x, u)$, is continuous.

Then f is said to be $\widetilde{C}^2_W$ -map on U, and in this case we define $\widetilde{d}^2 f(x, v, u) = d(df(\cdot, v))(x, u)$, for all $(x, v, u) \in U \times E \times E$.

Then for all $(x, v, u) \in B$, $\widetilde{d}^2 f(x, v, u) = d_u(df(\cdot, v))(x)$; $\widetilde{d}^2 f : U \times E \times E \to F$ is continuous; $\widetilde{d}^2 f(x, v, \cdot) : E \to F$ is a linear map for all $(x, v) \in U \times E$, and $\widetilde{d}^2 f$ is unique.

(*ii*) Let $f : U \to F$ be a $\widetilde{C}^2_W$-map on U. Suppose that:

(1) For all $(v, u) \in E \times E$, $\widetilde{d}^2 f(\cdot, v, u) : U \to F$ is a C^1_W-map on U (**3.19**).

(2) The map $\{(x, v, u, w) \in U \times E \times E \times E | (x, w) \in A_U\}(= B_3) \to F$, $(x, v, u, w) \mapsto d_w(\widetilde{d}^2 f(\cdot, v, u))(x)$, is continuous.

Then we say that f is a $\widetilde{C}^3_W$-map on U. In this case for all $(x, v, u, w) \in B_3$ one has $d(\,\widetilde{d}^2 f(\cdot, v, u))(x, w) = d_w(\,\widetilde{d}^2 f(\cdot, v, u))(x)$,and we define

$\widetilde{d}^3 f(x, v, u, w) = d(\,\widetilde{d}^2 f(\cdot, v, u))(x, w) \in F$ for all $(x, v, u, w) \in U \times E \times E \times E$. Then for all $(x, v, u, w) \in B_3$, $\widetilde{d}^3 f(x, v, u, w) = d_w(\widetilde{d}^2 f(\cdot, v, u))(x)$; $\widetilde{d}^3 f : U \times E \times E \times E \to F$ is continuous; $\widetilde{d}^3 f(x, v, u, \cdot) : E \to F$ is a linear map for all $(x, v, u) \in U \times E \times E$, and $\widetilde{d}^3 f$ is unique.

Inductively we have the following definition:

(*iii*) Let $f : U \to F$ be a C^1_W-map on U (**3.19**), with $int(Q) \neq \varnothing$. The map $f : U \to F$ is said to be a $\widetilde{C}^k_W$-map on U ($k \in \mathbb{N}$, $k > 1$) if:

$df(\cdot, v) : U \to F$ is a C^1_W-map on U for all $v \in E$ and the map $B_2 = \{(x, v, u) \in U \times E \times E | (x, u) \in A_U\} \to F$, $(x, v, u) \mapsto d_u(df(\cdot, v))(x)$, is continuous; $\widetilde{d}^2 f(\cdot, v, u) : U \to F$ is a C^1_W-map on U (**3.19**) for all $(v, u) \in E \times E$ and the map $\{(x, v, u, w) \in U \times E \times E \times E | (x, w) \in A_U\} (= B_3) \to F$, $(x, v, u, w) \mapsto d_w(\widetilde{d}^2 f(\cdot, v, u))(x)$, is continuous;...; $\widetilde{d}^{k-1} f(\cdot, v_1, ..., v_{k-1}) : U \to F$ is a C^1_W-map on U (**3.19**) for all $(v_1, ..., v_{k-1}) \in E^{k-1}$ and the map:

$\{(x, v_1, ..., v_{k-1}, w) \in U \times E^{k-1} \times E | (x, w) \in A_U\} (= B_k) \to F$,

$(x, v_1, ..., v_{k-1}, w) \mapsto d_w(\widetilde{d}^{k-1} f(\cdot, v_1, ..., v_{k-1}))(x)$, is continuous.

(Then we have the extension $\widetilde{d}^k f : U \times E^k \to F$, which is continuous, is a linear map respect the last coordinate, verifies $\widetilde{d}^k f(x, v_1, ..., v_{k-1}, w) =$

$d_w(\widetilde{d}^{k-1} f(\cdot, v_1, ..., v_{k-1}))(x)$ for all $(x, v_1, ..., v_{k-1}, w) \in B_k$, and is unique).

Theorem *Let $f : U \to F$ be a C^1_W-map on U* (**3.19**)*, with $int(Q) \neq \varnothing$. Then f is a $\widetilde{C}^k_W$-map on U ($k > 1$) if and only if f is a C^k_W-map on U* (**3.25** (4))*. (See* [11]*).*

Differentiation theory in convenient vector spaces

For a detailed study of the differentiation theory in convenient vector spaces, the reader can consult the books [10] and [18].

3.29 Mackey-convergence

Definition Let E be a Hlcrtvs. We say that a net $\{x_\gamma\}_{\gamma \in \Gamma}$ in E Mackey-converges (or M-converges) to $x \in E$, if there exists a closed bounded absolutely convex subset B of E such that $\{x_\gamma | \gamma \in \Gamma\} \cup \{x\} \subset \langle B \rangle$ and the net $\{x_\gamma\}_{\gamma \in \Gamma}$ converges to x in the normed space $E_B = (\langle B \rangle, p_B)$, where $p_B(y) = \inf\{r > 0 | y \in r \cdot B\}$ for all $y \in \langle B \rangle$.

Proposition 1 *Let E be a Hlcrtvs, $\{x_\gamma\}_{\gamma \in \Gamma}$ a net in E and x a point of E. Then $\{x_\gamma\}_{\gamma \in \Gamma}$ Mackey-converges to $x \in E$, if and only if there exists a bounded absolutely convex subset B' of E such that $\{x_\gamma | \gamma \in \Gamma\} \cup \{x\} \subset \langle B' \rangle$ and the net $\{x_\gamma\}_{\gamma \in \Gamma}$ converges to x in the normed space $E_{B'} = (\langle B' \rangle, p_{B'})$.*

Note that if B and B' are bounded absolutely convex subsets of a Hlcrtvs E with $B' \subset B$, then $\langle B' \rangle \subset \langle B \rangle$, $p_B(y') \leqslant p_{B'}(y')$ for all $y' \in \langle B' \rangle$, and the inclusion $i : E_{B'} \hookrightarrow E_B$ is a continuous map, which is equivalent to $T_{p_B}|_{\langle B' \rangle} \subset T_{p_{B'}}$.

Note also that if (E, T) is a rtvs, then:

(*i*) The closure of a bounded subset of (E, T) is a bounded subset of (E, T).

(*ii*) The closure in (E, T) of an absolutely convex subset in E is again absolutely convex.

Remarks (1) A net $\{x_\gamma\}_{\gamma\in\Gamma}$ in E Mackey-converges to $x \in E$ if and only if $\{x_\gamma - x\}_{\gamma\in\Gamma}$ Mackey-converges to $0 \in E$.

(2) If the net $\{x_\gamma\}_{\gamma\in\Gamma}$ in E Mackey-converges to $x \in E$, then there exists a sequence $\{\gamma_n\}_{n\in\mathbb{N}}$ in Γ such that the sequence $\{x_{\gamma_n}\}_{n\in\mathbb{N}}$ Mackey-converges to $x \in E$.

(3) Let E be a Hlcrtvs and B a bounded absolutely convex subset of E. Since $\langle B\rangle = \bigcup_{t>0} t\cdot B$ and $\{x \in \langle B\rangle\, |p_B(x) < 1\} \subset B \subset \{x \in \langle B\rangle\, |p_B(x) \leqslant 1\}$, we have that the inclusion $i_B : E_B \hookrightarrow E$ is a continuous map, which is equivalent to $T|_{\langle B\rangle} \subset T_{p_B}$ $(E_B = (\langle B\rangle, p_B))$. Therefore, if $\{x_\gamma\}_{\gamma\in\Gamma}$ is a net in E that Mackey-converges to $x \in E$, then $\{x_\gamma\}_{\gamma\in\Gamma}$ converges to x in E.

(4) Let (E, T) be a Hlcrtvs. Then (E, T_{born}) has the same Mackey converging sequences as (E, T) and T_{born} is the final topology in E with respect to the inclusions $i_B : E_B \hookrightarrow E$ for all B bounded and absolutely convex in E. Moreover, if $\widetilde{T}$ is a locally convex topology on E and it has the same Mackey converging sequences as (E, T), then $\widetilde{T} \subset T_{born}$.

(5) Let (E, T) be a Hlcrtvs. Then, T_{born} is the final topology in E with respect to the inclusions $i_B : E_B \to E$, for all B closed bounded and absolutely convex in (E, T).

Proposition 2 *Let E be a Hlcrtvs, B a bounded absolutely convex subset of E, $\{x_\gamma\}_{\gamma\in\Gamma}$ a net in $\langle B\rangle$ and $x \in \langle B\rangle$. Then the following conditions are equivalent:*

(a) $\{x_\gamma\}_{\gamma\in\Gamma}$ converges to x in E_B.

(b) There exists a net $\{\mu_\gamma\}_{\gamma\in\Gamma}$ in $\mathbb{R}$ which converges to 0 such that $(x_\gamma - x) \in \mu_\gamma \cdot B$, for all $\gamma \in \Gamma$.

(Use the preceding Remark (3), and in the step "(*a*)$\Longrightarrow$ (*b*)" take $\mu_\gamma = \delta p_B(x_\gamma - x)$ for all $\gamma \in \Gamma$, where $\delta > 1$).

Proposition 3 *Let E be a Hlcrtvs. Then:*

(1) Let $c : \mathbb{R} \to E$ be a C^1-curve and $\{t_n\}_{n\in\mathbb{N}}$ a sequence of real numbers that converge to 0. Then, the sequence $\{c(t_n)\}_{n\in\mathbb{N}}$ Mackey-converge to $c(0) \in E$ (use the Mean value Theorem).

(2) If $c : \mathbb{R} \to E$ is a Lip^1-curve in E, then the curve $t \longmapsto \frac{1}{t}(\frac{1}{t}(c(t) - c(0)) - c'(0))$ is bounded on bounded subsets of $\mathbb{R} - \{0\}$. Therefore, if the sequence $\{t_n\}_{n\in\mathbb{N}}$ in $\mathbb{R} - \{0\}$ converges to 0, the sequence $\{\frac{c(t_n)-c(0)}{t_n}\}_{n\in\mathbb{N}}$ Mackey-converges to $c'(0) \in E$.

Proposition 4 *Let E be a Hlcrtvs. Then:*

(1) A subset A of E is closed in $C^\infty E$ if and only if for every sequence $\{x_n\}_{n\in\mathbb{N}}$ in A, which Mackey-converges to $x \in E$, the point x belongs to A.

(2) *If U is a subset of E such that $U \cap \langle B \rangle$ is open in E_B for all B bounded and absolutely convex in E, then U is C^∞-open in E.*

3.30 Mackey complete spaces

Definition 1 (Mackey-Cauchy net) Let $\{x_\gamma, \gamma \in \Gamma, \leqslant\}$ be a net in a Hlcrtvs E. This net will be called Mackey-Cauchy net if there exist a bounded absolutely convex subset B of E and a net $\{\mu_{\gamma,\gamma'}, (\gamma,\gamma') \in \Gamma \times \Gamma, \leqslant \times \leqslant\}$ in $\mathbb{R}$ converging to 0 such that $x_\gamma - x_{\gamma'} \in \mu_{\gamma,\gamma'} \cdot B$ for all $(\gamma,\gamma') \in \Gamma \times \Gamma$.

Let B be a bounded absolutely convex subset of a Hlcrtvs E. Let $\{x_\gamma\}_{\gamma\in\Gamma}$ be a net in $\langle B \rangle$. Then $\{x_\gamma\}_{\gamma\in\Gamma}$ is a Cauchy net in the normed space E_B if and only if there exists a net $\{\mu_{\gamma,\gamma'}, (\gamma,\gamma') \in \Gamma \times \Gamma, \leqslant \times \leqslant\}$ in $\mathbb{R}$ converging to 0 such that $x_\gamma - x_{\gamma'} \in \mu_{\gamma,\gamma'} \cdot B$, for all $(\gamma,\gamma') \in \Gamma \times \Gamma$ (i.e., $\{x_\gamma\}_{\gamma\in\Gamma}$ is a Mackey-Cauchy net in E). Therefore, if $\{y_\gamma\}_{\gamma\in\Gamma}$ Mackey-converges to y in E, then $\{y_\gamma\}_{\gamma\in\Gamma}$ is a Mackey-Cauchy net.

Proposition 1 *Let $\{x_\gamma\}_{\gamma\in\Gamma}$ be a Mackey-Cauchy net in a Hlcrtvs E and $x \in E$. Then, $\{x_\gamma\}_{\gamma\in\Gamma}$ converges to x in E if and only if $\{x_\gamma\}_{\gamma\in\Gamma}$ Mackey converges to $x \in E$.*

(Use the following result: *Let E be a Hrtvs, B a bounded closed subset of E and J a closed bounded interval of $\mathbb{R}$. Then $J \cdot B = \{t \cdot v | t \in J, v \in B\}$ is closed in E.*)

Note that if the Hlcrtvs E is metrizable, the condition Mackey-Cauchy for the net can be omitted. [17].

Definition 2 (Mackey complete space) The Hlcrtvs E is called Mackey complete (or convenient) if every Mackey-Cauchy net in E converges in E.

Theorem *Let E be a Hlcrtvs. Then the following conditions are equivalent:*

(i) *E is Mackey-complete.*

(ii) *Every Mackey-Cauchy net in E, Mackey converges in E.*

(iii) *Every Mackey-Cauchy sequence in E converges in E.*

(iv) *Every Mackey-Cauchy sequence in E, Mackey converges in E.*

(v) *For every absolutely convex bounded subset B of E, the normed space E_B is complete.*

(vi) *For every absolutely convex closed bounded subset B of E, the normed space E_B is complete.*

(vii) *For every bounded subset B of E there exists an absolutely convex closed bounded subset B' of E such that $B \subset B'$ and $E_{B'}$ is complete.*

(viii) *For every bounded subset B of E there exists an absolutely convex bounded subset B' of E such that $B \subset B'$ and $E_{B'}$ is complete.*

(ix) *Any Lipschitz curve in E is locally Riemann integrable.*

(x) *For any $c_1 \in C^\infty(\mathbb{R}, E)$ there is $c_2 \in C^\infty(\mathbb{R}, E)$ such that $c_2' = c_1$.*

(xi) *E is closed in the C^∞-topology of any Hlcrtvs $\widetilde{E}$, where E is a topological vector subspace of $\widetilde{E}$ (Recall that the C^∞-topology on $\widetilde{E}$ is the final topology with respect to all smooth curves $c : \mathbb{R} \to \widetilde{E}$).*

(xii) *If $c : \mathbb{R} \to E$ is a curve such that $l \circ c : \mathbb{R} \to \mathbb{R}$ is smooth for all $l \in \mathcal{L}(E, \mathbb{R})$, then c is smooth.*

(xiii) *Any continuous linear mapping from a normed space F into E has a continuous extension to the completion of the normed space F.*

Recall that $c : [a, b] \to E$ is Riemann integrable (by definition) if the net of Riemann sums converges in E. Moreover if c is continuous and E is sequentially complete (hence convenient), then c is Riemann integrable (in fact, the net of Riemann sums is a Cauchy net that has a subnet that is a sequence, and on the other hand an agglomeration point of a Cauchy net is a point of convergence of this net).

Proposition 2 *Let (E, T) be a Hlcrtvs. Then we have:*

(1) *If E is complete (every Cauchy net in E converges in E), then E is sequentially complete (every Cauchy sequence in E converges in E).*

(2) *If E is sequentially complete, then E is Mackey complete.*

(3) *If E is metrizable, then: the statements "E is complete", "E is sequentially complete", and "E is Mackey complete", are equivalents.*

(4) *(E, T) is Mackey complete if and only if (E, T_{born}) is Mackey complete (use the preceding theorem and the results of* **3.29***).*

3.31

(*i*) Let E, F be Hlcrtv spaces, and let $l : E \to F$ be a continuous linear map. Then l maps Lip^k-curves in E to Lip^k-curves in F, for all $0 \leqslant k \leqslant \infty$, and for $k > 0$ one has $(l \circ c)' = l(c'(t)),\ t \in \mathbb{R}$.

(*ii*) Let E be a Mackey complete space and $c : \mathbb{R} \to E$ a curve such that $l \circ c : \mathbb{R} \to \mathbb{R}$ is Lip^n for all $l \in \mathcal{L}(E, \mathbb{R})$. Then c is Lip^n.

(*iii*) A linear mapping $l : E \to F$ between Hlcrtv spaces is bounded (it maps bounded sets to bounded sets) if and only if it maps smooth curves in E to smooth curves in F.

3.32 Smooth maps

Let E, F be Hlcrtv spaces, K a subset of E and $f : K \to F$ a map. Then f is called smooth if it maps smooth curves in K to smooth curves in F.

Note that in this case $f : K \to F$ is continuous, where the topology of K is the C^∞-topology (recall that $C^\infty K$ is a subspace of $C^\infty E$ whenever K is C^∞-open in E or locally C^∞-closed and convex; $C^\infty K$ is the topological space K with the C^∞-topology, i.e.,the final topology with respect to all smooth curves $c : \mathbb{R} \to K \hookrightarrow E$).

Remark If $K = E = \mathbb{R}$, then f is smooth (with the preceding definition) if and only if f is a smooth curve.

Proposition 1 *Let $f : U \to F$ be a smooth map, where U is a C^∞-open subset of E (U is open in $C^\infty E$). Then:*

(i) For all $x \in U$ and all $v \in E$, there exists $\lim\limits_{t\to 0} \frac{f(x+tv)-f(x)}{t} (= d_v f(x))$.

(ii) For all $x \in U$, $d.f(x)(= df(x)) : E \to F$, $v \mapsto d_v f(x)$, is a linear and bounded map, and therefore it is smooth and, finally, it is continuous from (E, T_{born}) into F (2.16) (E and (E, T_{born}) have the same collection of bounded subsets).

(iii) $d.f(\cdot)(= df) : U \times E \to F$, $(x, v) \mapsto d_v f(x)$ is smooth.

Hint:

(1) $c : \mathbb{R} \to E$, $t \mapsto x + tv$, is a smooth curve. Then there exists $\varepsilon > 0$ such that $c(t) \in U$ for all $t \in (-\varepsilon, \varepsilon)$. Let $\sigma : \mathbb{R} \to (-\varepsilon, \varepsilon)$ be a smooth curve such that $\sigma(t) = t$ for all $t \in (-\frac{\varepsilon}{2}, \frac{\varepsilon}{2})$. Then $c \circ \sigma : \mathbb{R} \to U \hookrightarrow E$ is a smooth curve and $f \circ c \circ \sigma : \mathbb{R} \to F$ is a smooth curve. Consequently $(f \circ c \circ \sigma)'(0) = \lim\limits_{t\to 0} \frac{f(x+tv)-f(x)}{t}$.

(2) Apply **3.8** (1) and the Boman's theorem: *Let $f : A \to F$ be a map, where A is an open set of $\mathbb{R}^n$ and F is a Hlcrtvs. Then f is smooth (with the preceding definition) if and only if f is a C^∞-map* (**3.3** (5)). In this case $df : A \times \mathbb{R}^n \to F$, $(x, v) \mapsto d_v f(x)$, is smooth and consequently $d.f(x) : \mathbb{R}^n \to F$ is smooth for all $x \in A$.

Proposition 2 (chain rule) *Let E, F, G be Hlcrtv spaces, U a C^∞-open subset in E, V a C^∞-open subset in F, $f : U \to F$ a map with $f(U) \subset V$ and $g : V \to G$ a map. Suppose that f, g are smooth maps. Then $g \circ f : U \to G$ is a smooth map and $d(g \circ f)(x, v) = dg(f(x), df(x, v))$, for all $(x, v) \in U \times E$. Consequently $d(g \circ f)(x, \cdot) = dg(f(x), \cdot) \circ df(x, \cdot) : E \to G$ for all $x \in U$.*

3.33 Smooth maps on quadrants

Definition Let E, F be convenient spaces (i.e., Mackey complete spaces), Q a quadrant of E with $int_{C^\infty E}(Q) \neq \varnothing$, U a C^∞-open subset of Q and $f : U \to F$ a map. We say that f is smooth if for all smooth curve $c : \mathbb{R} \to U$, $f \circ c : \mathbb{R} \to F$ is a smooth curve (recall that if $U \neq \varnothing$, then $int_{C^\infty E}(U) \neq \varnothing$).

One has:

(*i*) Let E, F, G be convenient spaces, Q a quadrant of E with $int_{C^\infty E}(Q) \neq \varnothing$, U a C^∞-open subset of Q, P a quadrant of F with $int_{C^\infty F}(P) \neq \varnothing$, V a C^∞-open subset of P, $f : U \to V$ a smooth map and $g : V \to G$ a smooth map. Then $g \circ f : U \to G$ is a smooth map.

(*ii*) Let $f : U \to F$ be a smooth map (preceding definition). Then the map $(\rho=)f|_{int_{C^\infty E}(U)} : int_{C^\infty E}(U) \to F$ is smooth (**3.32**) and the map, $x \in int_{C^\infty E}(U) \mapsto d.\rho(x) = d\rho(x) \in \{\lambda : E \to F | \lambda \textit{ is linear and bounded}\}$, extends uniquely to a smooth map (preceding definition) $df : U \to \{\lambda : E \to F | \lambda$

$is\ linear\ and\ bounded\}(= L_b(E,F))$. (Then $df(x)$ is linear and continuous from (E, T_{born}) into F). Note that the map, $(x,v) \in int_{C^\infty E}(U) \times E \mapsto d_v\rho(x) \in F$, is smooth (**3.32**). Moreover, for all $x \in int_{C^\infty E}(U)$, $d\rho(x) : E \to F$ is smooth **(3.32)** and continuous from (E, T_{born}) into F. Note, finally, that $L_b(E,F)$ is a convenient space.

(*iii*) Let $f : U \to F$ be a smooth map and $c : \mathbb{R} \to U$ a smooth curve. Then $f \circ c : \mathbb{R} \to F$ is a smooth curve and $(f \circ c)'(t) = df(c(t))(c'(t))$, for all $t \in \mathbb{R}$.

(*iv*) **Chain rule** Let $f : U \to F$ be a smooth map and $g : V \to G$ a smooth map with $f(U) \subset V$, ($V \subset P \subset F$). Then $g \circ f : U \to G$ is a smooth map and $d(g \circ f)(x) = dg(f(x)) \circ df(x)$, for all $x \in U$.

(*v*) **Restriction to open sets** Let $f : U \to F$ be a smooth map and V a C^∞-open subset of U, ($C^\infty U$ is a topological subspace of $C^\infty E$). Then $f|_V : V \to F$ is a smooth map (of course V is a C^∞-open subset of Q) and $df(x) = d(f|_V)(x)$, for all $x \in V$.

(*vi*) **Restriction to vector subspaces** Let $f : U \to F$ be a smooth map, G a closed linear subspace of E (then a convenient space), Q_G a quadrant of G with $int_{C^\infty G}(Q_G) \neq \varnothing$. Suppose that $U \cap G$ is a C^∞-open subset of Q_G. Then $f|_{U\cap G} : U \cap G \to F$ is a smooth map and $df(x)|_G = d(f|_{U\cap G})(x)$, for all $x \in U \cap G$.

(*vii*) **Open covering property** Let E, F be convenient spaces, Q a quadrant of E with $int_{C^\infty E}(Q) \neq \varnothing$, U a C^∞-open subset of Q and $f : U \to F$ a map. Let $\{V_j | j \in J\}$ be a C^∞-open covering of U. Then $f : U \to F$ is smooth if and only if $f|_{V_j} : V_j \to F$ is smooth for all $j \in J$.

Again, we have the essential results (together with the theorem of invariance of the boundary) to construct manifolds with corners modeled on convenient spaces.

3.34

(1) Let E, F, Q, U, $f : U \to F$ be as in Definition of **3.33.** Then $f : U \to F$ is a smooth map if and only if $\lambda \circ f : U \to \mathbb{R}$ is a smooth map for all $\lambda \in \mathcal{L}(F, \mathbb{R})$ (Definition of **3.33**), (in particular F may be Fréchet or Banach).

(2) Let E, F be convenient spaces, Q a quadrant of E with $int(Q) \neq \varnothing$ (which implies $int_{C^\infty E}(Q) \neq \varnothing$), U an open subset of Q (hence U is a C^∞-open subset of Q (in this case $C^\infty Q$ is a topological subspace of $C^\infty E$)), and $f : U \to F$ a map. Then f is a smooth map (Definition of **3.33)** if and only if $\lambda \circ f : U \to \mathbb{R}$ is a smooth map (**3.33)** for all $\lambda \in \mathcal{L}(F, \mathbb{R})$. Therefore f is a smooth map **(3.33)** if and only if f is a C^∞_W-map on U (**3.25, 3.28).**

3.35

(1) Let E, F, Q, U, f be as in Definition of **3.33.** Let P be a quadrant of F with $int_{C^\infty F}(P) \neq \varnothing$ and V a C^∞-open subset of P. Suppose that $f(U) \subset V$. We say that $f : U \to V$ is a smooth diffeomorphism if $f : U \to V$ is a bijective map and f, f^{-1} are smooth maps (**3.33**). Note that in this case $f^{-1} : V \to U$ is also a smooth diffeomorphism.

(2) The composition of smooth diffeomorphisms is again a smooth diffeomorphism, and the identity map is a smooth diffeomorphism.

(3) Let $f : U \to V$ be a smooth diffeomorphism. Then by **3.33** (*ii*), we have the extensions $df : U \to L_b(E,F)$ and $d(f^{-1}) : V \to L_b(F,E)$ and we know that $df(x) \in \mathcal{L}((E,T_{born}),F)$ and $d(f^{-1})(f(x)) \in \mathcal{L}((F,T'_{born}),E)$ for all $x \in U$, where T' is the topology of F. By **3.33** (*iv*), $df(x)$ is an isomorphism and $(df(x))^{-1} = d(f^{-1})(f(x))$ for all $x \in U$. Finally $df(x)$ is a linear homeomorphism from (E,T_{born}) onto (F,T'_{born}) for all $x \in U$.

3.36 Invariance of the boundary for smooth diffeomorphisms

Lemma *Let E, F be convenient spaces, U a C^∞-open subset of E, $\lambda : F \to \mathbb{R}$ a bounded linear function with $\lambda \neq 0$, $x \in U$ and $f : U \to F^+_\lambda$ a smooth map* (**3.33**) *such that $f(x) \in F^0_\lambda$. Then $df(x)(E) \subset F^0_\lambda$* (**3.33** (*ii*)).

Theorem (Invariance of the boundary) *Let $f : U \to V$ be a smooth diffeomorphism* (**3.35** (1), **3.33**). *(By* **2.19** (*i*)*, $index(Q)$ is finite and Q^0 is closed in (E,T_{born}), $index(P)$ is finite and P^0 is closed in (F,T'_{born}), where T and T' are the topologies of E and F, respectively). Then we have:*

(a) $index(x) = index(f(x))$ *for all* $x \in U$, (**2.9** (*c*) (*ii*)).

(b) $\{x \in U |\, index(x) \geqslant 1\} \neq \varnothing$ *if and only if* $\{y \in V |\, index(y) \geqslant 1\} \neq \varnothing$. *(Note that* $int_{T_{born}}(Q) \neq \varnothing$ *and* $int_{T'_{born}}(P) \neq \varnothing$, *2.19 (i)).*

(c) If $U \neq \varnothing$, $int_{C^\infty E}(U) \neq \varnothing$ *and* $int_{C^\infty F}(V) \neq \varnothing$.

(d) $f(\{x \in U |\, index(x) \geqslant k\}) = \{y \in V |\, index(y) \geqslant k\}$, *for all* $0 \leqslant k \leqslant index(Q)$.

(e) $f(\{x \in U |\, index(x) = k\}) = \{y \in V |\, index(y) = k\}$, *for all* $0 \leqslant k \leqslant index(Q)$.

(f) $f(int_{C^\infty E}(U)) = int_{C^\infty F}(V)$ *(Note that* $int_{C^\infty T}(Q) = int_{T_{born}}(Q)$ *and* $int_{C^\infty T'}(P) = int_{T'_{born}}(P)$*).*

(g) $f|_{int_{C^\infty E}(U)} : int_{C^\infty E}(U) \to int_{C^\infty F}(V)$ *is a smooth diffeomorphism* (**3.35**) *and* $d(f|_{int_{C^\infty E}(U)})(x) = df(x)$ *for all* $x \in int_{C^\infty E}(U)$.

3.37 Let E be a convenient space and Q, Q' quadrants in E with non-empty C^∞- interior. Then Q, Q' are smooth diffeomorphic (**3.35**) if and only if $index(Q) = index(Q')$ (**3.36**, **2.13** (*iii*)).

Indeed: if $index(Q) = index(Q')$,by **2.13** (*iii*), there exists a linear homeomorphism $\alpha : (E,T_{born}) \to (E,T_{born})$ such that $\alpha(Q) = Q'$ and $\alpha(Q^0) = Q'^0$. By **2.17**, $\alpha|_Q : Q \to Q'$ is a smooth diffeomorphism (**3.35**). For the converse, apply the Theorem of **3.36.**

On constructing smooth manifolds with corners modeled on convenient vector spaces, the C^∞-open subsets of quadrants, with non-empty C^∞-interior, of convenient vector spaces will be the local models.

4 Manifolds with corners

The preceding notions of differentiability (Section 3) will be used to introduce the corresponding notions of manifolds with corners.

Manifolds with corners modeled on normable spaces

4.1 Let X be a non-void set. We say that $(U, \varphi, (E, Q))$ is a chart on X if: U is a subset of X, E is a normable rtvs, Q is a quadrant in E with $int(Q) \neq \varnothing$, $\varphi : U \to Q$ is an injective map and $\varphi(U)$ is an open subset of Q. In this case U will be called the domain, φ the morphism and E the model of the chart. Recall that (**2.12**) if $index(Q) = n \in \mathbb{N}$, there exists a linearly independent system $\Lambda = \{\lambda_1, ..., \lambda_n\}$ of elements of $\mathcal{L}(E, \mathbb{R})$ such that $Q = E^+_\Lambda$, and consequently $Q^0 = E^0_\Lambda$ (see **2.8** for the (certain) unicity).

Let $c = (U, \varphi, (E, Q))$ and $c' = (U', \varphi', (E', Q'))$ be charts on X. We say that they are compatible of class r or C^r-compatible ($r \in \{0\} \cup \mathbb{N} \cup \{\infty\} (= \mathbb{N}^*)$) if: $\varphi(U \cap U')$ and $\varphi'(U \cap U')$ are open subsets of Q and Q', respectively, and the maps $\varphi'\varphi^{-1} : \varphi(U \cap U') \to \varphi'(U \cap U')$ and $\varphi\varphi'^{-1} : \varphi'(U \cap U') \to \varphi(U \cap U')$ (transition functions) are C^r-maps (**3.3**), (hence inverse homeomorphisms). In this case we shall write $c \overset{C^r}{\backsim} c'$.

A collection $\mathcal{A}$ of charts on X is called an atlas of class r or C^r-atlas on X ($r \in \mathbb{N}^*$), if the domains of the charts of $\mathcal{A}$ cover X and any two of them are C^r-compatible.

Two atlases $\mathcal{A}$, $\mathcal{A}'$ of class r on X are called equivalent of class r or C^r-equivalent, if $\mathcal{A} \cup \mathcal{A}'$ is an atlas of class r on X. In this case we shall write $\mathcal{A} \overset{C^r}{\backsim} \mathcal{A}'$.

The properties of **3.4** permit us to prove that the preceding binary relation $\overset{C^r}{\backsim}$ is an equivalence relation over the atlases of class r on X ([3]).

If $\mathcal{A}$ is an atlas of class r on X, the equivalence class (C^r-class), $[\mathcal{A}]$, is called differentiable structure of class r on X and the pair $(X, [\mathcal{A}])$ is called manifold with corners of class r or C^r-manifold with corners (or C^r-manifold).

If $(X, [\mathcal{A}])$ is a C^r-manifold with corners, then $\bigcup\{\mathcal{B} | \mathcal{B} \in [\mathcal{A}]\} \in [\mathcal{A}]$.

The C^0-manifolds with corners will be called topological manifolds with borders. In this case for every $x \in X$, there exists a chart $(U, \varphi, (E, Q))$ such that $x \in U$ and $\operatorname{codim}(Q^0) = 1$.

If a C^r-manifold $(X, [\mathcal{A}])$ with corners admits an atlas whose charts are modeled over Banachable (or Hilbertizable) rtv spaces, then all the charts of $(X, [\mathcal{A}])$ are modeled over Banachable (or Hilbertizable) rtv spaces. In this case, we say that $(X, [\mathcal{A}])$ is a Banach (or a Hilbert) C^r-manifold with corners.

4.2 The C^r-manifolds are endowed with a natural topology determined by a basis given by the domains of all the charts ([21], paragraph 1.2).

Since a normable rtvs is metrizable, the topology of a C^r-manifold is locally metrizable, verifies the first axiom of countability, and fulfils the T_1 axiom (i.e., the unitary subsets are closed). In general, this topology is not Hausdorff, and the axiom of Hausdorff does not imply the regularity. Finally, by a well known theorem of Smirnov, a Hausdorff C^r-manifold with corners is metrizable if and only if is paracompact ([21], paragraph 1.4).

The main feature of the C^r-manifolds ($r \geqslant 1$) is the existence of a boundary with a stratified structure.

Let $(X, [\mathcal{A}])$ be a C^r-manifold with $r \geqslant 1$, $x \in X$ and let $(U, \varphi, (E, Q))$, $(U', \varphi', (E', Q'))$ be charts of $(X, [\mathcal{A}])$ with $x \in U \cap U'$. Then, by **3.6**:

$$index(\varphi(x)) = index(\varphi'(x))(= ind(x)).$$

The non-negative integer number $ind(x)$ will be called index of x in $(X, [\mathcal{A}])$, (**2.9** (c)).

Let $(X, [\mathcal{A}])$ be a C^r-manifold ($r \geqslant 1$). Then:

(a) For all $k \in \mathbb{N} \cup \{0\}$, the set $\{x \in X | \, ind(x) \geqslant k\}(= \partial^k X)$ is called k-boundary of $(X, [\mathcal{A}])$. The set $\partial^1 X$ is also denoted by ∂X. It is clear that $\partial^0 X = X$.

(b) For all $k \in \mathbb{N} \cup \{0\}$, the set $\{x \in X | \, ind(x) = k\}$ is denoted by $B_k X$. The set $B_0 X$ is called interior of X and denoted by $IntX$.

It is clear that $\{B_k X | k \in \mathbb{N} \cup \{0\}\}$ is a partition of X, $\partial^k X = \bigcup_{k' \geqslant k} B_{k'} X$, and $\partial^k X \subset \partial^{k'} X$ whenever $k' \leqslant k$.

Let X be a C^0-manifold such that every $x \in X$ has a chart of X modeled over an euclidean space. Then we can speak of $dim_x X$ for all $x \in X$ (Riesz's theorem and Brouwer's theorem). Moreover if $x \in X$, then: *there is a chart at x, $(U, \varphi, (E, Q))$, with $index(\varphi(x)) > 0$ if and only if there is a chart at x, $(U', \varphi', (E', Q'))$, with $index(\varphi'(x))$* $= 1$. In this case we say that x is a boundary point of X and all these points will be denoted by $Bord(X)$. Finally $X - Bord(X)$ will be called interior of X.

The $B_k X$ sets have a natural structure of C^r-manifold:

Lemma *Let $(X, [\mathcal{A}])$ be a C^r-manifold ($r \geqslant 1$) and $x \in X$. Then there is a chart $(U, \varphi, (E, Q))$ of $(X, [\mathcal{A}])$ such that $x \in U$ and $\varphi(x) = 0$ (centred chart at x), and hence $ind(x) = index(Q)$.*

Proposition 1 *Let X be a C^r-manifold ($r \geqslant 1$). Then $\partial^k X$ is a closed subset of X, for all $k \in \mathbb{N}$, and $IntX$ is a dense open subset of X.*

Proposition 2 *Let $(X, [\mathcal{A}])$ be a C^r-manifold ($r \geqslant 1$) and $k \in \mathbb{N} \cup \{0\}$. Then there exists a unique differentiable structure of class r on $B_k X$ such that for all $x \in B_k X$ and all chart $(U, \varphi, (E, Q))$ of $(X, [\mathcal{A}])$ with $x \in U$ and $\varphi(x) = 0$, the triplet $(U \cap B_k X, \varphi|_{U \cap B_k X}, Q^0)$ is a chart of that structure.*

Furthermore $B_k X$ (with this structure) has not boundary, that is, $\partial(B_k X) = \varnothing$, and the topology of the manifold $B_k X$ is the topology induced by X.

Note that if $x \in B_k X$ and $(U, \varphi, (E, Q))$ is a chart of $(X, [\mathcal{A}])$ with $x \in U$ and $\varphi(x) = 0$, then $\varphi(U \cap B_k X) = \varphi(U) \cap Q^0$.

Corollary *Let $(X, [\mathcal{A}])$ be a C^r-manifold ($r \geqslant 1$). Then:*

(a) There is a unique differentiable structure of class r on $Int(X) = B_0 X$, such that for all $x \in Int(X)$ and all chart $(U, \varphi, (E, Q))$ of $(X, [\mathcal{A}])$ with $x \in U$ and $\varphi(x) = 0$ (Hence $Q^0 = E$ and $U \subset Int(X)$), the triplet (U, φ, E) is a chart of $Int(X)$ with that structure. Furthermore $Int(X)$ has not boundary and the topology of the manifold $Int(X)$ is the topology induced by X.

(b) If $\partial^2 X = \varnothing$, (hence $\partial X = B_1 X$), there is a unique differentiable structure of class r on ∂X such that for all $x \in \partial X$ and all chart $(U, \varphi, (E, Q))$ of $(X, [\mathcal{A}])$ with $x \in U$ and $\varphi(x) = 0$ (hence $index(Q) = 1$ and $\mathrm{codim}(Q^0) = 1$) the triplet

$(U \cap \partial X, \varphi|_{U\cap\partial X}, Q^0)$ is a chart of that structure. Furthermore $\partial(\partial X) = \varnothing$ and the topology of the manifold ∂X is the topology induced by X.

4.3 Given a C^r-manifold X ($r \geqslant 1$) and $x \in X$, the definition of $dim_x X$ is the natural one. We note that if X is a topological manifold, $x \in X$ and $(U, \varphi, (E, Q))$, $(U', \varphi', (E', Q'))$ are charts of X at x, then in general E and E' are not linearly homeomorphic.

If X is a C^r-manifold ($r \geqslant 1$), C is a connected component of X and x, z are elements of C, then $dim_x X = dim_z X$. Finally if X is a C^r-manifold ($r \geqslant 1$), then: *X is a locally compact space if and only if $dim_x X$ is finite for all $x \in X$*. Manifolds with this property are called locally finite-dimensional manifolds. On the other hand, even though the infinite-dimensional Banach C^r-manifolds are not locally compact, they always are Baire spaces ([21], 1.4.7).

The quadrants used in **4.1** are themselves C^∞-manifolds with corners. If X is a C^r-manifold ($r \geqslant 1$) and G is an open subset of X, then G is a C^r-manifold whose topology is the induced by X and $B_k(G) = B_k(X) \cap G$ for all $k \geqslant 0$.

4.4 Weakened differentiable manifolds

Let $(X, [\mathcal{A}])$ be a C^r-manifold and $s \in \mathbb{N} \cup \{0\}$ with $s < r \leqslant \infty$. Then, since a C^r-map is a C^s-map (**3.3** (5) (ii)), we have the C^s-manifold $(X, [\mathcal{A}]_s)$, (weakened manifold), ([21], 1.3.1).

Note that the manifolds $(X, [\mathcal{A}])$ and $(X, [\mathcal{A}]_s)$ have the same topology.

4.5 Calculus on manifolds modeled on normable spaces

Let $(X, [\mathcal{A}])$ be a C^r-manifold ($r \geqslant 1$), F a Hlcrtvs and $f : X \to F$ a map. We say that f is a C^r-map if for every $x \in X$ there exists a chart $(U, \varphi, (E, Q))$ of $(X, [\mathcal{A}])$ at x such that $f \circ \varphi^{-1} : \varphi(U) \to F$ is a C^r-map (**3.3**).

Note that if $f : (X, [\mathcal{A}]) \to F$ is a C^r-map, then for all $s < r$, the map $f : (X, [\mathcal{A}]_s) \to F$ is a C^s-map.

It is easy to prove that if $f : (X, [\mathcal{A}]) \to F$ is a C^r-map and $(V, \psi, (E', Q'))$ is a chart of $(X, [\mathcal{A}])$, then $f \circ \psi^{-1} : \psi(V) \to F$ is a C^r-map (**3.3**).

Of course when X is a local manifold, this definition and **3.3** are equivalent.

Let X, X' be C^r-manifolds ($r \geqslant 1$) and $f : X \to X'$ a map. We say that f is a C^r-map if for every $x \in X$ there is a chart $(U, \varphi, (E, Q))$ of X at x and there is a chart $(U', \varphi', (E', Q'))$ of X' at $f(x)$ such that $f(U) \subset U'$ and the map $\varphi' \circ f \circ \varphi^{-1} : \varphi(U) \to \varphi'(U')$ is a C^r-map (**3.3**).

This definition extends naturally to the case $r = 0$, that is, to topological manifolds.

See [21], 1.3, for elementary results (now for normable spaces).

4.6 Diffeomorphisms of class r ($r \geqslant 1$) preserve the boundary

Let X, X' be C^r-manifolds ($r \geqslant 1$) and $f : X \to X'$ a map. Then f is called a C^r-diffeomorphism if f is a bijective map and f, f^{-1} are C^r-maps (**4.5**).

This definition extends naturally to the case $r = 0$, that is, to the case of topological manifolds and then the notions of diffeomorphism of class 0 and homeomorphism coincide.

See [21], 1.3 (now for normable spaces).

Lemma *Let $(X, [\mathcal{A}])$ be a C^r-manifold ($r \geqslant 1$), U an open subset of X, E a normable rtvs, Q a quadrant of E with $int(Q) \neq \varnothing$ and $\varphi : U \to E$ a map such that $\varphi(U)$ is an open subset of Q. Then the following statements are equivalent:*

(a) $(U, \varphi, (E, Q))$ *is a chart of* $(X, [\mathcal{A}])$.

(b) $\varphi : U \to \varphi(U)$ *is a* C^r*-diffeomorphism.*

Theorem *Let X, X' be C^r-manifolds ($r \geqslant 1$) and $f : X \to X'$ a C^r-diffeomorphism. Then we have that:*

(i) $ind(x) = ind(f(x))$ *for all* $x \in X$, (**4.2**).

(*ii*) $f(\partial^k X) = \partial^k X'$ and $f(B_k X) = B_k X'$ for all $k \in \mathbb{N} \cup \{0\}$, (**4.2**).

Proposition *Let $f : X \to X'$ be a C^r-diffeomorphism ($r \geqslant 1$). Then, for all $k \in \mathbb{N} \cup \{0\}$, $f|_{B_k X} : B_k X \to B_k X'$ is a C^r-diffeomorphism* (**4.2**). In particular, if $\partial^2 X = \varnothing$, f is a C^r-diffeomorphism of $\partial X = B_1 X$ onto $\partial X' = B_1 X'$.

Remark Let X, X' be locally finite dimensional topological manifolds and $f : X \to X'$ a homeomorphism. Then (by Brouwer's theorem) $f(Bord(X)) = Bord(X')$, (**4.2**).

The main purpose of Differential Topology is to study the properties of the differentiable manifolds which are preserved by diffeomorphisms.

Remarks (1) If two topological manifolds, over a set, have the same associated topology, then these topological manifolds coincide.

(2) J. Milnor constructed several differentiable structures of class ∞ over the sphere S^7 whose associated topologies are the usual topology of S^7 and any two of them are not diffeomorphic, [23].

(3) S. Donaldson has proved that there are topological manifolds, modeled over $\mathbb{R}^4$, which do not admit compatible differentiable structures, [6].

(4) The only euclidean space that admits non-diffeomorphic differentiable structures with the usual topology as associated topology is $\mathbb{R}^4$, [12].

4.7 Tangent space and tangent linear map

Let $(X, [\mathcal{A}])$ be a C^r-manifold ($r \in \mathbb{N} \cup \{\infty\}$) and, a, a point of X. Consider the class $\{(c, v) | c = (U, \varphi, (E, Q))\ is\ a\ chart\ of\ (X, [\mathcal{A}])\ at\ a\ and\ v \in E\}$(=$C_a(X)$) and the binary relation $\backsim$ on $C_a(X)$ defined by: $(c, v) \backsim (c', v')$ if and only if $D(\varphi' \circ \varphi^{-1})(\varphi(a))(v) = v'$ (**3.3**). This binary relation is an equivalence relation on $C_a(X)$ and the quotient set $C_a(X)/\sim$ will be denoted by $T_a((X, [\mathcal{A}]))$.

Then we have the following properties:

(1) For every chart $c = (U, \varphi, (E, Q))$ of $(X, [\mathcal{A}])$ at a, the map $\theta_c^a : E \to T_a X$, $v \mapsto [(c, v)]$, is bijective.

(2) There is a unique structure of rtvs on T_aX such that for every chart $c = (U, \varphi, (E, Q))$ of $(X, [\mathcal{A}])$ at a, the map $\theta_c^a : E \to T_aX$ is a linear homeomorphism. This structure of rtvs on T_aX is normable (Banachable in the realm of Banach C^r-manifolds)

(3) If c and c' are charts of X at a, then $(\theta_{c'}^a)^{-1}\theta_c^a = D(\varphi'\varphi^{-1})(\varphi(a))$. The normable rtvs T_aX will be called tangent space of $(X, [\mathcal{A}])$ at a.

(4) For $1 < s < r$, $T_a((X, [\mathcal{A}]))$ and $T_a((X, [\mathcal{A}]_s))$ are linearly homeomorphic.

Let $f : X \to X'$ be a C^r-map ($r \geqslant 1$) and $a \in X$. Then, there is a unique continuous linear map, $T_af : T_aX \to T_{f(a)}X'$, such that for every chart c of X at a and every chart c' of X' at $f(a)$ it holds $T_af = \theta_{c'}^{f(a)} \circ D(\varphi' \circ f \circ \varphi^{-1})(\varphi(a)) \circ (\theta_c^a)^{-1}$. This map, T_af, will be called tangent linear map to f at the point a of X.

See [21], 1.6 (now for normable spaces).

4.8 Kinematic interpretation of the tangent vectors

Let X be a C^r-manifold ($r \geqslant 1$) and $x \in X$. A curve of class s on X with origin x, $0 \leqslant s \leqslant r$, is a map $\alpha : J \to X$ of class s, where $J = [0, a)$ or $J = (b, 0]$ or $J = (c, d)$ with $0 \in (c, d)$, such that $\alpha(0) = x$.

If α is a curve of class s on X ($1 \leqslant s \leqslant r$) with origin x defined on $J = [0, a)$, then the element $T_0(\alpha)\theta_{c_0}^0(1)$ of T_xX, where $c_0 = ([0, a), i, (\mathbb{R}, [0, \to)))$, ($\theta_{c_0}^0 : \mathbb{R} \to T_oJ$, $T_0(\alpha) : T_0J \to T_xX$), will be called tangent vector to α at 0 and will be denoted by $\dot{\alpha}(0)$. Analogously if $J = (b, 0]$ or $J = (c, d)$. We note that if $c = (U, \varphi, (E, Q))$ is a chart of X at x, then $\dot{\alpha}(0) = \theta_c^x \lim_{t\to 0+} \frac{\varphi\alpha(t) - \varphi\alpha(0)}{t} = \theta_c^x(\varphi\alpha)'(0)$, where $\theta_c^x : E \to T_xX$ is the natural linear homeomorphism. Analogously for $(b, 0]$ and (c, d).

Let X be a C^r-manifold ($r \geqslant 1$) and $x \in X$. Then $v \in T_xX$ is said to be inner (or interior) tangent vector if there is $\alpha : [0, a) \to X$, curve of class 1 on X with origin x, such that $\dot{\alpha}(0) = v$ (analogously $v \in T_xX$ is said to be outer (or exterior) tangent vector if there is $\beta : (b, 0] \to X$, curve of class 1 on X with origin x, such that $\dot{\beta}(0) = v$). The set of the inner tangent vectors at $x \in X$ will be denoted by $(T_xX)^i$ and the set of the outer tangent vectors at $x \in X$ will be denoted by $(T_xX)^e$. The set $(T_xX)^i \cap (T_xX)^e$ will be denoted by $(T_xX)^{ie}$.

Let X be a C^r-manifold ($r \geqslant 1$), $x \in X$ and $v \in T_xX$. Then, $v \in (T_xX)^{ie}$ if and only if there exists $\alpha : (c, d) \to X$, curve of class 1 on X with origin x, such that $\dot{\alpha}(0) = v$.

Proposition 1 *Let X be a C^r-manifold ($r \geqslant 1$) and $x \in X$. Then $(T_xX)^i = -(T_x)^e$ and $T_xX = L((T_xX)^i) = L((T_xX)^e)$.*

Proposition 2 *Let X be a C^r-manifold ($r \geqslant 1$), $x \in X$ and $c = (U, \varphi, (E, Q))$ a chart of X such that $x \in U$ and $\varphi(x) = 0$,* (**4.2**). Let Λ be as in (**2.12, 2.8, 4.1**). Then we have $\theta_c^x(Q) = (T_xX)^i = (T_xX)^+_{\Lambda\circ(\theta_c^x)^{-1}}$, $\theta_c^x(-Q) = (T_xX)^e$, $(\theta_c^x : E \to T_xX)$.

Proposition 3 *Let X be a C^r-manifold ($r \geqslant 1$), $x \in X$ and $c = (U, \varphi, (E, Q))$, $c' = (U', \varphi', (E', Q'))$ charts of X such that $x \in U \cap U'$ and $\varphi(x) = \varphi'(x) = 0$. Then $\theta_c^x(int(Q)) = \theta_{c'}^x(int(Q')) \subset (T_xX)^i$.*

The elements of $\theta_c^x(int(Q))$ will be called strictly inner (interior) tangent vectors at x and this set of vectors will be denoted by $(T_xX)_s^i$. The elements of $-\theta_c^x(int(Q))$ will be called strictly outer (exterior) tangent vectors at x. The set $-(T_xX)_s^i$ will be denoted by $(T_xX)_s^e$.

Note that if X is a C^r-manifold ($r \geqslant 1$) and $x \in Int(X) = B_0X$, then $T_xX = (T_xX)^i = (T_xX)^e$.

Proposition 4 *Let $c = (U, \varphi, (E, Q))$ be a chart at a point x of a C^r-manifold X ($r \geqslant 1$), $v \in T_xX$, and let Λ be as in* (**2.12, 2.8, 4.1**). Then:

(1) $v \in (T_xX)^i$ if and only if $\lambda \in \Lambda$ and $\lambda\varphi(x) = 0$ imply $\lambda(\theta_c^x)^{-1}(v) \geqslant 0$.

(2) $v \in (T_xX)_s^i$ if and only if $\lambda \in \Lambda$ and $\lambda\varphi(x) = 0$ imply $\lambda(\theta_c^x)^{-1}(v) > 0$.

Proposition 5 *Let $f : X \to X'$ be a C^r-map ($r \geqslant 1$) and x a point of X. Then $T_xf((T_xX)^d) \subset (T_{f(x)}X')^d$ for $d = i, e, ie$.*

4.9 Tangent bundle manifold

Let X be a C^r-manifold ($r \geqslant 1$). We denote by TX the set $\{(x, v)|x \in X,\ v \in T_xX\}$ and by τ_X the map $\tau_X : TX \to X,\ (x, v) \mapsto x$. We construct, in a natural way, a unique structure of differentiable manifold of class $r - 1$ on TX (topological manifold, if $r = 1$) such that for every chart $c = (U, \varphi, (E, Q))$ of X, d_c is a chart of this structure, where $d_c = (\tau_X^{-1}(U), \varphi_c, (E \times E, Q \times E))$, $\varphi_c : \tau_X^{-1}(U) \to E \times E$, $\varphi_c((x, v)) = (\varphi(x), (\theta_c^x)^{-1}(v))$. One has that: *if $(x, v) \in TX$, then $ind((x, v)) = ind(x)$*. (See [21], 1.6, now for normable spaces).

If $f : X \to X'$ is a C^r-map ($r \geqslant 1$), then $Tf : TX \to TX'$ defined by $Tf(x, v) = (f(x), T_xf(v))$, is a C^{r-1}-map. Moreover, one has that $T(1_X) = 1_{TX}$ and $T(g \circ f) = T(g) \circ T(f)$, and therefore, $T(f)$ is a C^{r-1}-diffeomorphism whenever f is a C^r-diffeomorphism. Thus TX is a differential invariant of the C^r-manifold X, and the principal differential invariants on X are constructed on TX.

4.10 Product of manifolds with corners

The category of manifolds with corners is the suitable category in which we can define finite products.

Let X,Y be C^r-manifolds with corners ($r \geqslant 1$). Then there is a unique structure of differentiable manifold with corners of class r, $[\mathcal{A}]$, in $X \times Y$ such that for every chart $c = (U, \varphi, (E, Q))$ of X and every chart $d = (V, \psi, (F, P))$ of Y, $c \times d = (U \times V, \varphi \times \psi, (E \times F, Q \times P))$, (**2.10**), is a chart of $(X \times Y, [\mathcal{A}])$. The pair $(X \times Y, [\mathcal{A}])$ will be called product manifold of X and Y. Recall that if $index(Q) = n$, $index(P) = m$, $Q = E_\Lambda^+$ and $P = F_M^+$ (**2.12** (ii)), then $(E \times F)^+_{\Lambda p_1 \cup M p_2} = Q \times P$ and $(E \times F)^0_{\Lambda p_1 \cup M p_2} = Q^0 \times P^0$ (**2.10**).

One has the following properties:

(i) The topology of the product manifold $X \times Y$ is the product of the topologies of X and Y.

(ii) For all $(x, y) \in X \times Y$, $ind(x, y) = ind(x) + ind(y)$, (**4.2**).

(*iii*) For all $k \in \mathbb{N} \cup \{0\}$, $\partial^k(X \times Y) = \bigcup_{n+m=k} (\partial^n X \times \partial^m Y)$. In particular, $\partial(X \times Y) = (\partial X \times Y) \cup (X \times \partial Y)$.

(*iv*) For all $k \in \mathbb{N} \cup \{0\}$, $B_k(X \times Y) = \bigcup_{n+m=k} (B_n X \times B_m Y)$ and these $B_n X \times B_m Y$ are pairwise disjoint open subsets of $B_k(X \times Y)$. In particular, $Int(X \times Y) = Int(X) \times Int(Y)$.

(*v*) The projections p_1 and p_2 are C^r-maps, and for every $(x, y) \in X \times Y$, the map $(T_{(x,y)}p_1, T_{(x,y)}p_2) : T_{(x,y)}(X \times Y) \to (T_x X) \times (T_y Y)$ is a linear homeomorphism. Moreover, $(T_{(x,y)}p_1, T_{(x,y)}p_2) = (\theta_c^x \times \theta_d^y) \circ (\theta_{c\times d}^{(x,y)})^{-1}$, and $(Tp_1, Tp_2) : T(X \times Y) \to TX \times TY$ is a C^{r-1}-diffeomorphism.

(*vi*) For every $(x, y) \in X \times Y$, the maps $j_x : Y \to X \times Y$, $y \mapsto (x, y)$, and $j_y : X \to X \times Y$, $x \mapsto (x, y)$, are of class r, and $(T_{(x,y)}p_1, T_{(x,y)}p_2) \circ T_x j_y(v) = (v, 0)$, $(T_{(x,y)}p_1, T_{(x,y)}p_2) \circ T_y j_x(u) = (0, u)$.

(*vii*) Let X, Y, Z be C^r-manifolds ($r \geqslant 1$) and $f : X \to Y$, $g : X \to Z$ maps. Then:

- (*a*) (f, g) is a C^r-map if and only if f, g are C^r-maps.
- (*b*) If (f, g) is a C^r-map, then $(T_{(f(x),g(x))}p_1, T_{(f(x),g(x))}p_2) \circ T_x(f, g) = (T_x f, T_x g)$ for all $x \in X$, and $(Tp_1, Tp_2) \circ T((f, g)) = (Tf, Tg)$.

(*viii*) Let X, Y, X', Y' be C^r-manifolds ($r \geqslant 1$) and $f : X \to X'$, $g : Y \to Y'$ maps. Then:

- (*a*) $f \times g$ is a C^r-map if and only if f, g are C^r-maps.
- (*b*) If $f \times g$ is a C^r-map, then $(T_{(f(x),g(y))}p'_1, T_{(f(x),g(y))}p'_2) \circ T_{(x,y)}(f \times g) = ((T_x f) \times (T_y g)) \circ (T_{(x,y)}p_1, T_{(x,y)}p_2)$, for all $(x, y) \in X \times Y$, and $(Tp'_1, Tp'_2) \circ T(f \times g) = (Tf \times Tg) \circ (Tp_1, Tp_2)$ (See [21], 2.3, now for normable spaces).

By induction we construct the finite products, in this category of manifolds with corners. Note that in the category of manifolds without boundary, the construction of products follows in a natural way without dificulties.

In order to develop this theory (manifolds with corners over normable rtv spaces) we have a great inconvenience: the inverse mapping theorem fails (See [21], 2.2, for the Banach case).

4.11 Submanifolds

Let X be a C^r-manifold ($r \geqslant 1$) and X' a subset of X. We say that X' is a submanifold of class r of X if for every point x' of X' there are a chart $c = (U, \varphi, (E, Q))$ of the C^r-manifold X with $x' \in U$ and $\varphi(x') = 0$, a closed linear subspace E' of E that admits a topological supplement G in E (that is, $E = F + G$, $F \cap G = \{0\}$ and the map $\alpha : F \times G \to E$, $(u, v) \mapsto u + v$, is a linear homeomorphism) and a quadrant Q' of E' with $int_{E'}(Q') \neq \varnothing$ such that $\varphi(U \cap X') = \varphi(U) \cap Q'$ and $\varphi(U) \cap Q'$ is an open subset of Q' (See [21], 3.1). In this case we say that c is adapted to X' at x' through (E', Q').

If we omit (for E') "that admits a topological supplement G in E", then we say that X' is a non-splitting-submanifold of class r of X.

Proposition 1 *Let X be a C^r-manifold ($r \geqslant 1$), X' a subset of X, x' a point of X', $c = (U, \varphi, (E, Q))$ a chart of X at x' with $\varphi(x') = 0$, E' a closed linear subspace of E that admits a topological supplement in E and Q' a quadrant of E' with $int_{E'}(Q') \neq \varnothing$. Then the following statements are equivalent:*

(a) $\varphi(U \cap X') = \varphi(U) \cap Q'$ *and this set is open in* Q'.

(b) $\varphi(U \cap X') = \varphi(U) \cap Q'$ *and* $Q' \subset Q$.

Note that "non-splitting-submanifold" does not imply "submanifold'.

Proposition 2 *Let X' be a submanifold of class r of X. Then there is a unique differentiable structure $[\mathcal{A}']$ of class r in X' such that for every point x' of X' and every chart $c = (U, \varphi, (E, Q))$ of X at x' with $\varphi(x') = 0$, adapted to X' at x' through (E', Q'), $c' = (U \cap X', \varphi|_{U \cap X'}, (E', Q'))$ is a chart of $[\mathcal{A}']$.*

Proposition 3 *Let X' be a submanifold of class r of X. Then the topology of X' as a manifold, is the topology induced by the topology of X. Moreover $X' = A \cap C$, where A is an open subset of X and C is a closed subset of X.*

Proposition 4 *Let X' be a submanifold of class r of X and $j : X' \hookrightarrow X$ the inclusion map. Then we have:*

(i) *j is a C^r-map, and for all C^r-manifold Y and all map $f : Y \to X'$ one has that: f is a C^r-map if and only if $j \circ f : Y \to X$ is a C^r-map.*

(ii) *Let $[\mathcal{A}'']$ be a differentiable structure of class r on X' such that for all C^r-manifold Y and all map $f : Y \to X'$ one has that: f is C^r if and only if $j \circ f : Y \to X$ is C^r. Then $[\mathcal{A}''] = [\mathcal{A}']$, (Proposition 2).*

(iii) *For every point x' of X', $T_{x'}(j)$ is an injective continuous linear map whose image $im(T_{x'}(j))$ admits a topological supplement in $T_{x'}X$.*

(iv) *If $c = (U, \varphi, (E, Q))$ is a chart of X at $x' \in X'$ with $\varphi(x') = 0$, adapted to X' at x' through (E', Q'), then for all $y' \in U \cap X'$, $T_{y'}j(T_{y'}X') = \theta_c^{y'}(E')$, $(\theta_c^{y'} : E \to T_{y'}X)$.*

By analogy with the *general topology*, we may say that $[\mathcal{A}']$ (Proposition 2) is the initial differentiable structure of class r on X' respect the map $j : X' \hookrightarrow (X, [\mathcal{A}])$.

Definition (Special submanifolds) Let X' be a submanifold of class r of X. Then:

(i) X' is a neat submanifold of class r of X if $\partial X' = (\partial X) \cap X'$.

(ii) X' is a totally neat submanifold of class r of X if the following conditions holds:

(1) For all point x' of X', $ind_{X'}(x') = ind_X(x')$ (that is, $B_k X' = X' \cap B_k X$ for all $k \in \mathbb{N} \cup \{0\}$).

In fact (1) is equivalent to:

(2) $\partial X' = (\partial X) \cap X'$ and $T_{x'}X = (T_{x'}j')(T_{x'}X') + (T_{x'}j)(T_{x'}B_k X)$ for all $x' \in X' \cap B_k X$, where $j' : X' \hookrightarrow X$ and $j : B_k X \hookrightarrow X$ are the inclusion maps. (See [21], 3.1, now for normable spaces).

In [1] the authors (et al.) give conditions, over the Banach C^r-manifolds with corners, which imply that these manifolds can be considered as closed submanifolds of a Banach space. The proof of this result can be adapted to prove the following:

Theorem *Let X be a Hausdorff paracompact C^r-manifold ($r \in \mathbb{N}$ if $\partial X \neq \varnothing$) with corners whose charts are modeled over C^r-normal normable rtv spaces (i.e. normable rtv spaces where every two disjoint closed subsets can be functionally separated by a function of class r). Then X is C^r-diffeomorphic to a closed C^r-submanifold of a normed space.*

For manifolds without boundary, J. Eells and K.D. Elworthy proved in [9] the following immersion theorem: *Let E be a C^∞-smooth Banach space of infinite dimension, with a Schauder base. Suppose that X is a separable metrizable C^∞-manifold without boundary modeled on E. If X is parallelizable, then X is C^∞-diffeomorphic to an open subset of E.*

4.12 Quotient manifolds

An important question in the theory of manifolds is the construction of quotient manifolds. If $(X, [\mathcal{A}])$ is a C^r-manifold with corners ($r \geqslant 1$) and R is an equivalence relation on X, it is not always possible to construct a differentiable structure of class r on the quotient set X/R compatible with the quotient topology. To obtain positive results one proceeds as follows:

Definition 1 A C^r-map $f : X \to X'$ is a C^r-submersion at $x \in X$ if admits a C^r local section at x. In the case that f is a C^r-submersion at every point x of X, we say that f is a C^r-submersion of X into X'.

We take this definition in order to have the following universal property: *If $f : X \to X'$ is a surjective C^r-submersion and $g : X' \to X''$ is a map, where X'' is a C^r-manifold, then g is a C^r-map if and only if $g \circ f$ is a C^r-map.*

Definition 2 Let $(X, [\mathcal{A}])$ be a C^r-manifold with corners and R an equivalence relation on X. Then we say that R is regular if there is a differentiable structure of class r in X/R such that the natural projection $p : X \to X/R$ is a C^r-submersion.

If R is a regular equivalence relation on the C^r-manifold X, then the differentiable structure of class r on X/R such that $p : X \to X/R$ is a C^r-submersion is unique and its topology is the quotient topology of X respect to R.

Remark If X is a Banach (or a Hilbert, or a locally finite-dimensional) C^r-manifold and R is a regular equivalence relation on X, then the quotient manifold X/R is a Banach (or a Hilbert, or a locally finite-dimensional, respectively) C^r-manifold.

Proposition *Let $(X, [\mathcal{A}])$ be a C^r-manifold and R a regular equivalence relation on X. Let $[\mathcal{A}'']$ be a differentiable structure of class r on X/R such that for all C^r-manifold Y and all map $f : X/R \to Y$, one has that: f is C^r if and only if $f \circ p : X \to Y$ is C^r. Then $(X/R, [\mathcal{A}''])$ is the quotient C^r-manifold introduced in the preceding Definition* 2.

By analogy with the *general topology*,we may say that the quotient manifold is the final differentiable structure of class r on X/R respect the map $p : (X, [\mathcal{A}]) \to X/R$.

Theorem *Let X be a Banach C^r-manifold with corners and R an equivalence relation on X. Then the following statements are equivalent:*

(1) R is regular and $p(\partial X) = \partial(X/R)$.

(2) R is a neat C^r-submanifold of $X \times X$ such that $R[\partial X] = \partial X$, and $(p_1)|_R : R \to X$

is a C^r-submersion ($p_1 : X \times X \to X$ is the first projection).

For a proof of the preceding Theorem see [20], and for a characterization of regular equivalence relations such that the quotient Banach manifold has not boundary see [21], paragraph 4.3.

The authors (et al.) have extensively developed the theory of C^r-manifolds with corners in the realm of Banach spaces, i.e. Banach C^r-manifolds with corners. The development of the theory of C^r-manifolds with corners in the context of normable spaces may run analogously, taking account that now the inverse mapping theorem fails.

Manifolds with corners modeled on Hlcrtv spaces

Now we introduce the manifolds with corners in the context of Hlcrtv spaces.

4.13 Let X be a non-void set. We say that $(U, \varphi, (E, Q))$ is a chart on X if one verifies : U is a subset of X, E is a Hlcrtvs, Q is a quadrant in E with $int(Q) \neq \varnothing$, $\varphi : U \to Q$ is an injective map and $\varphi(U)$ is an open subset of Q (if $index(Q) = n \in \mathbb{N}$, there exists a linearly independent system $\Lambda = \{\lambda_1, ..., \lambda_n\}$ of elements of $\mathcal{L}(E, \mathbb{R})$ such that $Q = E^+_\Lambda$).

Let $(U, \varphi, (E, Q)), (U', \varphi', (E', Q'))$ be charts on X. We say that they are C^r_W-compatible ($r \in \{0\} \cup \mathbb{N} \cup \{\infty\} (= \mathbb{N}^*)$) if:

$\varphi(U \cap U')$ is open in Q, $\varphi'(U \cap U')$ is open in Q' and the maps $\varphi'\varphi^{-1} : \varphi(U \cap U') \to \varphi'(U \cap U')$ and $\varphi\varphi'^{-1} : \varphi'(U \cap U') \to \varphi(U \cap U')$ are C^r_W-maps (**3.19, 3.25, 3.28**). Note that C^0_W-map means continuous map.

A collection $\mathcal{A}$ of charts on X is called a C^r_W-atlas on X ($r \in \mathbb{N}^*$), if the domains of the charts of $\mathcal{A}$ cover X and any two of them are C^r_W-compatible.

Two C^r_W-atlases $\mathcal{A}$, $\mathcal{A}'$ on X are called C^r_W-equivalent, if $\mathcal{A} \cup \mathcal{A}'$ is a C^r_W-atlas on X. This binary relation is an equivalence relation over the C^r_W-atlases on X (**3.26**).

If $\mathcal{A}$ is a C^r_W-atlas on X, the equivalence class $[\mathcal{A}]$ is called C^r_W-differentiable structure on X and the pair $(X, [\mathcal{A}])$ is called C^r_W-manifold with corners (or C^r_W-manifold).

If $(X, [\mathcal{A}])$ is a C^r_W-manifold, the set $\{U \subset X | U \ is\ a\ domain\ of\ a\ chart\ of\ (X, [\mathcal{A}])\}$ is a basis of a topology $T_{[\mathcal{A}]}$ on X (called the natural topology induced by $[\mathcal{A}]$). This topology verifies the T_1 axiom, and fulfils the first axiom of countability if and only if the manifold is modeled on metrizable Hlcrtv spaces (hence locally metrizable). Moreover, a C^r_W- manifold is metrizable if and only if it is Hausdorff paracompact and is modeled on metrizable Hlcrtv spaces (use the quoted Smirnov Theorem).

Let $(X, [\mathcal{A}])$ be a C^r_W-manifold with $r \geqslant 1$, x a point of X and $(U, \varphi, (E, Q))$, $(U', \varphi', (E', Q'))$ charts of $(X, [\mathcal{A}])$ with $x \in U \cap U'$.Then, by **3.24**, $index(\varphi(x)) = index(\varphi'(x))(= ind(x))$. The non-negative integer number $ind(x)$ will be called index of x in $(X, [\mathcal{A}])$.

Let $(X, [\mathcal{A}])$ be a C^r_W-manifold ($r \geqslant 1$). Then:

(*a*) For all $k \in \mathbb{N} \cup \{0\}$, the set $\{x \in X |\, ind(x) \geqslant k\} (= \partial^k X)$ is called k-boundary of $(X, [\mathcal{A}])$.

(*b*) For all $k \in \mathbb{N} \cup \{0\}$, the set $\{x \in X |\, ind(x) = k\}$ is denoted by $B_k X$.

In particular we have the Fréchet-manifolds, when we deal with Fréchet spaces as a models. Since Fréchet spaces are complete, we have that Fréchet-manifolds are Baire spaces. Moreover, by the Theorem of **2.12**, the topology of a Fréchet C^r-manifold with corners is the topology of a Hilbert C^0-manifold. This type of manifolds (Fréchet-manifolds) has been studied by R. Hamilton, in the context of manifolds without boundary, in [14].

In the realm of normable rtv spaces we have the concept **4.1** of manifold and the present one. The result **3.27** gives the relations between these two concepts.

Lemma *Let $(X,[\mathcal{A}])$ be a C^r_W-manifold ($r \geqslant 1$) and x a point of X. Then there is a chart $(U,\varphi,(E,Q))$ of $(X,[\mathcal{A}])$ such that $x \in U$ and $\varphi(x) = 0$ (centred chart at x), and hence $ind(x) = index(Q)$.*

Proposition 1 *Let $(X,[\mathcal{A}])$ be a C^r_W-manifold ($r \geqslant 1$). Then $\partial^k X$ is a closed subset of X, for all $k \in \mathbb{N}$, and $Int(X) = B_0 X$ is a dense open subset of X.*

Proposition 2 *Let $(X,[\mathcal{A}])$ be a C^r_W-manifold ($r \geqslant 1$) and $k \in \mathbb{N} \cup \{0\}$. Then there exists a unique C^r_W-differentiable structure on $B_k X$ such that for all point x of $B_k X$ and all chart $(U,\varphi,(E,Q))$ of $(X,[\mathcal{A}])$ with $x \in U$ and $\varphi(x) = 0$, the triplet $(U \cap B_k X, \varphi|_{U \cap B_k X}, Q^0)$ is a chart of that C^r_W-differentiable structure. Furthermore $\partial(B_k X) = \varnothing$ (with this C^r_W-differentiable structure), and the topology of the C^r_Wmanifold $B_k X$ is the topology induced by X.*

Note that if x is a point of $B_k X$ and $(U,\varphi,(E,Q))$ is a chart of $(X,[\mathcal{A}])$ with $x \in U$ and $\varphi(x) = 0$, then $\varphi(U \cap B_k X) = \varphi(U) \cap Q^0$.

The reader can study, for C^r_W-manifolds with corners, the *general properties* (when they have meaning) mentioned in **4.3, ...,4.12,** for C^r-manifolds with corners.

Manifolds with corners modeled on convenient real vector spaces

4.14 Let X be a non-void set. We say that $(U,\varphi,(E,Q))$ is a chart on X if: U is a subset of X, E is a convenient space (i.e., a Hlcrtvs which is Mackey complete (**3.30**)), Q is a quadrant in E with $int_{C^\infty T}(Q) \neq \varnothing$ (where, T is the topology of E), $\varphi : U \to Q$ is an injective map, and $\varphi(U)$ is a C^∞-open subset of Q (note that in this case the C^∞-topology of Q (**2.17**) is $(C^\infty T)|_Q$).

Recall that (**2.19**) $index(Q)$ is finite,Q^0 is closed in (E, T_{born}) and $int_{T_{born}}(Q)$ is non-void. Consequently if $index(Q) = n \in \mathbb{N}$, there exists a linearly independent system $\Lambda = \{\lambda_1, ..., \lambda_n\}$ of elements of $LB_T(E,\mathbb{R})$ such that $Q = E^+_\Lambda$, (hence $Q^0 = E^0_\Lambda$, Q and Q^0 are closed in (E, T_{born}), and $int_{C^\infty T}(Q) = int_{T_{born}}(Q)$).

Let $(U,\varphi,(E,Q))$, $(U',\varphi',(E',Q'))$ be charts on X. We say that they are smooth-compatible (*sth*-compatible) if: $\varphi(U \cap U')$ and $\varphi'(U \cap U')$ are C^∞-open subsets of Q and Q', respectively, and $\varphi'\varphi^{-1} : \varphi(U \cap U') \to \varphi'(U \cap U')$, $\varphi\varphi'^{-1} : \varphi'(U \cap U') \to \varphi(U \cap U')$ are smooth maps (**3.33**).

A collection $\mathcal{A}$ of charts on X is called a smooth-atlas on X (*sth*-atlas) if the domains of the charts of $\mathcal{A}$ cover X and any two of them are *sth*-compatible.

Two *sth*-atlases $\mathcal{A}$, $\mathcal{A}'$ on X are called smooth-equivalent (*sth*-equivalent) if $\mathcal{A} \cup \mathcal{A}'$ is a *sth*-atlas on X. This binary relation is an equivalence relation over the *sth*-atlases on X.

If $\mathcal{A}$ is a *sth*-atlas on X, the equivalence class $[\mathcal{A}]$ is called *sth*-structure on X and the pair $(X,[\mathcal{A}])$ is called *sth*-manifold with corners (or *sth*-manifold).

Remark For all Hlcrtvs (E, T) we have: $C^{\infty}(T_{born}) = C^{\infty}T$ (**2.17**), (E, T_{born}) is a bornological Hlcrtvs (**2.16**), and: *(E, T) is convenient if and only if (E, T_{born}) is convenient* (Proposition 2 (4) of **3.30**). Moreover, if (E, T) is a convient vector space, then the identity map, $1_E : (E, T) \to (E, T_{born})$, is a smooth diffeomorphism (**3.35**). In this way, we can consider that the *sth*-manifolds are modeled on bornological convenient vector spaces (see **2.19**).

If we deal with the class of Fréchet spaces (which are convenient spaces (Proposition 2 of **3.30**)), then we have:

(1) $C^{\infty}T = T_{born} = T$ for all Fréchet space (hence a Hlcrtvs) (E, T), (**2.17).**

(2) Let E, F be Fréchet spaces, Q a quadrant of E with $int(Q) \neq \varnothing$ (hence $int_{C^{\infty}E}(Q) \neq \varnothing$), U an open subset of Q (which is equivalent to: U is a C^{∞}-open subset of Q), and $f : U \to F$ a map. Then f is a smooth map (Definition of **3.33)** if and only if f is a C^{∞}_{W}-map on U (**3.19, 3.25, 3.28**).

(3) The constructions (of manifolds) developed in **4.13** (C^{∞}-class) and **4.14** are indistinguishable.

If $(X, [\mathcal{A}])$ is a *sth*-manifold, the set $\{U \subset X | U\ is\ a\ domain\ of\ a\ chart\ of\ (X, [\mathcal{A}])\}$ is a basis of a topology $T_{[\mathcal{A}]}$ on X (called the natural topology induced by $[\mathcal{A}]$).

If $(X, [\mathcal{A}])$ is a *sth*-manifold, the *sth*-curves $c : \mathbb{R} \to X$ are defined in a natural way (for each chart $\varphi \circ c$ is a *sth*-curve) and it can be proved that $T_{[\mathcal{A}]}$ is the final topology in X respect to the family $\{c : \mathbb{R} \to X | c\ is\ a\ sth\text{-}curve\}$. Moreover, the topology of a *sth*-manifold verifies the T_1 axiom (the C^{∞}-topology of a Hlcrtvs is Hausdorff), and finally a *sth*-manifold fulfils the first axiom of countability if and only if it is a Fréchet manifold (see the proof of 4.19 in [18]).

Theorem *Let $(X, [\mathcal{A}])$ be a sth-manifold, $x \in X$ and $c = (U, \varphi, (E, Q))$, $c' = (U', \varphi', (E', Q'))$ charts of $(X, [\mathcal{A}])$with $x \in U \cap U'$. Then, by* **3.36**, $index(\varphi(x)) = index(\varphi'(x))$ $(= ind(x))$. The non-negative integer number $ind(x)$ will be called index of x in $(X, [\mathcal{A}])$.

Let $(X, [\mathcal{A}])$ be a *sth*-manifold. Then:

(*a*) For all $k \in \mathbb{N} \cup \{0\}$, the set $\{x \in X |\ ind(x) \geqslant k\}$ $(= \partial^k X)$ is called k-boundary of $(X, [\mathcal{A}])$.

(*b*) For all $k \in \mathbb{N} \cup \{0\}$, the set $\{x \in X |\ ind(x) = k\}$ is denoted by $B_k X$.

If the quadrants are omitted, the *sth*-manifolds that we obtain were already defined in [18], p. 264.

4.15 The smooth maps between *sth*-manifolds are defined as usually (by localization) and one has that: *$f : X \to X'$ is a smooth map if and only if f maps sth-curves in X to sth-curves in X'.*

The reader can study, for *sth*-manifolds with corners, the *general properties* (when they have meaning) mentioned in **4.4,...,4.12,** for C^r-manifolds with corners.

5 Manifolds with generalized boundary

The C^r_S-maps, introduced in **3.14**, permit us to establish the foundations of the manifolds with generalized boundary. This theory has been developed by G. Graham in [13].

5.1 Let X be a non-void set. We say that (U, φ, E) is a chart on X if: U is a subset of X, E is a normable rtvs, $\varphi : U \to E$ is an injective map and $\varphi(U)$ is an admissible subset of E (**3.11**).

Let (U, φ, E), (U', φ', E') be charts on X.We say that they are C^r_S-compatible ($r \in \mathbb{N} \cup \{\infty\}$) if: $\varphi(U \cap U')$ is open in $\varphi(U)$ (hence admissible in E), $\varphi'(U \cap U')$ is open in $\varphi'(U')$ (hence admissible in E') and the maps $\varphi'\varphi^{-1} : \varphi(U \cap U') \to \varphi'(U \cap U')$, $\varphi\varphi'^{-1} : \varphi'(U \cap U') \to \varphi(U \cap U')$ are C^r_S-maps (**3.14**), (hence they are inverse homeomorphisms).

A set $\mathcal{A}$ of charts on X is called a C^r_S-atlas on X, if the domains of the charts of $\mathcal{A}$ cover X and any two of them are C^r_S-compatible. Two C^r_S-atlases $\mathcal{A}$, $\mathcal{A}'$ on X are called C^r_S-equivalent if $\mathcal{A} \cup \mathcal{A}'$ is an C^r_S-atlas on X. This binary relation is an equivalence relation over the C^r_S-atlases on X (use the results of **3.16**).

If $\mathcal{A}$ is a C^r_S-atlas on X, the equivalence class $[\mathcal{A}]$ is called C^r_S-structure on X and the pair $(X, [\mathcal{A}])$ is called C^r_S-manifold with generalized boundary (or C^r_S-manifold).

Let $(X, [\mathcal{A}])$ be a C^r_S-manifold. Then the set $\{U \subset X | U\ is\ the\ domain\ of\ a\ chart\ of\ (X, [\mathcal{A}])\}$ is a basis of a topology on X, which will be denoted by $T_{[\mathcal{A}]}$ and will be called natural topology induced by $[\mathcal{A}]$ on X.

Now we deal with C^r_S-manifolds modeled over real Banach spaces. In this case we can define the generalized boundary as follows: Let $(X, [\mathcal{A}])$ be a C^r_S-manifold, x a point of X, and (U, φ, E), (U', φ', E') charts of $(X, [\mathcal{A}])$ with $x \in U \cap U'$. Then, by **3.18**, $\varphi(x) \in int(\varphi(U \cap U'))$ if and only if $\varphi'(x) \in int(\varphi'(U \cap U'))$.

Definition Let $(X, [\mathcal{A}])$ be a C^r_S-manifold and x a point of X. We say that x is interior point of $(X, [\mathcal{A}])$ if there exists a chart (U, φ, E) of $(X, [\mathcal{A}])$ such that $x \in U$ and $\varphi(x) \in int(\varphi(U))$. The set of these points will be denoted by $Int(X)$ and will be called interior of X. The set $X - Int(X)$ will be denoted by ∂X and will be called boundary of X. The points of ∂X will be called boundary points of the manifold $(X, [\mathcal{A}])$.

Proposition *Let $(X, [\mathcal{A}])$ be a C^r_S-manifold. Then we have that $Int(X)$ is open and dense in $(X, T_{[\mathcal{A}]})$ and ∂X is closed in this topological space. Moreover, $Int(X)$ becomes, in a natural way, C^r-manifold without boundary (**3.15, 4.1**).*

The boundary of a C^r_S-manifold X can be a rare subset of X and, in general, we can not define a differentiable structure on it.

Example If B is any subset of $\{y \in \mathbb{R}^2 | \, \|y\| = 1\}$, then $X = B \cup \{y \in \mathbb{R}^2 | \, \|y\| < 1\}$ is a C^∞_S-manifold with generalized boundary such that $\partial X = B$ and $Int(X) = \{y \in \mathbb{R}^2 | \, \|y\| < 1\}$ (note that X is an admissible subset of $\mathbb{R}^2$).

5.2 The C^r_S-maps between C^r_S-manifolds are defined as usually by localization.

Proposition *Let $f : X \to X'$ be a C^r_S-diffeomorphism between C^r_S-manifolds modeled over Banach spaces. Then, $f(\partial X) = \partial X'$, $f(Int(X)) = Int(X')$, and $f|_{Int(X)} : Int(X) \to Int(X')$ is a C^r-diffeomorphism (between C^r-manifolds without boundary).*

Bibliography

[1] S. Armas-Gómez, J. Margalef-Roig, E. Outerelo-Domínguez and E. Padrón Fernández: Embedding of a Urysohn Differentiable Manifold with corners in a real Banach space *In: The Proceedings of the winter school Geometry and Physics Srni-January 1991* (Supplemento ai Rendiconti del Circolo Matematico di Palermo, Serie II, numero 30, 1993) 143–152

[2] N. Bourbaki: *Variétés différentielles et analytiques. Fascicule de résultats* (Paris: Hermann, 1967)

[3] N. Bourbaki: *Théorie des ensembles* (Paris: Hermann, 1970)

[4] J. Cerf: Topologie de certains espaces de plongements *Bull. Soc. Math. France* **89** (1961) 227–380

[5] J. Dieudonné: *Foundations of Modern Analysis* (New York and London: Academic Press, 1960)

[6] S. Donaldson: An application of Gauge Theory to four-dimensional Topology *J. Differential Geometry* **18** (1983) 279–315

[7] A. Douady: Variétés à bord anguleux et voisinages tubulaires, théorèmes d'isotopie et de recollement *In: Séminaire Henri Cartan 1961/62, exposé 1* (Secrétariat Math., Paris, 1961/1962)

[8] J. Eells: On the Geometry of Function Spaces *In: Symp. Int. Algebraic Topology* (Mexico, 1956, 1958) 303–308

[9] J. Eells and K. D. Elworthy: Open embeddings of certain Banach manifolds *Ann. of Math.* **91** (1970) 465–485

[10] A. Frölicher and A. Kriegl: *Linear Spaces and Differentiation Theory* (Chichester and New York: John Wiley & Sons, 1988)

[11] H. Glöckner: Infinite-Dimensional Lie Groups without Completeness Restrictions *In: Geometry and Analysis of Finite and Infinite-Dimensional Lie Groups, A. Strasburger et Al. (Eds)* (Warsaw, Banach Center Publications, Vol. 55, 2002)

[12] R. E. Gompf: An infinite set of exotic $\mathbb{R}^{4\prime}$s *J. Differential Geometry* **21** (1985) 283–300

[13] G. Graham: Differentiable Manifolds with Generalized Boundary *Czechoslovak Math. J.* **34** (1984) 46–63

[14] R. Hamilton: The inverse function theorem of Nash and Moser *Bull. Amer. Math. Soc.* **7** (1982) 65–222

[15] K. Jänich: On the Classification of O(n)-Manifolds *Math. Annalen* **176** (1968) 53–76

[16] H. Jarchow: *Locally Convex Spaces* (Stuttgart: B. G. Teubner, 1981)

[17] G. Köthe: *Topological Vector Spaces I* (Berlin: Springer-Verlag, 1969)

[18] A. Kriegl and P. W. Michor: *The Convenient Setting of Global Analysis* (Providence, R.I.: A.M.S., 1997)

[19] J. M. Lee: *Introduction to Smooth Manifolds* (Graduate Texts in Mathematics, Vol. 218, Berlin: Springer-Verlag, 2003)

[20] J. Margalef-Roig, E. Outerelo-Domínguez and E. Padrón-Fernández: On Submersions preserving the boundary and Quotient Manifolds *In: Proceedings of the International Conference on Differential Geometry and its applications, J. Janyška, D. Krupka (Eds.)* (World Scientific, 1990) 119–128

[21] J. Margalef-Roig and E. Outerelo Dominguez: *Differential Topology* (Amsterdam: North-Holland, 1992)

[22] R. B. Melrose: Differential Analysis on Manifolds with Corners *Preprint* rbm@math.mit.edu (1993)

[23] J. Milnor: On Manifolds Homeomorphic to the 7-sphere *Ann. Math.* **64** (1956) 399–405

[24] J. Milnor: On infinite-dimensional Lie groups *In: Relativité, Groups et Topologie II, B. DeWitt and R. Stora (Eds.)* (Amsterdam: North-Holland, 1983) 1007–1057

[25] M. Morse: *Calculus of Variations in the Large* (Amer. Math. Soc. Colloq. Publ. 18, Providence, RI, 1934)

[26] B. Riemann: *Ueber die Hypothesen welche der Geometrie zu grunde liegen* (Habilitations vortrag 10. Juni 1854, Gesammelte werke (H. Weber, Ed.), Leipzig: Teubner, 1876)

[27] H. Toruńczyk: Characterizing Hilbert space topology *Fundamenta Mathematicae* **CXI** (1981) 247–262

[28] Ch. Wockel: Smooth Extensions and Spaces of Smooth and Holomorphic Mappings *Preprint* `arXiv:math.DG/0511064` v1, 1–9 (2005)

J. Margalef-Roig

Consejo Superior de Investigaciones Científicas,
Instituto de Matemáticas y Física Fundamental,
Calle Serrano 123, 28006 Madrid, Spain

E-mail: `margalef@imaff.cfmac.csic.es`

E. Outerelo Domínguez

Universidad Complutense de Madrid, Facultad de Ciencias Matemáticas,
Departamento de Geometría y Topología, Ciudad Universitaria,
28040. Madrid, Spain

E-mail: `outerelo@mat.ucm.es`

This page intentionally left blank

Jet manifolds and natural bundles[1]

D. J. Saunders

Contents

1 Introduction

The two concepts forming the subject of this short article were both originally formulated in the nineteen fifties, at a time when the importance of fibre bundles was becoming clear. The idea of a *jet* appeared in the work of Charles Ehresmann [7, 8, 9, 10, 11]; this is an object which encapsulates the values taken at a point by a map and its derivatives up to some given order. Jets are useful as tools to provide coordinate-free ways of describing constructions such as differential equations, and are particularly convenient where a space of maps, which would normally be infinite-dimensional, can be replaced by a finite-dimensional space of jets. In Section 2 we give the basic definitions and describe various manifolds of jets, paying particular attention to the geometrical structures which are associated with these manifolds.

The two following sections cover two major applications where the language of jets can profitably be employed. In Section 3 we describe differential equations and, in particular, look at the problem of when systems of partial differential equations have no solutions because cross-differentiating gives rise to inconsistencies; in Section 4 we express problems in the calculus of variations using this language, and see in particular how the classical Euler-Lagrange equations make natural use of jet coordinates.

The concept of a *natural bundle* also uses jets, although the underlying ideas go back to the early descriptions of vectors and tensors as 'objects transforming in certain ways'. With the development of the the theory of fibre bundles, these objects were seen to be elements of certain bundles associated to the frame bundle of a manifold. There were,

[1] This work was supported by grant no. 201/06/0922 for Global Analysis and its Applications from the Czech Science Foundation.

nevertheless, other significant constructions which did not fit into this pattern, the most important of these being connections.

A more general approach to these ideas appeared in the work of Nijenhuis [29], where a functorial method of associating bundles with manifolds was used, and we discuss this in Section 5. By insisting that the association is 'local' in a specific way, it can be shown that a natural bundle is simply one which is associated to a higher-order frame bundle, in the same way that a bundle of tensors is associated to the (first-order) frame bundle. The language of jets is again appropriate here, as higher-order frames are just jets of a particular kind, and the structure group of a higher-order frame bundle is a jet group. Finally, we make a further generalisation to *gauge-natural bundles*, where the functorial correspondence associates the new bundles with principal bundles rather than manifolds.

The scope of this article is quite broad, and it is not intended to provide a comprehensive coverage of all the topics mentioned. It is instead introductory, with several of topics being discussed in more depth elsewhere in this Handbook (see, in particular, [13, 19, 21, 22, 43]; a detailed discussion of natural bundles may be found in [18].

2 Jets

2.1 What is a jet?

A jet is an object which can be constructed from a map (between differentiable manifolds) and a point in the domain of the map; it is, essentially, the Taylor polynomial of the map about that point.

Definition 2.1 If $\phi, \hat{\phi} : M \to N$ are C^p maps between C^p manifolds $(1 \leq p \leq \infty)$ then they are *k-equivalent at $x \in M$* $(k \leq p,\ k < \infty)$ if $\hat{\phi}(x) = \phi(x)$ and, for every C^p curve $\gamma : (a, b) \to M$ with $0 \in (a, b) \subset \mathbb{R}$ and $\gamma(0) = x$, and for every C^p function $f : N \to \mathbb{R}$,

$$\left.\frac{d^r(f \circ \hat{\phi} \circ \gamma)}{dt^r}\right|_0 = \left.\frac{d^r(f \circ \phi \circ \gamma)}{dt^r}\right|_0$$

whenever $1 \leq r \leq k$. The equivalence class containing the map ϕ is called the *k-jet of ϕ at $x \in M$* and is denoted $j^k_x\phi$. The number k is called the *order* of the jet. The definition of an *infinite jet* is similar, and will be given in subsection 2.3.

Some remarks about this definition are needed. First, it is clear that the equivalence relation depends, not on the maps themselves, but on their germs at x; we would get the same result by considering maps defined only in neighbourhoods of x, because we could always extend such maps to be defined on the whole of M. (Of course this wouldn't work in the same way for real-analytic maps, or for holomorphic maps on complex manifolds, and there it would be more appropriate to concentrate on germs rather than maps.) And secondly, we haven't imposed any conditions about the manifolds being Hausdorff, paracompact or finite-dimensional, although normally all three conditions hold: indeed, one of the reasons for studying jets is that we can often avoid the complexities of infinite-dimensional manifold theory. We shall henceforth assume that these conditions do hold. We shall also restrict attention to manifolds and maps of class C^∞; the reason for this will be explained below.

The definition of a jet is often given in chart form, where it is common to use multi-index notation. A multi-index I is an m-vector of non-negative integers with i-th element

$I(i)$, and the length $|I|$ of the multi-index is $\sum_i I(i)$. We also write $I!$ for the factorial $\prod_i I(i)!$ and we let 1_i denote the multi-index $(0,\ldots,0,1,0,\ldots,0)$ with a single 1 in the i-th position. So for instance, if $I=(3,1,0,\ldots,0)$ then

$$\frac{\partial^{|I|}f}{\partial x^I}=\frac{\partial^4 f}{(\partial x^1)^3\partial x^2}\,,\qquad |I|=4\,,\qquad I!=3!\,1!\,0!\ldots0!=6\,.$$

Taking a chart on M with coordinates x^i around x, and another chart on N with coordinates u^α around $\phi(x)$, we say that $\hat{\phi}$ is k-equivalent to ϕ at x if $\hat{\phi}(x)=\phi(x)$ and also

$$\left.\frac{\partial^{|I|}\hat{\phi}^\alpha}{\partial x^I}\right|_x=\left.\frac{\partial^{|I|}\phi^\alpha}{\partial x^I}\right|_x$$

for every multi-index I with $1\le|I|\le k$. Given this alternative definition in one chart, the chain rule then shows that the same condition holds for any chart, and the definition of partial derivatives then gives us our original definition back again. In any chart there is exactly one representative of the equivalence class that is a polynomial of degree k, and that is indeed the Taylor polynomial of the map.

2.2 Manifolds of jets

We often use manifolds of jets, and several of these are already familiar in the first-order case. First-order jets at zero of maps from $\mathbb{R}$ to N are tangent vectors, and the set of all of them forms the tangent bundle TN; the higher-order jets form the higher-order tangent bundles T^kN. Note that, for instance, T^2N is *not* the repeated tangent bundle TTN. It is in fact a proper subset; we shall return to this point, and its generalizations, in subsection 2.4. On the other hand, first-order jets of maps from M to $\mathbb{R}$ are cotangent vectors, and the set of all of them, at all points of M, forms the cotangent bundle T^*M; higher-order cotangent vectors are used rather less frequently.

In a similar way, first-order jets at zero of *non-degenerate* maps from $\mathbb{R}^n$ to N, where $n=\dim N$, are frames in N, where by non-degenerate we mean maps with non-vanishing Jacobian at zero; the set of all of these jets forms the frame bundle $\mathcal{F}N$. The set of first-order jets of *all* maps from $\mathbb{R}^m$ to N (with $m\le n$) is perhaps less familiar: it is the bundle $T_{(m)}N$ of m-velocities in N; the higher-order jets form the bundles of higher-order m-velocities $T^k_{(m)}N$. Each of these has a sub-bundle containing the non-degenerate m-velocities $\mathcal{F}^k_{(m)}N$, where in particular $\mathcal{F}^1_{(n)}N$ is just the frame bundle $\mathcal{F}N$; we also write $\mathcal{F}^kN$ for the bundle $\mathcal{F}^k_{(n)}N$ of k-th order frames.

The example of the tangent bundle TN will explain why we insist on smooth, that is C^∞, manifolds and maps, when our definition of the k-jet applies equally to a C^p map when $p\ge k$ is finite. Let $\mathbb{R}[X]$ be the real algebra of polynomials in an indeterminate X, and let (X^2) denote the ideal generated by the monomial X^2. The quotient $D=\mathbb{R}[X]/(X^2)$ is a local algebra: that is, it has a unique maximal ideal (ε), where ε is the image of the monomial X under the projection $\mathbb{R}[X]\to D$. Elements of D are often called *dual numbers*, and were studied in detail by André Weil [44]. If N is a smooth manifold with an algebra of smooth real-valued functions $C^\infty(N)$ then a tangent vector $v\in T_yN$ may be identified with an algebra homomorphism $\phi_v\in C^\infty(N)\to D$ satisfying

$\phi_v(f) = f(y) \mod \varepsilon$, so that if we write $\phi_v(f) = f_0 + f_1\varepsilon$ then

$$\phi_v(fg) = f_0 g_0 + (f_0 g_1 + g_0 f_1)\varepsilon\,,$$

and the coefficient of ε incorporates Leibniz rule for derivations. Other manifolds of jets may be generated algebraically in a similar way by using other local algebras; see, in particular, the article by I. Kolář in this volume [17]. But the algebraic construction doesn't work if the manifold isn't smooth, because in such cases there are derivations which cannot be obtained by this identification. For instance, it can been shown that the space of derivations at a given point y of a finite-dimensional C^p manifold N (where p is finite) is an infinite dimensional space, whereas the tangent space $T_y N$ has the same finite dimension as N [28].

We return to the case of smooth manifolds and maps. One important type of structure arises when we start with a fibred manifold $\pi : E \to M$, so that π is a surjective submersion. This might be a fibre bundle, or it might be a locally trivial bundle with no particular specification about the group in which the transition functions take their values. But neither of these conditions is a requirement; all that's needed is that we can take local sections $\phi : U \to E$ (where $U \subset M$) and consider their k-jets at points of M. Details of the constructions may be found in [36]. We write, as before $j^k_x\phi$ for the k-jet, and put $J^k\pi$ as the set of all these k-jets; we define the source and target projections $\pi_k : J^k\pi \to M$, $\pi_{k,0} : J^k\pi \to E$ by

$$\pi_k(j^k_x\phi) = x\,, \qquad \pi_{k,0}(j^k_x\phi) = \phi(x)\,.$$

The set $J^k\pi$ is a C^∞ manifold; if we take fibred charts on E with coordinates of the form (x^i, u^α), where the projected chart on M has coordinates x^i, then the coordinates on the preimage of the chart in $J^k\pi$ are $(x^i, u^\alpha, u^\alpha_I)$ where

$$u^\alpha_I(j^k_x\phi) = \left.\frac{\partial^{|I|}\phi}{\partial x^I}\right|_x .$$

(If $m = 1$ then there is only one multi-index of a given length, and so we tend to write the jet coordinates as $u^\alpha_{(r)}$ rather than u^α_I, where $r = |I|$.) These coordinates define a C^∞ manifold structure on $J^k\pi$ of dimension $n(1 + (m + k)!/m!k!)$.

Note that we could use ordinary indices instead of multi-indices as subscripts of the jet coordinates, on the understanding that $u^\alpha_{i_1\cdots i_r}$ was symmetric in its subscripts. Changing from a sum over multi-indices to a sum over ordinary indices introduces a numerical factor to compensate for repetition due to the symmetry: we have, for any quantity $\Phi(I)$ depending on a multi-index I, the relationship

$$\sum_{|I|=r} \frac{|I|!}{I!}\Phi(I) = \sum_{i_1=1}^{m} \cdots \sum_{i_r=1}^{m} \Phi(1_{i_1} + \cdots + 1_{i_r})\,,$$

where the integer $|I|!/I!$ is called the *weight* of the multi-index I.

With this manifold structure on $J^k\pi$ the source and target maps become smooth surjective submersions, and in fact the target map defines a bundle structure. If we let $\pi_{l,k} : J^l\pi \to J^k\pi$ be given by $\pi_{l,k}(j^l_x\phi) = j^k_x\phi$ for $l > k$, and identify the zero-jet manifold with E, then, in particular, each $\pi_{k,k-1}$ defines an *affine* bundle structure. The associated vector bundle is $S^k T^* M \otimes_{J^{k-1}\pi} V\pi$ containing $(k+1)$-fold symmetric cotangent vectors from the base, tensored with vertical tangent vectors from the total space, all pulled back to $J^{k-1}\pi$. The action of such a tensor on a jet is evident in coordinates: if

$$\xi = \sum_{|I|=k} \xi^\alpha_I dx^I \otimes \frac{\partial}{\partial u^\alpha} \in S^k T^* M \otimes_{J^{k-1}\pi} V\pi\big|_{j^{k-1}_x \phi}$$

then

$$u^\alpha_I(\xi \cdot j^k_x \phi) = \xi^\alpha_I + u^\alpha_I(j^k_x \phi) \qquad (|I| = k)\,.$$

Another important type of structure arises when we have a single manifold E of dimension $(m+n)$ without a fibration over a base manifold M. We can then obtain 'jets of immersed submanifolds'. As a jet is a purely local construction, the only invariant of the immersed manifold is its dimension, so we may assume that the immersed manifold is $\mathbb{R}^m$. We therefore start with a map $\sigma : \mathbb{R}^m \to E$ that is non-degenerate at the origin, and take its k-jet $j^k_0 \sigma \in \mathcal{F}^k_{(m)} E$; but we wish to ignore the parametrization of the submanifold, so we must take a further equivalence by setting $j^k_0 \sigma \sim j^k_0(\sigma \circ f)$ for any diffeomorphism f of $\mathbb{R}^m$ satisfing $f(0) = 0$. We call the resulting set $J^k(E, m)$, the set of k-jets of m-dimensional submanifolds of E, or sometimes the set of m-dimensional k-th order contact elements of E. In the case $k = 1$ this is the Grassmannian bundle of m-planes in E, so by extension $J^k(E, m)$ is also called the bundle of k-th order Grassmannians, or the bundle of k-th order contact elements. This set, too, is a C^∞ manifold, of the same dimension as $J^k\pi$, and is the latter's 'projective completion'. We use the same coordinates on $J^k(E, m)$ as on $J^k\pi$ because, given an immersion, we can always define a local fibration of E over $\mathbb{R}^m$ such that locally the immersion is a section of the fibration. The simplest example of this is when $m = k = 1$; the (affine) jet bundle may then be considered as an affine sub-bundle of the tangent bundle TE, whereas the bundle of jets of submanifolds may be identified with the projective tangent bundle PTE. As with the case of jets of sections, the maps $\pi_{k+1,k} : J^{k+1}(E, m) \to J^k(E, m)$ define affine bundles, provided $k \geq 1$; the map $\pi_{1,0} : J^1(E, m) \to E$ does not define an affine bundle but, as described above, defines a Grassmannian bundle, of which the projective tangent bundle is a special case. In this article, we shall tend to concentrate on the affine sub-bundle $J^k\pi$ rather than the bundle $J^k(E, m)$ of immersed submanifolds; many constructions involving the latter are described in an article by D. R. Grigore in this volume [13].

Incidentally, when constructing $J^k(E, m)$, we could have chosen to specify that the further equivalence involved orientation-preserving diffeomorphisms f. This would have given us the bundle of *oriented* jets of submanifolds $J^k_+(E, m)$; in the simple case $m = k = 1$ we would obtain the sphere bundle, as studied in Finsler geometry, instead of the projective tangent bundle.

Yet another variant of the jet construction arises when we consider jets of local diffeomorphisms: that is, maps defined globally with the property that each point has a neighbourhood where the restriction is a diffeomorphism onto its image. If we consider jets of local diffeomorphisms from a manifold to itself, and consider only those maps taking a given fixed point to itself, then we might as well look at jets of local diffeomorphisms $f : \mathbb{R}^m \to \mathbb{R}^m$ satisfying $f(0) = 0$. These are the *invertible jets*, and they form a group, the *jet group* L^k_m, under the operation of composition of jets: for the case $k = 1$ the group L^1_m

is just the general linear group $\mathrm{GL}(m,\mathbb{R})$. If we restrict further to orientation-preserving diffeomorphisms then we obtain the oriented jet group L^k_{m+}. From the construction of $J^k(E,m)$ it is evident that

$$\mathcal{F}^k_{(m)}E \to J^k(E,m)$$

is a principal L^k_m-bundle, and similarly that

$$\mathcal{F}^k_{(m)}E \to J^k_+(E,m)$$

is a principal L^k_{m+}-bundle. Special cases of these bundles are the k-th order frame bundles

$$\mathcal{F}^k E = \mathcal{F}^k_{(m+n)}E \to J^k(E,m+n) = E\,.$$

The algebraic structure of the jet groups is important. As usual we denote by $\pi_{l,k} : L^k_m \to L^l_m$ the natural projection, and this is clearly a group homomorphism. We can then write L^k_m as a semidirect product

$$L^k_m \cong \mathrm{GL}(m,\mathbb{R}) \rtimes \ker \pi_{k,1}\,,$$

where $A \in \mathrm{GL}(m,\mathbb{R})$ is regarded as a non-degenerate linear map $\mathbb{R}^m \to \mathbb{R}^m$, and where the action of $\mathrm{GL}(m,\mathbb{R})$ on $\ker\pi_{k,1}$ is left multiplication by $j^k_0 A$ ([18], Proposition 13.4).

In all these cases, the operation of taking jets is a covariant functor. For jets of sections, it is a functor on the category of fibred manifolds and fibred maps whose projections are local diffeomorphisms on the base. For jets of submanifolds, it is a functor on the category of manifolds and smooth maps. The action of the functor on a map will be described shortly, when we have considered prolongations.

2.3 Infinite jets

For some purposes it is convenient to use infinite jets. The definition is the obvious extension of the definition of a finite-order jet, so that two maps are ∞-equivalent when they are k-equivalent for each k.

Definition 2.2 If $f, g : M \to N$ are C^∞ maps between C^∞ manifolds then they are *∞-equivalent at $x \in M$* if $f(x) = g(x)$ and, for every C^∞ curve $\gamma : (a,b) \to M$ with $0 \in (a,b) \subset \mathbb{R}$ and $\gamma(0) = x$, and for every C^∞ function $\phi : N \to \mathbb{R}$,

$$\left.\frac{d^r(\phi\circ f\circ\gamma)}{dt^r}\right|_0 = \left.\frac{d^r(\phi\circ g\circ\gamma)}{dt^r}\right|_0$$

whenever $r \geq 1$. The equivalence class containing the map f is called the *∞-jet of f at $x \in M$* and is denoted $j^\infty_x f$.

As with finite-order jets, we can construct manifolds of infinite jets; but some care is needed, because these will be infinite-dimensional manifolds.

The first observation here is that infinite jet manifolds will be Fréchet manifolds, rather than Banach manifolds. The model vector space will be the space $\mathbb{R}[[X_1,\ldots,X_m]]$ of formal power series in m indeterminates, perhaps taking a product with a finite-dimensional space depending upon the particular manifold being considered. The natural topology to impose on this space is the inverse limit topology using the spaces $\mathbb{R}_k[X_1, \ldots, X_m]$ of polynomials of total degree no greater than k and the sequence

$$\mathbb{R}_0[X_1, \ldots, X_m] \leftarrow \mathbb{R}_1[X_1, \ldots, X_m] \leftarrow \ldots \leftarrow \mathbb{R}_k[X_1, \ldots, X_m] \leftarrow \ldots$$

of projection maps which discard the terms of highest degree. The topology constructed in this way is locally convex, because it is defined by seminorms p_k obtained by taking the pull-backs of the norms on each polynomial space. It is metrizable, because there are countably many seminorms and so the function

$$d(x, y) = \sum_k \frac{1}{2^k} p_k(x - y)$$

is a translation-invariant pseudometric which generates the topology; the seminorms separate points, and so the space is Hausdorff and hence the pseudometric is a metric. And finally the space is complete, because if (x_n) is a Cauchy sequence in this metric then $(x_n|_k)$ is a Cauchy sequence in $\mathbb{R}_k[X_1, \ldots, X_m]$, where $x_n|_k$ denotes the truncation of the power series x_n to a polynomial of degree k. Each Cauchy sequence of polynomials tends to a limit polynomial $x|_k$ in the finite-dimensional polynomial space, and these limit polynomials determine a unique formal power series x which is easily seen to be the limit of (x_n) in the metric d.

Thus the space of formal power series satisfies the conditions for a Fréchet space.

Another technical detail to consider when defining a manifold structure concerns the natural extension of the finite-order jet charts with jet coordinates,

$$u^\alpha_I(j^\infty_x f) = \left.\frac{\partial^{|I|} f}{\partial x^I}\right|_x,$$

where now the multi-index I may have arbitrary length. We would like to know that every formal power series is indeed the Taylor series of some smooth function, even if the series converges only at the origin: the fact that this is the case is Borel's Theorem. This is a local theorem, and so its proof may be carried out on vector spaces ([42], Theorem 38.1). We consider the Taylor map from the Fréchet space of smooth functions to the Fréchet space of formal power series, and the transpose of this map between the dual spaces. The topological dual of the space of formal power series is the space of polynomials; and the topological dual of the space of smooth functions is the space of Schwarzian distributions with compact support. It may be shown in the general case of continuous linear maps between Fréchet spaces that if the transpose map is injective and has a closed image under the weak topology then the original map is surjective ([42], Theorem 37.2); in this particular case it is clear that the transpose is injective, and it may also be shown that the image of the transpose is the space of finite linear combinations of derivatives of Dirac delta-distributions, and this is closed under the weak topology ([42], Theorem 24.6). Thus the Taylor map, which is evidently linear and is also continuous, must be surjective.

We therefore see that we may impose the structure of a Fréchet manifold upon suitable sets of infinite jets; we shall concentrate on the situation of a fibred manifold $\pi : E \to M$ and its infinite jet manifold $J^\infty\pi$. As with all infinite-dimensional manifolds, one needs to take care about the validity of the theorems being used. In particular, the proof for Banach spaces that an ordinary differential equation written in solved form has a local solution is not valid in the absence of a norm, and so we cannot guarantee that a vector field on an infinite jet manifold will have a flow: there is, indeed, an important structural vector field which definitely does not have a flow.

One final observation here about infinite jet manifolds concerns their vector fields and differential forms. A tangent vector at any point may be considered as an element of the model Fréchet space, a formal power series, and so may have infinitely many 'components'; the same obviously applies to a vector field. On the other hand, a cotangent vector at that point may be regarded as an element of the dual of the model space, namely a polynomial, and so will have a finite order: it will always be a pull-back from a cotangent vector on a finite-order jet manifold. But a differential form, a section of the cotangent bundle or its wedge products, need not be a pull-back: at each point it will have a finite order, but these orders need not be bounded over the whole manifold. In many applications, for instance in the calculus of variations, such generality is not required, and so consideration is restricted to differential forms of globally finite order.

2.4 Prolongations and holonomic jets

One usually finds the words 'prolongation' and 'holonomic' associated with jets: the former just means differentiation, in a suitable sense, and the latter refers to an object that has been obtained by prolongation. In this subsection we shall concentrate on jets of sections of a fibration; similar considerations apply to other manifolds of jets.

The basic act of prolongation occurs in the construction of jets themselves. Starting with a local section $\phi : U \to E$, we obtain a jet $j^k_x\phi$ for each $x \in U$; the correspondence $U \to J^k\pi$ given by $x \mapsto j^k_x\phi$ is then a local section of $\pi_k : J^k\pi \to M$ which we denote $j^k\phi$ and call the *k-th prolongation of* ϕ. By construction the coordinate representation of $j^k\phi$ satisfies

$$u^\alpha_I \circ j^k\phi = \frac{\partial^{|I|}(u^\alpha \circ \phi)}{\partial x^I} \,,$$

from which it is clear that $j^k\phi$ is indeed smooth. Not every local section ψ of π_k is a prolongation: there is no particular reason why, in general, the coordinate representation of ψ should satisfy

$$u^\alpha_I \circ \psi = \frac{\partial^{|I|}(u^\alpha \circ \psi)}{\partial x^I} \,.$$

But if it does, we say that the local section ψ is *holonomic*.

We can use this idea to investigate some properties of manifolds of repeated jets. By construction, $\pi_k : J^k\pi \to M$ is a fibred manifold, and so we can take l-jets of its local sections $\psi : U \to J^k\pi$; the set of all these l-jets will form the manifold $J^l\pi_k$. For a holonomic section $\psi = j^k\phi$, the resulting l-jet would then be $j^l_x\psi = j^l_x(j^k\phi)$. Such an l-jet is called *holonomic*. Not every l-jet in $J^l\pi_k$ is holonomic, and the subset containing holonomic l-jets is a submanifold that may be identified with $J^{k+l}\pi$, essentially because of the commutativity of partial derivatives. In coordinate terms, the jet coordinates on $J^l\pi_k$ are $u^\alpha_{I,J}$, indexed with a double multi-index, where $|I| \le k$ and $|J| \le l$ (it is convenient here to include the cases where $|I| = 0$ or $|J| = 0$). In contrast, the jet coordinates on $J^{k+l}\pi$ are u^α_{I+J}, with a single, symmetrized, multi-index. Using this identification, we call $J^{k+l}\pi$ the *holonomic submanifold* of $J^l\pi_k$. There is also a larger submanifold of 'semi-holonomic jets' which we shall describe below in (2.6). In general, there is *no* canonically-defined projection from the repeated jet manifold $J^l\pi_k$ to the holonomic submanifold $J^{k+l}\pi$.

We can generalize this idea to prolong certain fibred maps between two fibred manifolds. If the manifolds are $\pi : E \to M$ and $\rho : H \to N$, and if $f : E \to H$ projects to a local diffeomorphism $\bar{f} : M \to N$ then the prolonged map $j^k f : J^k\pi \to J^k\rho$ is defined by

$$j^k f(j^k_x\phi) = j^k_{\bar{f}(x)}(f \circ \phi \circ \bar{f}^{-1}),$$

where $\bar{f}^{-1}$ denotes the inverse of $\bar{f}$ in a neighbourhood of $\bar{f}(x)$. In the special case where $N = M$ and $\bar{f} = \mathrm{id}_M$ then this formula simplifies considerably, to

$$j^k f(j^k_x\phi) = j^k_x(f \circ \phi).$$

The prolonged map $j^k f$ is fibred over both f and $\bar{f}$, and in fact the correspondence between fibred manifolds $(E, \pi, M) \mapsto (J^k\pi, \pi_k, M)$ and fibred maps $(f, \bar{f}) \mapsto (j^k f, \bar{f})$ is a functor on the category of fibred manifolds and fibred maps projecting to local difformorphisms. Note that if E and H are each fibred over two different base manifolds, and if f is thus fibred over two different maps, then the same symbol $j^k f$ would represent two different prolonged maps between two different pairs of jet manifolds. In these circumstances a more explicit notation such as $j^1(f, \bar{f})$ might be helpful.

In charts (x^i, u^α) on E and (y^a, v^A) on H, if the coordinate representation of f is (f^a, f^A) then the coordinate representation of the first prolongation $j^1 f$ is

$$v^A_a \circ j^1 f = \frac{df^A}{dx^i}\frac{\partial x^i}{\partial y^a}$$

where

$$\frac{df^A}{dx^i} = \frac{\partial f^A}{\partial x^i} + u^\alpha_i \frac{\partial f^A}{\partial u^\alpha}$$

is the *total derivative* of f^A; total derivative operators are described more geometrically in the next subsection. Similar formulæ may be found for the coordinate representation of higher prolongations, but a general formula is rather unwieldy unless $M = N$ and $\bar{f} = \mathrm{id}_M$, when the general formula is simply

$$v^A_I \circ j^k f = \frac{d^{|I|} f^A}{dx^I}.$$

We can also prolong vector fields on the total space E of a fibred manifold $\pi : E \to M$. If the vector field $X \in \mathfrak{X}(E)$ is projectable to $\bar{X} \in \mathfrak{X}(M)$ then the flow ψ_t of X is projectable to the flow $\bar{\psi}_t$ of $\bar{X}$. Thus ψ_t is (at least locally) a fibred map, and so may be prolonged to give a (local) fibred map $j^k\psi_t$ which will be the flow of a vector field $X^{(k)}$ on $J^k\pi$. In a chart, if

$$X = X^i \frac{\partial}{\partial x^i} + X^\alpha \frac{\partial}{\partial u^\alpha}$$

with the functions X^i constant on each fibre, then

$$X^{(k)} = X^i \frac{\partial}{\partial x^i} + \sum_{|I|=0}^{k} \left(\frac{d^{|I|} X^\alpha}{dx^I} - \sum_{\substack{J+K=I \\ J \neq 0}} \frac{I!}{J!\,K!} \frac{\partial^{|J|} X^j}{\partial x^J} u^\alpha_{K+1_j} \right) \frac{\partial}{\partial u^\alpha_I} \, ;$$

if X is vertical over M so that the projected flow $\bar{\psi}_t = \mathrm{id}_M$ then the coordinate formula is simply

$$X^{(k)} = \sum_{|I|=0}^{k} \frac{d^{|I|} X^\alpha}{dx^I} \frac{\partial}{\partial u^\alpha_I} \, .$$

In fact the restriction that X be projectable is not necessary. An arbitrary vector field on E may be prolonged to a vector field on $J^k\pi$, and in the coordinate formula above the partial derivatives of the functions X^i are just replaced by total derivatives. An intrinsic construction of $X^{(k)}$ in this more general case may be found in ([36], Definition 6.4.16) as a composition

$$X^{(k)} = r_k \circ j^k X \, ,$$

where $j^k X : J^k\pi \to J^k(\pi \circ \tau_E)$ is the prolongation of $X : E \to TE$ regarded as a map fibred over the identity on M, and $r_k : J^k(\pi \circ \tau_E) \to TJ^k\pi$ is a canonical map described in ([36], Definition 6.4.14). It is for this reason that we use the notation $X^{(k)}$ for the prolonged vector field; the symbol $j^k X$ used by some authors has for us a different meaning.

2.5 Contact forms and total derivatives

Every manifold of jets has a distinguished class of differential forms called *contact forms*. These forms capture, in an invariant way, the fact that some coordinates on a jet manifold are 'derivatives' of others. The discussion in this subsection will be given in terms of affine jet bundles, but similar (although more complicated) arguments apply to other types of jet manifold.

Definition 2.3 A differential form θ on a jet manifold $J^k\pi$ is called a contact form if, for every local section $\phi : U \to E$, $U \subset M$, we have

$$(j^k\phi)^*\theta = 0 \, .$$

A local basis for the contact 1-forms on $J^k\pi$ is given in coordinates by

$$\theta^\alpha_I = du^\alpha_I - u^\alpha_{I+1_i} dx^i \, , \qquad |I| \leq k-1 \, .$$

Note in particular that the contraction of a contact 1-form with a vector field on $J^k\pi$ vertical over $J^{k-1}\pi$ will be zero (in other words, a contact 1-form is *horizontal* over $J^{k-1}\pi$). But this need not be true for contact r-forms with $r \geq 2$; for instance

$$d\theta^\alpha_I = -du^\alpha_{I+1_i} \wedge dx^i$$

is a contact form which is not horizontal over $J^{k-1}\pi$.

A simple example of a contact 1-form arises when $k = 1$ and the fibre dimension of E over M is 1; there is then a single contact form $\theta = du - u_i dx^i$ which satisfies $(d\theta)^m \wedge \theta \neq 0$, so that in this case the $(2m+1)$-dimensional manifold $J^1\pi$ is a contact manifold in the classical sense.

It is clear from the coordinate formula that any 1-form on $J^k\pi$ horizontal over $J^{k-1}\pi$ may be written uniquely as the sum of a contact form and a form horizontal over M; in particular, this applies to the pull-back by $\pi_{k,k-1}$ of a 1-form on $J^{k-1}\pi$. We may extend this decomposition to r-forms by defining *s-contact* r-forms. We say that an r-form θ on $J^k\pi$ horizontal over $J^{k-1}\pi$ is 1-contact if it is contact and if, for every vector field Y on $J^k\pi$ vertical over M, the contraction $i_Y\theta$ is horizontal over M. We can then say that θ is s-contact if, for every vector field Y on $J^k\pi$ vertical over M, the contraction $i_Y\theta$ is $(s-1)$-contact. Given these definitions, every r-form θ on $J^l\pi$ horizontal over $J^{k-1}\pi$ may be written uniquely as

$$\theta = \theta_q + \theta_{q+1} + \ldots + \theta_r$$

where θ_s is s-contact and where $q = \max\{r - m, 0\}$.

We have used prolonged local sections $j^k\phi$ to characterise contact forms; we may conversely use contact forms to characterise prolonged sections.

Proposition 2.4 *If ψ is a local section of $\pi_k : J^k\pi \to M$ satisfying the condition that $\psi^*\theta = 0$ for every contact 1-form θ on $J^k\pi$ then there is a local section ϕ of $\pi : E \to M$ with $\psi = j^k\phi$.*

The proof is a straightforward computation in local coordinates. We have

$$0 = \psi^*(du^\alpha_I - u^\alpha_{I+1_i}dx^i) = d\psi^\alpha_I - \psi^\alpha_{I+1_i}dx^i$$

so that

$$\psi^\alpha_{I+1_i} = \frac{\partial\psi^\alpha_I}{\partial x^i}$$

and we may take $\phi = \pi_{k,0} \circ \psi$.

We now turn to the dual objects. There are two different points of view here; on the one hand, we may think of contact 1-forms as living on the jet manifold $J^k\pi$, and look at the vector distribution annihilated by the contact forms. This is called the *Cartan distribution* or, alternatively, the *contact distribution*, and denoted $C\pi_{k,k-1}$. It is not an integrable distribution in the sense of Frobenius, and in fact its maximal integral manifolds are of two types. By construction, the images of prolonged sections are integral manifolds, and they are spanned by vectors of the form

$$\frac{\partial}{\partial x^i} + \sum_{|I| \leq k-1} u^\alpha_{I+1_i}\frac{\partial}{\partial u^\alpha_I}\,. \tag{2.1}$$

But the fibres of the affine bundle $\pi_{k,k-1} : J^k\pi \to J^{k-1}\pi$ are also integral manifolds, and they are spanned by vectors of the form

$$\frac{\partial}{\partial u^\alpha_J} \qquad |J| = k\,.$$

It is clear that taking the Lie bracket of a vector field of one type with a vector field of the other type takes us outside the distribution.

Any symmetry (or infinitesimal symmetry) of a jet bundle which preserves the derivative relationship between the coordinates must also preserve the Cartan distribution, and in fact we have the following results.

Theorem 2.5 *Suppose that the fibre dimension n of $\pi : E \to M$ satisfies $n \geq 2$. If f is a diffeomorphism of $J^k\pi$ satisfying $f_*(C\pi_{k,k-1}) = C\pi_{k,k-1}$ then f is fibred over E. If X is a vector field on $J^k\pi$ such that $[X, Y]$ is in $C\pi_{k,k-1}$ whenever Y is in $C\pi_{k,k-1}$ then X is the prolongation of a vector field on E.*

Proofs of these results for the case $k = 1$ may be found in ([36], Theorems 4.5.12 and 4.5.15); the proofs in the general case are similar. Note that the results are false when $n = 1$: for instance, the map f given by

$$\begin{aligned} x^i \circ f &= u_i \\ u \circ f &= x^i u_i - u \\ u_i \circ f &= x^i \end{aligned}$$

is a diffeomorphism of $J^1\pi$ satisfying $f_*(C\pi_{1,0}) = (C\pi_{1,0})$ but which is not fibred over E.

Whereas the Cartan distribution arises by duality when we regard the contact 1-forms as being differential forms on $J^k\pi$, we may also consider the latter as 'differential forms along the map $\pi_{k,k-1}$'. With this interpretation, the dual objects are 'vector fields along $\pi_{k,k-1}$', namely maps X from $J^k\pi$ to $TJ^{k-1}\pi$ satisfying

$$X_{j^k_x\phi} \in T_{j^{k-1}_x\phi} J^{k-1}\pi\,,$$

or equivalently sections X of the pullback bundle $\pi^*_{k,k-1}(TJ^{k-1}\pi) \to J^k\pi$.

Definition 2.6 A *total derivative* is a section X of the pull-back bundle

$$\pi^*_{k,k-1}(TJ^{k-1}\pi) \to J^k\pi$$

which annihilates the contact 1-forms on $J^k\pi$.

A local basis for the total derivatives is given in coordinates by

$$\frac{\partial}{\partial x^i} + \sum_{|I|\leq k-1} u^\alpha_{I+1_i} \frac{\partial}{\partial u^\alpha_I}\,.$$

Although this is identical to the formula given above (2.1) for certain vector fields in the Cartan distribution, the total derivatives are conceptually quite different: in particular, they are *not* vector fields on a manifold, and so the concept of a flow does not arise. They may be used to differentiate functions, as the notation suggests; but whereas the functions to be differentiated will be defined on $J^{k-1}\pi$, the results will be defined on $J^k\pi$.

Both the set of contact forms and set of total derivativatives are evidently modules over the ring of functions on $J^k\pi$; but the information contained in these modules may be captured in a single tensorial object. It is clear that the property of being a contact 1-form is a pointwise property, and so we may define the sub-bundle of contact cotangent vectors $C^*(\pi_{k,k-1})$. Regarded as a sub-bundle of the cotangent bundle $T^*J^k\pi$ this is by definition the Cartan distribution; but regarded as a sub-bundle of the pull-back bundle $\pi^*_{k,k-1}(T^*J^{k-1}\pi)$ it is complementary to the horizontal sub-bundle $\pi^*_k(T^*M)$. The projections on the two components are type $(1,1)$ tensor fields, each of which captures the contact information and may be called the *contact structure* of the jet bundle. In coordinates, they are

$$h_k = dx^i \otimes \frac{d}{dx^i}\,, \qquad v_k = \sum_{|I|=0}^{k-1} \theta^\alpha_I \otimes \frac{\partial}{\partial u^\alpha_I}\,.$$

The first of these, h_k, may be regarded as a kind of universal connection; indeed, if γ is a section of $\pi_{1,0} : J^1\pi \to E$ then the composition $h_1 \circ \gamma$ is a type $(1,1)$ tensor field on E whose image is complementary to the vertical bundle $V\pi$: see subsection 2.7 below.

It is worth mentioning that the definition of the contact structure makes essential use of the fibration of E over a base manifold M, and so cannot be constructed in this way for some other types of jet structure such as manifolds of jets of immersions.

2.6 Semiholonomic jets

One of the significant features of manifolds of jets is that the highest-order derivatives play a rather distinctive rôle. We are able to take account of this property to construct a new manifold, the manifold of semi-holonomic jets, that lies between the manifolds of holonomic and non-holonomic jets. We restrict attention again in this subsection to affine jet bundles, although similar considerations apply to bundles of jets of immersions.

We start by regarding $\pi_{k,k-1} : J^k\pi \to J^{k-1}\pi$ as a map fibred over the identity on M; we may therefore prolong it to give the map

$$j^1\pi_{k,k-1} : J^1\pi_k \to J^1\pi_{k-1}\,.$$

On the other hand, if we consider the first jet bundle of the fibred manifold $\pi_k : J^k\pi \to M$ we obtain the map $(\pi_k)_{1,0} : J^1\pi_k \to J^k\pi$, and regarding $J^k\pi$ as a submanifold of $J^1\pi_{k-1}$ we obtain a second map

$$(\pi_k)_{1,0} : J^1\pi_k \to J^1\pi_{k-1}\,.$$

It therefore makes sense to look at the subset of $J^1\pi_k$ where these two maps take the same values.

Definition 2.7 The *semiholonomic manifold* $\widehat{J}^{k+1}\pi$ is the submanifold of $J^1\pi_k$ given by

$$\widehat{J}^{k+1}\pi = \{ j^1_x\psi \in J^1\pi_k : j^1\pi_{k,k-1}(j^1_x\psi) = (\pi_k)_{1,0}(j^1_x\psi) \}\,.$$

We can see that $\widehat{J}^{k+1}\pi$ is indeed a submanifold of $J^1\pi_k$ by noting that it is given locally by the coordinate conditions $u^\alpha_{J,j} = u^\alpha_{J+1_j}$ for $|J| \le k-1$, so that the jet coordinates of order k or lower are completely symmetric. We may therefore take coordinates

$$(x^i, u^\alpha_I, u^\alpha_{K,i}) \qquad |I| \le k, |K| = k$$

on $\widehat{J}^{k+1}\pi$.

We mentioned earlier that there is, in general, no canonical projection $J^1\pi_k \to J^{k+1}\pi$; by contrast, there *is* a projection $\widehat{J}^{k+1}\pi \to J^{k+1}\pi$. We may construct this projection by noting that the condition for $j^1_x\psi \in J^1\pi_k$ to lie in the submanifold $\widehat{J}^{k+1}\pi$ may be written as

$$j^1_x(\pi_{k,k-1}\circ\psi) = \psi(x)\,, \tag{2.2}$$

and defining the projection by

$$j^1_x\psi \mapsto j^{k+1}_x(\pi_{k,0}\circ\psi)\,;$$

we may check that condition (2.2) implies that the definition does not depend upon the choice of representative section ψ.

In the particular case $k = 1$ we note that the coordinates on $\widehat{J}^2\pi$ are $(x^i, u^\alpha, u^\alpha_i, u^\alpha_{ij})$ in ordinary index notation, whereas those on $J^2\pi$ are $(x^i, u^\alpha, u^\alpha_i, u^\alpha_{(ij)})$ with symmetric second-order jet coordinates, so we might expect to be able to find a complementary manifold with skew-symmetric second-order coordinates. This is indeed the case, and we may write

$$\widehat{J}^2\pi \cong J^2\pi \times_{J^1\pi} \left(\textstyle\bigwedge^2 T^*M \otimes_{J^1\pi} V\pi\right) \tag{2.3}$$

([36], Theorem 5.3.4). The higher-order semiholonomic jet bundles may similarly be written as fibre products, but the complementary manifolds are more complicated as they have coordinates which are partly symmetric and partly skew-symmetric [26].

2.7 Connections

A connection (or, more precisely, an Ehresmann connection) on a principal bundle $P \to M$ with structure group G is usually taken to be a $\mathfrak{g}$-valued 1-form on P, where $\mathfrak{g}$ is the Lie algebra of G, satisfying certain conditions. An equivalent definition takes the horizontal subspaces of the connection as fundamental, and the equivariance condition on the connection form translates into a similar condition on the family of subspaces. But one might easily imagine a less restrictive definition with no equivariance condition on the subspaces, and then this could be applied to a more general situation where the principal bundle was replaced by a fibred manifold without any particular structure group.

Definition 2.8 A *connection* on the fibred manifold $\pi : E \to M$ is a sub-bundle $H \to E$ of the tangent bundle $TE \to E$ which is complementary to the vertical sub-bundle $V\pi \to E$. The fibres of the bundle H are the *horizontal subspaces* of the connection. A vector field on E is called *horizontal if it takes its values in the horizontal subspaces.*

A version of the connection form definition may be recovered in this situation by considering, rather than an algebra-valued 1-form, a tangent-valued 1-form Γ which is a projection and whose images are the horizontal subspaces. Taking a fibred chart (x^i, u^a) on E, we obtain

$$\Gamma = dx^i \otimes \left(\frac{\partial}{\partial x^i} + \Gamma^\alpha_i \frac{\partial}{\partial u^\alpha}\right).$$

The similarity between this formula and the coordinate formula for the horizontal contact structure h on the jet bundle $J^1\pi$ should be immediate. The differences are that Γ is defined on E whereas h is defined on $J^1\pi$, and that the functions Γ^α_i defined locally on E have replaced the jet coordinate functions u^α_i. This leads us to an alternative definition of a connection on a fibred manifold.

Definition 2.9 A connection on the fibred manifold $\pi : E \to M$ is a section $\gamma : E \to J^1\pi$ of the affine bundle $J^1\pi \to E$.

The relationship between the two definitions is simply that $\Gamma = h \circ \gamma$: locally, the section γ substitutes the concrete functions Γ^α_i for the jet coordinates u^α_i. The correspondence is bijective ([36], Proposition 4.6.3).

On a principal bundle, a fundamental property of a connection is its *curvature*, a $\mathfrak{g}$-valued 2-form obtained by differentiating the connection form. In the more general framework we may also construct the curvature, and again there are two different approaches using the two definitions of a connection. Using Γ, we may define the curvature by taking the Lie bracket of two horizontal vector fields and measuring its deviation from the horizontal.

Definition 2.10 The curvature of a connection Γ is the map $\Omega : \mathfrak{X}(E) \times \mathfrak{X}(E) \to \mathfrak{X}(E)$ given by

$$\Omega(X, Y) = [\Gamma(X), \Gamma(Y)] - \Gamma([\Gamma(X), \Gamma(Y)])\,.$$

It is easy to check that Ω is bilinear over the functions on E and so is a vertical tangent-valued 2-form. In coordinates it is

$$\Omega = \Omega^\alpha_{ij}\, dx^i \wedge dx^j \otimes \frac{\partial}{\partial u^\alpha} = \tfrac{1}{2}\left(\frac{\partial \Gamma^\alpha_j}{\partial x^i} + \Gamma^\beta_i \frac{\partial \Gamma^\alpha_j}{\partial u^\beta} - \frac{\partial \Gamma^\alpha_i}{\partial x^j} - \Gamma^\beta_j \frac{\partial \Gamma^\alpha_i}{\partial u^\beta} \right) dx^i \wedge dx^j \otimes \frac{\partial}{\partial u^\alpha}\,. \tag{2.4}$$

On the other hand, we may take the section γ as the connection, and prolong it to give a map $j^1\gamma : J^1\pi \to J^1\pi_1$. Taking the composite

$$j^1\gamma \circ g : E \to J^1\pi_1$$

we may check that, in a chart, $u^\alpha_i \circ j^1\gamma \circ g = u^\alpha_{,i} \circ j^1\gamma \circ g$ so that the composite map takes its values in the semiholonomic manifold $\widehat{J}^2\pi$. Recalling (equation 2.3) the decomposition of $\widehat{J}^2\pi$ as a fibre product, we may take the projection on the second component $\bigwedge^2 \pi_1^* T^* M \otimes \pi_{1,0}^* V\pi$ to obtain a vertical tangent-valued 2-form which may be shown to equal $-\Omega$ ([36], Theorem 5.3.5).

A connection with vanishing curvature has special properties.

Definition 2.11 If $\gamma : E \to J^1\pi$ is a connection then any local section $\phi : U \to E$, $U \subset M$ satisfying

$$j^1\phi = \gamma \circ \phi$$

is called an *integral section* of γ.

Theorem 2.12 *A connection γ has integral sections if, and only if, it has vanishing curvature. We also say that such a connection is* integrable.

The proof of this theorem involves noting that, in coordinates, the condition for a local section ϕ to be an integral section is

$$\frac{\partial \phi^\alpha}{\partial x_i} = \Gamma^\alpha_i \circ \phi\,, \tag{2.5}$$

so that ϕ is a solution of the partial differential equation given in (2.5); the vanishing of the curvature (2.4) is just the Frobenius condition for this equation to have solutions ([36], Proposition 5.3.1).

3 Differential equations

3.1 Differential operators

One way in which differential equations can be constructed is by the use of differential operators. We shall shall consider two fibred manifolds $\pi_1 : E_1 \to M$, $\pi_2 : E_2 \to M$ over the same base M; a map $\Delta : \Gamma(\pi_1) \to \Gamma(\pi_2)$ between sections of these two fibred manifolds is called an *operator*. A simple type of operator arises from a map $f : E_1 \to E_2$ fibred over the identity on M by setting

$$\Delta_f \phi = f \circ \phi$$

for $\phi \in \Gamma(\pi_1)$. Such an operator Δ_f is a *pointwise* operator because, for each $x \in M$, the value of $\Delta_f \phi$ at x depends only on the value of ϕ there. A more general type of operator Δ is called *local*, where $\Delta\phi(x)$ depends on the germ of ϕ at x rather than just its value. By definition, a *differential operator* is one where $\Delta\phi(x)$ depends on the derivatives of ϕ at x; in the language of jets there is a map $f : J^k\pi_1 \to E_2$ such that $\Delta = \Delta_f$ where now

$$\Delta_f \phi = f \circ j^k \phi$$

for some k. The least such k is called the *order* of the differential operator.

We say that an operator Δ is *regular* if, whenever ϕ_t is a smoothly parametrised family of sections of π_1, $\Delta\phi_t$ is a smoothly parametrised family of sections of π_2. If Δ_f is a differential operator then it is regular precisely when the associated map f is smooth ([18], Proposition 14.14).

It is clear that a differential operator is local, because the jet of a section is determined by its germ. A study of the converse involves *Peetre's Theorem*. The original version of Peetre's Theorem involved linear differential operators; we may define these on vector bundles $\pi_1 : E_1 \to M$, $\pi_2 : E_2 \to M$ as the spaces of sections $\Gamma(\pi_1)$, $\Gamma(\pi_2)$ are then vector spaces. It now makes sense to talk about the *support* of a section, defined in the usual way as the closure of the set where the section does not vanish. We may also introduce linear operators which do not increase support; it is clear that a linear local operator satisfies this condition.

Theorem 3.1 (Peetre's Theorem [33, 34]; see also ([18], Theorem 19.1) and the article by J. Szilasi and R. L. Lovas in this volume[41])

Let Δ be a linear support non-increasing operator $\Gamma(\pi_1) \to \Gamma(\pi_2)$. For every compact set $K \subset M$ there is a natural number k such that, whenever $\phi_1, \phi_2 \in \Gamma(\pi_1)$ and $x \in K$, the condition $j^k_x\phi_1 = j^k_x\phi_2$ implies $\Delta\phi_1(x) = \Delta\phi_2(x)$.

Thus we may say that every linear local operator is a differential operator. Note that differentiability of the image sections $\Delta\phi$ is required for the theorem to hold: see [18], Example 19.3.

A version of Peetre's Theorem for non-linear operators is significantly more complicated; see [39], and also ([18], Sections 19.4–19.15) and [41].

3.2 Differential equations on jet bundles

We saw in subsection 2.7 that a connection defines a first-order partial differential equation of a particular type. But we can use jet bundles to give intrinsic descriptions of much more general differential equations. If $R \subset J^k\pi$ is a closed embedded submanifold then it will be defined locally by the vanishing of some functions f^μ. If we add some conditions to ensure that $\pi_k|_R : R \to M$ is a fibred submanifold of $\pi_k : J^k\pi \to M$, and that R is genuinely defined on $J^k\pi$ rather than on a lower-order jet bundle, then we can say that R 'is' a k-th order partial differential equation in m independent and n dependent variables. (If $m = 1$ then it is an ordinary differential equation.) A local solution will then be a local section $\phi : U \to E$, $U \subset M$, whose prolongation satisfies $j^k\phi(U) \subset R$. This is like a generalization of regarding a first-order ordinary differential equation, not as a vector field, but as the image of a vector field in the tangent bundle. It isn't quite the same, because the solutions of the vector field equation are the integral curves, and the vector field itself is needed (rather than its image) in order to fix the parametrization; in $J^k\pi$ the solutions are local sections, carrying their own parametrization, so it is adequate to give the equation as a submanifold.

Differential equations often arise in this intrinsic format from differential operators between fibred manifolds. If $\Delta : \Gamma(\pi_1) \to \Gamma(\pi_2)$ is such an operator with $\Delta\phi = f \circ j^k\phi$, and if ψ is a fixed section of π_2, we may let $R = f^{-1}(\psi(M))$. Typically π_2 is a vector bundle and ψ is its zero section, so that we recover the specification of R in terms of the vanishing of some locally-defined functions $f^\mu = w^\mu \circ f$ where w^μ are fibre coordinates on E_2.

To make R 'look like' a differential equation, we have to use its local coordinate representation: in the case of partial differential equations, there may be integrability conditions to be obtained by cross-differentiating the equations. Nevertheless we can, using jet bundles, analyse the problem in a fairly abstract way. The idea is to say that the equation is *formally integrable* if at each point we can find a sequence of Taylor coefficients such that a formal power series with those coefficients would satisfy the equation; we then demonstrate the existence of these coefficients in an intrinsic way. Nothing is said about the convergence of the formal power series, and indeed there are known to be smooth (but non-analytic) differential equations which are formally integrable but do not have actual solutions: the formal power series thus cannot converge away from the point in question. The first example of such an equation was given by Lewy [25] (see also [35], Chapter 5 Section 7.2):

$$\frac{\partial y^1}{\partial x^1} - \frac{\partial y^2}{\partial x^2} - 2x^2\frac{\partial y^1}{\partial x^3} - 2x^1\frac{\partial y^2}{\partial x^3} = h$$
$$\frac{\partial y^2}{\partial x^1} + \frac{\partial y^1}{\partial x^2} + 2x^1\frac{\partial y^1}{\partial x^3} - 2x^2\frac{\partial y^2}{\partial x^3} = 0\,.$$

The equation is formally integrable, according to the Theorem in Section 3.5; but with a suitable choice of a smooth but non-analytic function h it will have no local solutions. The reason behind this may be found in complex analysis. Setting $z = x^1 + ix^2$, $w = y^1 + iy^2$ and $t = x^3$, we may rewrite the equations as

$$\frac{\partial w}{\partial \bar{z}} + iz\frac{\partial w}{\partial t} = \frac{h}{2}$$

and the operator on the left-hand side is the Cauchy-Riemann operator on a suitable hypersurface in $\mathbf{C}^2$. There is now a substantial theory behind such operators [15]. Of course there are many differential equations which are not of this form, and in particular cases (such as when the equation R is the image of a connection γ) we may be sure that a formally integrable equation does indeed have local solutions.

3.3 Prolonging differential equations

In order to analyse the formal integrability of a differential equation R, we proceed in several stages. The first stage is to prolong the equation, by differentiating it. By definition $\pi_k|_R : R \to M$ is a fibred manifold, and so we can construct its jet bundles $(\pi_k|_R)_l : J^l(\pi_k|_R) \to M$ where of course $J^l(\pi_k|_R) \subset J^l\pi_k$.

Definition 3.2 The l-th prolongation of the differential equation $R \subset J^k\pi$ is the holonomic subset

$$R^{(l)} = J^l(\pi_k|_R) \cap J^{k+l}\pi$$

of $J^l(\pi_k|_R)$.

In general, $\pi_{k+l}|_{R^{(l)}} : R^{(l)} \to M$ need not be a fibred submanifold of $J^{k+l}\pi$; but if it is then we may prolong $R^{(l)}$ again, and $(R^{(l)})^{(p)} = R^{(l+p)}$. If the l-th prolongation is a fibred submanifold for every $l \geq 1$ then we say that R is *regular*. A local section ϕ is a solution of R if, and only if, it is a solution of the prolonged equation $R^{(l)}$.

Definition 3.3 The regular differential equation $R \subset J^k\pi$ is *formally integrable* if, for each $l \geq 1$, $\pi_{k+l,k+l-1}(R^{(l)}) = R^{(l-1)}$.

The essential meaning of this definition is that, no matter how often we differentiate the original equation, we cannot obtain any new integrability conditions that would provide constraints on the possible values of the derivatives of a solution. We may therefore use the differentiated equations to construct the sequence of Taylor coefficients. In coordinates, if R is defined by the equations $f^\mu(x^i, u^\alpha, u^\alpha_I) = 0$, then $R^{(l)}$ will be defined by

$$f^\mu = 0\,, \qquad \frac{d^{|I|} f^\mu}{dx^I} \quad (1 \leq |I| \leq l)\,.$$

As the total derivatives are quasilinear expressions, it will always be possible to solve for the highest-order derivatives and so construct the Taylor coefficients.

Of course the trouble with this definition of formal integrability is that it is impossible to check it in a finite number of steps. We can deal with this problem using the algebraic technique of Spencer coholomogy to measure the additional constraints which arise when we differentiate the equation.

3.4 The symbol of a differential equation

An important object associated with a differential equation is its *symbol*. This is a family of vector spaces G_R defined on the equation manifold R. The idea is that G_R captures the dependence of the differential equation on the highest derivatives.

To work towards a definition of G_R, we recall first that $\pi_{k,k-1} : J^k\pi \to J^{k-1}\pi$ is an affine bundle, modelled on the vector bundle $S^k T^*M \otimes_{J^{k-1}\pi} V\pi$. Thus the vector bundle

$V\pi_{k,k-1}$ of tangent vectors on $J^k\pi$ vertical over $J^{k-1}\pi$ is isomorphic to $S^kT^*M \otimes_{J^{k-1}\pi} V\pi$, and so we may regard the restriction $V\pi_{k,k-1}|_R$ as satisfying

$$V\pi_{k,k-1}|_R \subset S^kT^*M \otimes_{J^{k-1}\pi} V\pi\Big|_R .$$

Definition 3.4 The symbol G_R of the differential equation $R \subset J^k\pi$ is the family of vector spaces $V\pi_{k,k-1}|_R$.

Thus, at each point $j^k_x\phi \in R$, the symbol $G_{j^k_x\phi}$ may be identified with a vector subspace of $S^kT^*M \otimes_{J^{k-1}\pi} V\pi$; to simplify the notation we shall, for the remainder of this section, write $S^kT^* \otimes V$ for the latter space, omitting the arguments of the functors T^* and V and the point, so that we have the identification

$$G_R \subset S^kT^* \otimes V$$

representing an inclusion of vector spaces at $j^k_x\phi$. In general the dimensions of these subspaces may vary; but if the dimensions are all the same then the symbol will be a vector bundle over R. In coordinates, the symbol at a point is the subspace defined by the equations

$$\sum_{|I|=k} \frac{\partial f^\mu}{\partial u^\alpha_I} v^\alpha_I = 0$$

where v^α_I are coordinates on $S^kT^* \otimes V$ with respect to bases dx^i of T^* and $\partial/\partial u^\alpha$ of V.

As with the equation itself, we may prolong the symbol. This, though, is an algebraic operation, carried out at each fixed point $j^k_x\phi$, and so we may describe it at the level of the individual vector spaces. We shall make use of the bundle $F \to J^k\pi$ defined by

$$F = V\pi_k|_R / V(\pi_k|_R) ;$$

as R is a fibred submanifold we may identify this with the normal bundle of R in $J^k\pi$. Once again we shall also write F for the fibre of this bundle at the point $j^k_x\phi$. As $S^kT^* \otimes V \subset V_{j^k_x\phi}\pi_k$ we may restrict the quotient map $V_{j^k_x\phi}\pi_k \to F$ to give a map

$$S^kT^* \otimes V \to F$$

and hence, by taking tensor products, maps

$$S^lT^* \otimes S^kT^* \otimes V \to S^lT^* \otimes F$$

for each $l \geq 0$. Combining these with the maps

$$S^{k+l}T^* \otimes V \to \bigotimes^l T^* \otimes \bigotimes^k T^* \otimes V \to S^lT^* \otimes S^kT^* \otimes V$$

given by inclusion and symmetrization, we obtain maps

$$S^{k+l}T^* \otimes V \to S^lT^* \otimes F .$$

Definition 3.5 The l-th prolongation of the symbol G of R is the kernel

$$G^{(l)}_R \subset S^{k+l}T^* \otimes V$$

of the map $S^{k+l}T^* \otimes V \to S^lT^* \otimes F$.

Note that we have defined each vector space $G_R^{(l)}$ at a point $j_x^k\phi$ of the original differential equation R, not at a point of the prolonged equation $R^{(l)}$; but in fact there is a canonical isomorphism between this vector space and the symbol $G_{R^{(l)}}$ at a point $j_x^{k+l}\phi \in R^{(l)}$. In coordinates, each of these vector spaces is defined by the equations

$$\sum_{|I|=k} \frac{\partial f^\mu}{\partial u_I^\alpha} v_{I+J}^\alpha = 0\,, \qquad |J| = l\,.$$

We can use the prolonged symbols to remove the need for checking that a differential equation is regular before applying the definition of formal integrability.

Proposition 3.6 (see [35], Chapter 2 Theorem 3.16)

The differential equation $R \subset J^k\pi$ is formally integrable if, and only if, for each $l \geq 1$ the prolonged equation satisfies $\pi_{k+l,k+l-1}(R^{(l)}) = R^{(l-1)}$ and the prolonged symbol $G_R^{(l)}$ is a vector bundle.

3.5 Spencer sequences and formal integrability

The key to analyzing the formal integrability of a differential equation is to look, not so much at the prolongations of the equation, as at the prolongations of its symbol. These may be related by a map called the *Spencer δ-map* which relates symmetric and skew-symmetric products of vector spaces (see [40]). We continue writing T^* for the pullback of the cotangent space T_x^*M to the point $j_x^k\phi$.

Definition 3.7 The Spencer δ-map is the map

$$\delta : \bigwedge^p T^* \otimes S^q T^* \to \bigwedge^{p+1} T^* \otimes S^{q-1} T^*$$

given by composition of the inclusion

$$\bigwedge^p T^* \otimes S^q T^* \to \bigotimes^p T^* \otimes \bigotimes^q T^* = \bigotimes^{p+q} T^*$$

and the symmetrization and skew-symmetrization

$$\bigotimes^{p+q} T^* = \bigotimes^{p+1} T^* \otimes \bigotimes^{q-1} T^* \to \bigwedge^{p+1} T^* \otimes S^{q-1} T^*\,.$$

By taking tensor products we obtain a map

$$\delta : \bigwedge^p T^* \otimes S^q T^* \otimes V \to \bigwedge^{p+1} T^* \otimes S^{q-1} T^* \otimes V\,.$$

The Spencer δ-maps clearly satisfy the condition $\delta^2 = 0$, and by repeated composition we obtain, for any $l \geq 0$, the Spencer sequence

$$0 \to S^{k+l} T^* \otimes V \xrightarrow{\delta} T^* \otimes S^{k+l-1} T^* \otimes V \xrightarrow{\delta} \ldots \xrightarrow{\delta} \bigwedge^m T^* \otimes S^{k+l-m} T^* \otimes V \to 0$$

where we take $S^{k+l-m} T^* = 0$ for $k + l < m$.

Theorem 3.8 (see [35], Chapter 3 Proposition 1.5)

The Spencer sequences are exact.

So far we have taken no account of any differential equation. If an equation $R \subset J^k\pi$ is given then we may consider the prolongations $G_R^{(l)} \subset S^{k+l}T^* \otimes V$ of its symbol, obtaining their Spencer sequences

$$0 \to G_R^{(l)} \xrightarrow{\delta} T^* \otimes G_R^{(l-1)} \xrightarrow{\delta} \ldots \xrightarrow{\delta} \bigwedge^m T^* \otimes G_R^{(l-m)} \to 0$$

with the conventions that $G_R^{(l-p)} = S^{k+l-p}T^* \otimes V$ if $0 \leq k+l-p < k$ and that $G_R^{(l-p)} = 0$ if $k+l-p < 0$.

Definition 3.9 The *Spencer cohomology* of the symbol G_R is the family of cohomology spaces $H^{p,k+l-p}$ at the terms $\bigwedge^p T^* \otimes G_R^{(l-p)}$ of the Spencer sequences for the prolongations $G_R^{(l)}$.

Definition 3.10 The symbol G_R is said to be *s-acyclic* if $H^{p,q} = 0$ for $0 \leq p \leq s$, $q \geq k$. It is said to be *involutive* if $H^{p,q} = 0$ for $0 \leq p \leq m$, $q \geq k$.

Given these properties of the symbol G_R, we now have a result which enables us to check that a given differential equation is formally integrable.

Theorem 3.11 (Goldschmidt [12])

If the differential equation R has a 2-acyclic symbol G_R, and if its first prolongation $R^{(1)}$ is a fibred submanifold of $J^{k+1}\pi$ and satisfies $\pi_{k+1,k}(R^{(1)}) = R$, then R is formally integrable.

A formally integrable equation with an involutive symbol is said to be an *involutive equation*, and if the fibred manifold $E \to M$ is real-analytic then such a system is guaranteed to have genuine solutions, unique for given initial conditions: this is the *Cartan-Kähler Theorem*, and is ultimately a consequence of the Cauchy-Kovalevskaya Theorem (see [35], Chapter 4 Theorem 4.6).

Of course one needs to be able to check whether a symbol is involutive in order to use these results, and there are different approaches to the task. A homological approach is described in [24]; a more traditional method [24, 35] is to use *Cartan characters*. To find these, we choose a coordinate system x^i in a neighbourhood of $x \in M$ and subspaces $G_{R,j} \subset G_R$ defined by

$$G_{R,j} = \{\xi \in G_R : i_{\partial/\partial x^i}\xi = 0,\, 1 \leq i \leq j\}\,.$$

These subspaces form a filtration

$$0 = G_{R,m} \subset G_{R,m-1} \subset \ldots \subset G_{R,1} \subset G_R$$

and the Cartan character α_j is the increase in dimension $\dim G_{R,j-1} - \dim G_{R,j}$.

Theorem 3.12 (see [35], Chapter 3, Remark 2.28)

The Cartan characters satisfy $\dim G_R^{(1)} \leq \sum_{j=1}^m j\alpha_j$, *and the symbol is involutive if, and only if, there are coordinates x^i such that equality is attained. Such coordinates are called* δ-regular *or* quasi-regular.

It is worth noting that the Cartan characters used here may be related to those which arise in the study of differential equations using exterior differential systems. The relationship between Cartan characters and Spencer comology was described in a letter by J. P. Serre, quoted (in French) in an appendix to [14].

There remains the question of how to deal with an equation whose symbol is not involutive. The answer is to prolong the equation, and hence to prolong the symbol; eventually, the symbol will become involutive at, say, $G^{(l)}$. We now need to check that prolonging the equation once more to $R^{(l+1)}$ does not introduce integrability conditions. If it does, we take the projection $\pi_{k+l+1,k+l}(R^{(l+1)})$ as a new equation (having the same solutions as R) and start the process again. The *Cartan-Kuranishi Theorem* guarantees than this process terminates. The final equation may, of course, be a zero-dimensional submanifold, and then the original equation cannot have solutions.

Theorem 3.13 (Kuranishi [23])
If the equation R is regular then there are integers l, p such that $\pi_{k+l+p,k+l}(R^{(l+p)})$ is involutive.

4 The calculus of variations

4.1 Variational problems on jet bundles

Problems in the calculus of variations can usefully be formulated on jet manifolds, and by doing so the need to study infinite-dimensional function spaces can often be avoided. We consider here only problems with fixed boundary conditions. There are essentially two types of problem, depending upon whether or not the parametrization of the solution manifold is important. For instance, a variational problem in mechanics might have trajectories as solutions, and the speed along the trajectory would be important; on the other hand, a variational problem in geometry might have paths giving the shortest distance between two points, and then only the image of the path would be significant.

In this subsection we consider variational problems on a jet bundle $J^k\pi$, where $\pi : E \to M$ is a fibred manifold; we suppose that the base manifold M is orientable. The problem is defined by a *Lagrangian*, an m-form λ which is horizontal over M. If M has a given volume form ω then we would have $\lambda = L\omega$ for some function L on $J^k\pi$. In a problem of this kind, a solution is a local section and so has a parametrisation.

Definition 4.1 An *extremal* of the variational problem defined by the Lagrangian λ is a local section $\phi : U \to E, U \subset M$, satisfying

$$\left.\frac{d}{dt}\right|_{t=0} \int_C (j^k\phi_t)^*\lambda = 0$$

whenever $C \subset U$ is a compact m-dimensional submanifold and ϕ_t is a 1-parameter family of local sections of π satisfying $\phi_0 = \phi$ and $\phi_t|_{\partial C} = \phi|_{\partial C}$.

We study this problem by regarding the family of local sections ϕ_t as being generated by a *variation field*, a vector field X on E which is vertical over M and which satisfies $\phi_t = \psi_t \circ \phi$ where ψ_t is the local flow of X. The condition $\phi_t|_{\partial C} = \phi|_{\partial C}$ translates into $X|_{\pi^{-1}(\partial C)} = 0$. Using this, we are able to reformulate the problem as

$$\int_C (j^k\phi)^* d_{X^k}\lambda = 0$$

where d_{X^k} denotes the Lie derivative by the prolongation X^k of the variation field.

We now suppose that we can find another m-form ϑ_λ on a related jet bundle $J^l\pi$ (where $l \geq k$) such that the difference $\vartheta_l - \pi^*_{l,k}\lambda$ is a contact form on $J^l\pi$, and such that the contraction of $d\vartheta_\lambda$ with any vector field Y on $J^l\pi$ which is vertical over E must also result in a contact form. Any form ϑ_λ satisfying these two conditions is called a *Lepage equivalent* of λ.

The first condition immediately implies that $(j^k\phi_t)^*\lambda = (j^l\phi_t)^*\vartheta_\lambda$ for any local section ϕ, so that ϕ is an extremal of the variational problem defined by λ precisely when

$$\int_C (j^l\phi)^* d_{X^l}\vartheta_\lambda = 0\,.$$

We may write the Lie derivative d_{X^l} in terms of contraction as $i_{X_l}d + di_{X^l}$, giving two integrals; but then Stokes' Theorem gives

$$\begin{aligned}\int_C (j^l\phi)^* di_{X^l}\vartheta_\lambda &= \int_C d(j^l\phi)^* i_{X^l}\vartheta_\lambda \\ &= \int_{\partial C} (j^l\phi)^* i_{X^l}\vartheta_\lambda \\ &= 0\end{aligned}$$

as X^l vanishes on $\pi_l^{-1}(\partial C)$, so that ϕ is now an extremal precisely when

$$\int_C (j^l\phi)^* i_{X^l} d\vartheta_\lambda = 0\,.$$

We next use the second condition in the definition of a Lepage equivalent. If Y is an arbitrary vector field on $J^l\pi$ projectable to X then the difference $Y - X^l$ is vertical over E, so that $i_{(Y-X^l)}d\vartheta_\lambda$ is a contact form and hence $(j^l\phi)^* i_{(Y-X^l)}d\vartheta_\lambda$ vanishes. It follows that ϕ is an extremal precisely when

$$\int_C (j^l\phi)^* i_Y d\vartheta_\lambda = 0$$

for any projectable vector field Y on $J^l\pi$ vanishing on $\pi_l^{-1}(\partial C)$.

We now observe that the integrand $(j^l\phi)^* i_Y d\vartheta_\lambda$ may be written as

$$(j^{l+1}\phi)^* i_{\bar{Y}}\pi^*_{l+1,l}d\vartheta_\lambda$$

for any vector field $\bar{Y}$ on $J^{l+1}\pi$ projecting to Y. Write the $(m+1)$-form $\psi = \pi^*_{l+1,l}d\vartheta_\lambda$ as the sum

$$\psi = \psi_1 + \psi_2 + \ldots + \psi_{m+1}$$

of 1-contact, 2-contact, …, $(m+1)$-contact components (subsection 2.5); then by definition only the 1-contact part ψ_1 makes a contribution to the integrand, and the contraction $i_{\bar{Y}}\psi_1$ must be horizontal over M.

Now take a fibred chart (x^i, u^α) on E with domain U such that $\pi(U) \subset M$ is contained in the interior of C. Let $x \in \pi(U)$, and let f be a bump function on M vanishing outside $\pi(U)$ such that $f(x) = 1$. Take $\bar{Y}$ to be the vector field $\partial/\partial u^\alpha$ on $J^{l+1}\pi$, and write

$$i_{(f\bar{Y})}\psi_1 = fE_\alpha d^m x\,;$$

the functions E_α must then satisfy $(j^{l+1}\phi)^* E_\alpha = 0$ at x (and hence throughout $\pi(U)$) using the vanishing of the integral and the arbitrariness of the bump function. These equations for the local section ϕ are the *Euler-Lagrange equations* for the Lagrangian λ, and they take the usual coordinate form where

$$E_\alpha = \sum_{|I|=0}^{k} (-1)^{|I|} \frac{d^{|I|}}{dx^I} \left(\frac{\partial L}{\partial u_I^\alpha} \right) ;$$

in general we must therefore have $l+1 \geq 2k$, although for particular Lagrangians it may be possible to take smaller values of l. The $(m+1)$-form ε_λ on $J^{l+1}\pi$ given in coordinates by $E_\alpha du^\alpha \wedge d^m x$ is, indeed, globally well-defined and is just the 1-contact part of $\pi^*_{l+1,l} d\vartheta_\lambda$; the zero set of ε_λ is a submanifold of $J^l\pi$ defining the global Euler-Lagrange equation.

Of course, the calculation above presupposes the existence of a Lepage equivalent for an arbitrary Lagrangian λ. The following results are known.

Theorem 4.2 LEPAGE EQUIVALENT THEOREM

(1) Every Lagrangian has at least one global Lepage equivalent defined on $J^{2k-1}\pi$.

(2) If $\dim M = 1$ then there is a unique global Lepage equivalent of each Lagrangian λ, defined on $J^{2k-1}\pi$. If, in coordinates, $\lambda = L\, dt$ then

$$\vartheta_\lambda = L\, dt + \sum_{r=0}^{k-1} \left(\sum_{s=0}^{k-r-1} (-1)^s \frac{d^s}{dt^s} \left(\frac{\partial L}{\partial u^\alpha_{(r+s+1)}} \right) \right) \theta^\alpha_{(r)} .$$

(3) If $\dim M \geq 2$ and $k = 1$ then three of the global Lepage equivalents are natural operators (subsection 5.3) defined on $J^1\pi$. If, in coordinates, $\lambda = L\, d^m x$ then they are

$$\text{(a)} \quad \vartheta_\lambda = L\, d^m x + \frac{\partial L}{\partial u_i^\alpha} \theta^\alpha \wedge d^{m-1} x_i$$

$$\text{(b)} \quad \widehat{\vartheta}_\lambda = \frac{1}{L^{m-1}} \bigwedge\nolimits_{i=1}^{m} \left(L\, dx^i + \frac{\partial L}{\partial u_i^\alpha} \theta^\alpha \right)$$

$$\text{(c)} \quad \widetilde{\vartheta}_\lambda = \sum_{r=0}^{\min\{m,n\}} \frac{1}{(r!)^2} \frac{\partial^r L}{\partial y_{i_1}^{\alpha_1} \ldots \partial y_{i_r}^{\alpha_r}} \theta^{\alpha_1} \wedge \ldots \wedge \theta^{\alpha_r} \wedge d^{m-r} x_{i_1 \cdots i_r} ,$$

where formula (b) is valid for a non-vanishing Lagrangian.

(4) If $\dim M \geq 2$ and $k = 2$ then two of the global Lepage equivalents are natural operators defined on $J^3\pi$. If, in coordinates, $\lambda = L\, d^m x$ then they are

$$\text{(a)} \quad \vartheta_\lambda = L\, d^m x + \left(\left(\frac{\partial L}{\partial u_i^\alpha} - \frac{1}{\#(ij)} \frac{d}{dx^j} \frac{\partial L}{\partial u_{ij}^\alpha} \right) \theta^\alpha + \frac{1}{\#(ij)} \frac{\partial L}{\partial u_{ij}^\alpha} \theta_j^\alpha \right) \wedge d^{m-1} x_i$$

$$\text{(b)} \quad \widehat{\vartheta}_\lambda = \frac{1}{L^{m-1}} \bigwedge\nolimits_{i=1}^{m} \big(L\, dx^i +$$

$$\left(\frac{\partial L}{\partial u^\alpha_i} - \frac{1}{\#(ij)} \frac{d}{dx^j} \frac{\partial L}{\partial u^\alpha_{ij}} \right) \theta^\alpha + \frac{1}{\#(ij)} \frac{\partial L}{\partial u^\alpha_{ij}} \theta^\alpha_j \right) ,$$

where formula (b) is valid for a non-vanishing Lagrangian.

(5) If $\dim M \geq 2$ *and* $k \geq 2$ *then the global Lepage equivalent is not unique.*

There are several different proofs of (1); one approach is described in ([36], Section 6.5). This constructs, in an intrinsic way, a (unique) global Lepage equivalent for first-order Lagrangians (formula (3a)), and then uses the identification $J^{k+1}\pi \subset J^1\pi_k$ to construct, recursively, Lepage equivalents for higher-order Lagrangians. The fact that there is no canonical projection $J^1\pi_k \to J^{k+1}\pi$ means, though, that the latter construction is not unique for $m \geq 2$ (although see the comment on formula (4a) below).

The fact that the formulæ (2)–(4) do indeed describe Lepage equivalents, and that these formulæ are unaffected by changes of fibred coordinates on E, may be verified by direct calculation. An intrinsic method of constructing formula (2) uses a 'vertical endomorphism' operator; this is related to the operators of the same name described in subsection 4.2 below. Formula (3b) is classical [3], and is invariant under an arbitrary change of coordinates on E (that is, ignoring the fibration over M). This property also holds for formula (3c) ([20] and, independently, [2]), with the additional property that $d\widetilde{\vartheta}_\lambda = 0$ exactly when $\varepsilon_\lambda = 0$. Formula (4a) may be constructed using the recursive technique described above, but making use of the projection from the semiholonomic manifold $\widehat{J}^2\pi \to J^2\pi$ [37]. Formula (4b) (and also formulæ (2), (3b) and (3c)) may be obtained by considering an associated homogeneous problem and projecting [4, 5]; see also [30]. Finally, the negative result (5) may be found in [16].

4.2 Homogeneous variational problems

Variational problems whose solutions are unparametrised submanifolds may be specified in two different ways. A direct method uses jets of submanifolds rather than jets of sections, whereas the approach described in this subsection is to use instead spaces of non-degenerate velocities: in general, the solution of a variational problem in this context would have a preferred parametrisation, but we impose a homogeneity condition on the Lagrangian to ensure that any reparametrisation is also a solution. Both approaches are commonly used in Finsler geometry, and an interesting comparison between the two is given in ([1], Section 2.1). The more general relation between the two approaches is described in an article by D. Grigore in this Handbook [13].

We consider, therefore, variational problems on $\mathcal{F}^k_{(m)}E$, and the Lagrangian will be a function on this manifold, rather than an m-form. To describe the homogeneity condition, we shall use a version of the total derivative operators described earlier for jet bundles. In the present context these are the m vector fields d_j along the projection $\mathcal{F}^{k+1}_{(m)}E \to \mathcal{F}^k_{(m)}E$ given in coordinates as

$$d_j = \sum_{|J|=0}^{k} u^a_{J+1_j} \frac{\partial}{\partial u^a_J} .$$

An intrinsic construction of d_j takes a general point $j_0^{k+1}\gamma \in \mathcal{F}^{k+1}_{(m)}E$ and chooses a representative map $\gamma : \mathbb{R}^m \to E$; then

$$d_j|_{j_0^{k+1}\gamma} = (j^k\gamma)_* \left(\frac{\partial}{\partial t^j}\bigg|_0 \right) \in T_{j_0^k\gamma}\mathcal{F}^{k-1}_{(m)}E$$

where $\partial/\partial t^j$ is the j-th standard coordinate vector field on $\mathbb{R}^m$.

We also need a family of m type $(1,1)$ tensor fields S^i on $\mathcal{F}^k_{(m)}E$ whose intrinsic construction is a little more complicated (see [4]) but which may be expressed in coordinates as

$$S^i = \sum_{|I|=0}^{k} (I(i)+1)\frac{\partial}{\partial u^a_{I+1_i}} \otimes du^a_I \,.$$

These tensor fields commute, and so we may use multi-index notation and write S^I for composite tensor fields. We now define the vector fields Δ^I_i on $\mathcal{F}^k_{(m)}E$ by

$$\Delta^I_j = S^I(d_j)\,;$$

although in principle these would also be vector fields along the projection $\mathcal{F}^{k+1}_{(m)}E \to \mathcal{F}^k_{(m)}E$, they are constant on the fibres from the properties of S^i and so we regard them as vector fields on $\mathcal{F}^k_{(m)}E$. These vector fields Δ^I_j are, in fact, the fundamental vector fields of the principal L^k_{m+}-bundle $\mathcal{F}^k_{(m)}E \to J^k_+(E,m)$.

We can now give the condition for a Lagrangian function $L : \mathcal{F}^k_{(m)}E \to \mathbb{R}$ to be homogeneous: it is that

$$\Delta^i_j(L) = \delta^i_j L\,, \qquad \Delta^I_j(L) = 0 \text{ for } |I| \geq 2.$$

Definition 4.3 An *extremal* of the variational problem defined by the homogeneous Lagrangian L is a map $\sigma : \mathbb{R}^m \to E$ satisfying

$$\frac{d}{dt}\bigg|_{t=0} \int_C ((j^k\sigma_t)^* L)\, d^m t = 0$$

whenever $C \subset \mathbb{R}^m$ is a compact m-dimensional submanifold. σ_t is a 1-parameter family of maps $\mathbb{R}^m \to E$ satisfying $\sigma_0 = \sigma$ and $\sigma_t|_{\partial C} = \sigma|_{\partial C}$, and $d^m t$ is the standard volume form on $\mathbb{R}^m$.

Theorem 4.4 (see, for example, [4] Theorem 4.2)

If σ is an extremal and ψ is an orientation-preserving diffeomorphism of $\mathbb{R}^m$ then $\sigma\circ\psi$ is also an extremal.

There is no concept of a Lepage equivalent for homogeneous Lagrangians; there are, however, objects which may be used in a similar way, and these may be defined irrespective of the order of the Lagrangian. First, the *Hilbert forms* ϑ_i defined on $\mathcal{F}^{2k-1}_{(m)}E$ are generalisations of the 1-form of the same name used in Finsler geometry, and are specified by

$$\vartheta^i = \sum_I \frac{(-1)^{|I|}}{(|I|+1)I!} d_I(S^{I+1_i} dL)\,;$$

they are related to the Euler-Lagrange form ε_L on $\mathcal{F}^{2k}_{(m)}E$ by

$$\varepsilon_L = dL - d_i\vartheta^i$$

where the pull-back maps have been omitted. Note that, just as the Lagrangian is a function rather than an m-form, the Euler-Lagrange form in this homogeneous context is a 1-form rather than an $(m+1)$-form. In coordinates, the latter is

$$\varepsilon_L = E_a du^a = \sum_{|I|=0}^{k} (-1)^{|I|} d_I \left(\frac{\partial L}{\partial u^a_I} \right) du^a$$

so that the coefficients E_a are, once again, the Euler-Lagrange equations.

Next, we observe that if L is non-vanishing the the m-form

$$\widehat{\Theta}_L = \frac{1}{L^{m-1}} \bigwedge\nolimits_{i=1}^{m} \vartheta^i$$

is projectable to $J^{2k-1}_+(E, m)$ exactly when $m = 1$ or $k \leq 2$, and in suitable charts on $J^{2k-1}_+(E, m)$ the projections have precisely the coordinate representations given in formulæ (2), (3b) or (4b) of the Lepage equivalent theorem ([4], Section 6). As

$$J^k\pi \subset J^k_+(E, m)\,, \qquad J^{2k-1}\pi \subset J^{2k-1}_+(E, m)$$

are open submanifolds, and as a Lagrangian form λ on $J^k\pi$ gives rise to a homogeneous Lagrangian function L on an open submanifold of $\mathcal{F}^k_{(m)}E$ by

$$L(\xi) = \langle\, \lambda_{\rho(\xi)}, \xi \,\rangle$$

where $\xi \in \mathcal{F}^k_{(m)}E$ and where $\rho : \mathcal{F}^k_{(m)}E \to J^k_+(E, m)$ is the projection, this provides one explanation for the negative result of part (5) the Lepage equivalent theorem.

We also observe that, for a first-order homogeneous Lagrangian L, the m-form

$$\widetilde{\Theta}_L = \frac{1}{m!} S^1 dS^2 d \ldots S^m dL$$

satisfies $d\widetilde{\Theta}_L = 0$ exactly when $\varepsilon_L = 0$. This m-form is projectable to $J^1_+(E, m)$, and in suitable charts the projection has precisely the coordinate representation given in formula (3c) of the Lepage equivalent theorem [5].

A final remark is that many of these objects may be expressed concisely in terms of vector-valued forms, taking their values in the vector space $\bigwedge^s \mathbb{R}^{m*}$ [38]; we denote the space of r-forms on $\mathcal{F}^k_{(m)}E$ taking their values in this vector space as $\Omega^{r,s}_{(k)}$. Identifying the coordinate function t^i on $\mathbb{R}^m$ with the constant 1-form dt^i, we define the Lagrangian vector-valued 0-form Λ, the Hilbert vector-valued 1-form Θ_L and the Euler-Lagrange vector-valued 1-form $\mathcal{E}_L$ to be

$$\Lambda = L\, d^m t \in \Omega^{0,m}_{(k)}\,, \qquad \Theta_L = \vartheta^i \otimes d^{m-1}t_i \in \Omega^{1,m-1}_{(2k-1)}\,, \qquad \mathcal{E}_L = \varepsilon_L \otimes d^m t \in \Omega^{1,m}_{(2k)}\,.$$

The relationship between these vector-valued forms may be specified using a coboundary operator $d_{\mathbf{T}} : \Omega^{r,s}_{(k)} \to \Omega^{r,s+1}_{(k+1)}$ defined in terms of the total derivative operators d_i by

$$d_{\mathbf{T}}(\theta \otimes w) = d_i\theta \otimes (dt^i \wedge w)\,, \qquad w \in \bigwedge\nolimits^s \mathbb{R}^{m*}\,.$$

This operator is globally exact for $r \geq 1$ modulo pull-backs (and, if defined on an infinite-order manifold $\mathcal{F}^{\infty}_{(m)}E$ then it would be globally exact for $r \geq 1$ without further qualification); a canonical homotopy operator P is given by

$$P(\Phi) = P^{j}_{(s)}(\phi_{i_1 \cdots i_s}) \otimes \left\{ i_{\partial/\partial t^j} \left(dt^{i_1} \wedge \ldots \wedge dt^{i_s} \right) \right\}$$

where $\Phi = \phi_{i_1 \cdots i_s} \otimes \left(dt^{i_1} \wedge \ldots \wedge dt^{i_s} \right) \in \Omega^{r,s}_{(k)}$, the $\phi_{i_1 \cdots i_s}$ are scalar r-forms, completely skew-symmetric in the indices $i_1, \ldots, i_s$, and $P^{j}_{(s)}$ is the differential operator on scalar r-forms defined by

$$P^{j}_{(s)} = \sum_{|J|=0}^{rk-1} \frac{(-1)^{|J|}(m-s)!|J|!}{r^{|J|+1}(m-s+|J|+1)!J!} d_J S^{J+1_j} \, .$$

The relationship between Λ, Θ_L and $\mathcal{E}_L$ may now be given (omitting pull-back maps) by

$$\Theta_L = P d\Lambda \, , \qquad \mathcal{E}_L = d\Lambda - d_{\mathbf{T}}\Theta_L \, .$$

4.3 The variational bicomplex

When describing the decomposition of differential r-forms on manifolds of jets of sections in terms of their s-contact components (subsection 2.5) we made a point of restricting attention to forms defined on $J^k\pi$ horizontal over $J^{k-1}\pi$. The horizontal and vertical contact structures h_k and v_k (subsection 2.5) may be combined with the exterior derivative to give two new operators, the horizontal and vertical differentials d_{h} and d_{v}, by

$$d_{\mathrm{h}} = h_k d - d h_{k-1} \, , \qquad d_{\mathrm{v}} = v_k d - d v_{k-1} \, .$$

These map r-forms on $J^{k-1}\pi$ to $(r+1)$-forms on $J^k\pi$, and satisfy

$$d_{\mathrm{h}}^2 = d_{\mathrm{v}}^2 = 0 \, , \qquad d_{\mathrm{h}} d_{\mathrm{v}} + d_{\mathrm{v}} d_{\mathrm{h}} = 0 \, .$$

By construction if the r-form χ is s-contact then so is $d_{\mathrm{h}}\chi$, whereas $d_{\mathrm{v}}\chi$ is $(s+1)$-contact, so if we write Φ^s_{r-s} for the space of s-contact r-forms then we have a bicomplex whose typical rows and columns are

$$\xrightarrow{d_{\mathrm{h}}} \Phi^s_{r-s-1} \xrightarrow{d_{\mathrm{h}}} \Phi^s_{r-s} \xrightarrow{d_{\mathrm{h}}} \Phi^s_{r-s+1} \xrightarrow{d_{\mathrm{h}}} \, ,$$

$$\xrightarrow{d_{\mathrm{v}}} \Phi^{s-1}_{r-s} \xrightarrow{d_{\mathrm{v}}} \Phi^s_{r-s} \xrightarrow{d_{\mathrm{v}}} \Phi^{s+1}_{r-s} \xrightarrow{d_{\mathrm{v}}} \, .$$

A similar construction may be carried out on the manifold of infinite jets, and here the differentials are defined on the manifold $J^\infty\pi$. This construction, augmented by some additional spaces along the edge of the diagram, is known as the *variational bicomplex*. Both horizontal and vertical differentials locally exact, although the proof for d_{h} is considerably harder than that for d_{v}. More details about the variational bicomplex, together with a somewhat different construction of a variational sequence defined on $J^k\pi$ for a fixed finite order k, may be found in the article by R. Vitolo in this Handbook [43].

We may also consider whether it is possible to approach this problem in the context of jets of submanifolds, or of non-degenerate velocities. In these cases, a form cannot be called 'horizontal', and so some other approach must be adopted. For jets of submanifolds, one approach is to consider *pseudo-horizontal forms* [27]: these are the annihilators of *pseudo-vertical vectors*, which are equivalence classes of vectors differing by total derivatives.

For the context of non-degenerate velocities, the vector valued forms described in subsection 4.2 may be used instead; the operator $d_{\mathbf{T}}$ corresponds to the horizontal differential, and the ordinary exterior derivative corresponds to the vertical differential.

5 Natural bundles

5.1 Natural bundles and geometric objects

The idea of a natural bundle is that, starting with a manifold M of some given dimension m, we can construct a fibred manifold $FM \to M$ in a local, functorial way. Thus F should be a covariant functor from the category of m-dimensional manifolds and local diffeomorphisms, to the category of fibred manifolds and fibred maps over local diffeomorphisms, such that following F with the base manifold functor gives the identity functor. The local condition is that, if we look at a non-empty open submanifold $U \subset M$, then the resulting fibred manifold $FU \to U$ should be just the restriction of FM to U. There is, in addition, a regularity condition, that a smoothly parametrised family of local diffeomorphisms should give rise to a smoothly parametrised family of fibred maps, although it may be shown that this condition is always satisfied: see, for instance, [18] Corollary. 20.7. It will be clear from the context whether we use the phrase 'natural bundle' to refer to the functor F, or to the image FM.

A geometric object at $x \in M$ is an element of the fibre of a natural bundle FM at x. Such an object might be specified in terms of coordinates; a change of coordinates on M would be represented by a local diffeomorphism, and so the corresponding change of coordinates of the object would be determined by applying the functor: this is just the classical way of describing geometric objects, as objects which 'transform in a particular way'. A basic example of a natural bundle is the tangent bundle $TM \to M$; tangent vectors are geometric objects. The cotangent bundle $T^*M \to M$ is also a natural bundle, although in a rather more complicated way, as the cotangent functor T^* itself is contravariant. Here, we have to use the fact that we consider only local diffeomorphisms f between m-dimensional manifolds, so that T^*f is an isomorphism on each fibre and we may use its inverse to give a covariant functor. In this way, cotangent vectors are also geometric objects.

Although we have stated that each natural bundle functor takes its values in the category of fibred manifolds, the name suggests that each image is actually a locally trivial bundle, and this indeed the case. The standard fibre of F is the fibre $\mathcal{S}_F$ of $F\mathbb{R}^m$ at $0 \in \mathbb{R}^m$ (recall that F acts on the category of m-dimensional manifolds); smoothness of the translation map $t_x : y \mapsto x + y$ on $\mathbb{R}^m$ and the regularity condition then show that $F\mathbb{R}^m \cong \mathbb{R}^m \times \mathcal{S}_F$ is globally trivial. The local condition, and the existence of charts $U \to \mathbb{R}^m$ with $U \subset M$ open, then show that FM is locally trivial.

5.2 The order of a natural bundle

We can describe the *order* of a natural bundle by considering how much information in a local diffeomorphism is preserved by the action of the functor. We say that the order of F is at most k if, for each pair of local diffeomorphisms $f, g : M_1 \to M_2$ and for each $x \in M_1$, the restrictions of Ff and Fg to the fibre of FM_1 at x are equal whenever $j^k_x f_1 = j^k_x f_2$. The order of F is then the least such k for which this condition holds. So, for example, both the tangent bundle and the cotangent bundle are first-order natural bundles.

It is well-known that the tangent and cotangent bundles TM, T^*M are associated bundles of the frame bundle $\mathcal{F}M$. In fact every first-order natural bundle is an associated bundle of $\mathcal{F}M$, and indeed every k-th order natural bundle is an associated bundle of $\mathcal{F}^k M$, the k-th order frame bundle (see [18], Proposition 14.5). To see how this arises, note that the definition of order implies that there is a map

$$L^k_m \times_{\mathbb{R}^m} F\mathbb{R}^m \to F\mathbb{R}^m \,, \qquad (j^k_0 f, (0, s)) \mapsto Ff(0, s)$$

for any local diffeomorphism of $\mathbb{R}^m$ defined near zero, where we have written an element of $F\mathbb{R}^m$ projecting to zero as $(0, s) \in \mathbb{R}^m \times \mathcal{S}_F$; we thus obtain a left action of L^k_m on the standard fibre $\mathcal{S}_F$, and so obtain an associated bundle in the usual way.

Of course we cannot use this description of natural bundles in the case of infinite order; but it has been shown this case never arises, and that the order of every natural bundle is finite ([32]; see also [18] Theorem 22.3).

5.3 Natural operators

A natural operator is a family of regular operators (subsection 3.1) between two natural bundles satisfying certain conditions. If Δ is a natural operator between the natural bundles F_1, F_2 then, for each manifold M, we require Δ_M to be a regular operator from sections of $F_1 M$ to sections of $F_2 M$. We expect Δ to behave correctly with respect to diffeomorphic manifolds M_1, M_2 so that

$$\Delta_{M_2}(F_1 f \circ \phi \circ f^{-1}) = F_2 f \circ \Delta_{M_1} \phi \circ f^{-1}$$

as sections of $F_2 M_2 \to M_2$, whenever ϕ is a section of $F_1 M_1 \to M_1$ and $f : M_1 \to M_2$ is a diffeomorphism. We also require a local condition, that if $U \subset M$ is an open submanifold then

$$\Delta_U(\phi|_U) = \Delta_M \phi|_U$$

for every section ϕ of $F_1 M \to M$. These two conditions taken together show that a natural operator is represented by regular operators whose coordinate representations are independent of the choice of chart.

One of the most useful ways of classifying natural operators involves the following result.

Theorem 5.1 (see [18], Theorem 14.18) *Let F_1, F_2 be two natural bundles defined on m-dimensional manifolds of finite orders k_1, k_2 with standard fibres $\mathcal{S}_{F_1}$, $\mathcal{S}_{F_2}$. There is a canonical bijective correspondence between the set of all l-th order natural operators from F_1 to F_2 and the set of all smooth L^p_m-equivariant maps between $T^l_m \mathcal{S}_{F_1}$ and $\mathcal{S}_{F_2}$, where $p = \max\{k_1 + l, k_2\}$.*

As an example, we consider the exterior derivative acting on r-forms with $r \geq 1$. This is a first-order natural operator between the first-order natural bundles $\bigwedge^r T^*$ and $\bigwedge^{r+1} T^*$, and so corresponds to a L^2_m-equivariant map $T^1_m \bigwedge^r \mathbb{R}^{m*} \to \bigwedge^{r+1} \mathbb{R}^{m*}$; it is possible to show that any natural operator between these natural bundles is first-order, and is necessarily a constant multiple of the exterior derivative [31]. First, we need to see that any such natural operator has finite order; this is a consequence of some analytical arguments ([18], Proposition 23.5 and Example 23.6). Next, we use the correspondence between l-th order natural operators between these two natural bundles and L^{l+1}_m-equivariant maps; by considering the action of such a map on the jet coordinates and choosing suitable elements of L^{l+1}_m, it may be shown that the map depends only on the first-order jet coordinates, that this dependence is linear, and is just a constant multiple of the equivariant map corresponding to the exterior derivative ([18], Proposition 25.4).

5.4 Gauge natural bundles

A generalisation of the idea of a natural bundle arises when we consider, instead of just a manifold M, an arbitrary principal G-bundle $\pi : P \to M$ for some Lie group G [6]. Once again we fix the dimension $m = \dim M$. A *gauge natural bundle* is then a functor F such that

(1) every principal bundle $\pi : P \to M$ with $\dim M = m$ is transformed to a fibred manifold $F\pi : FP \to M$;

(2) every principal morphism $(f : P_1 \to P_2, \bar{f} : M_1 \to M_2)$ is transformed to a fibred map $(Ff : FP_1 \to FP_2, \bar{f} : M_1 \to M_2)$;

(3) for every non-empty open set $U \subset M_1$ the transform Fi of the inclusion principal morphism $(i : \pi^{-1}(U) \to P, \bar{\imath} : U \to M)$ is equal to the inclusion of fibred manifolds $((F\pi)^{-1}(U) \to FP, U \to M)$.

The order of the gauge-natural bundle may be described by considering principal morphisms. Suppose that two such morphisms $f, g : P_1 \to P_2$ have the same k-jet $j^k_y f = j^k_y g$ at some point $y \in P_1$; then by the group action they have the same k-jet at any point of the fibre at $x = \pi_1(y)$. We say that the order of F is at most k if, for each $x \in M$, the restrictions of Ff and Fg to the fibre of FP_1 at x are equal whenever the k-jets of f and g are equal at any point of the fibre of P_1 at x. The order of F is then the least such k for which this condition holds. We also say that the gauge-natural bundle is regular if a smoothly parametrised family of principal morphisms transforms to a smoothly parametrised family of maps of fibred manifolds. As with natural bundles, it may be shown that every gauge-natural bundle is regular ([18], Proposition 51.10), and that every gauge-natural bundle has finite order ([18], Proposition 51.7).

Bibliography

[1] D. Bao, S.-S. Chern and Z. Shen: *An introduction to Riemann-Finsler geometry* (Graduate Texts in Mathematics 200, Berlin: Springer-Verlag, 2000)

[2] D. E. Betounes: Extension of the classical Cartan form *Phys. Rev. D* **29** (1984) 599–606

[3] C. Carathéodory: Über die Variationsrechnung bei mehrfachen Integralen *Acta Szeged Sect. Scient. Mathem.* **4** (1929) 193–216

[4] M. Crampin, D. J. Saunders: The Hilbert-Carathéodory and Poincaré-Cartan forms for higher-order multiple-integral variational problems *Houston J. Math.* **30(3)** (2004) 657–689

[5] M. Crampin, D. J. Saunders: On null Lagrangians *Diff. Geom. Appl.* **22(2)** (2005) 131–146

[6] D. J. Eck: *Gauge-natural bundles and generalized gauge theories* (Mem. Amer. Math. Soc. 247, 1981)

[7] C. Ehresmann: Les Prolongements d'une Variété Différentiable: Calcul des Jets, Prolongement Principal *C. R. Acad. Sci. Paris* **233** (1951) 598–600

[8] C. Ehresmann: Les Prolongements d'une Variété Différentiable: L'espace des Jets d'ordre r de V_n dans V_m *C. R. Acad. Sci. Paris* **233** (1951) 777–779

[9] C. Ehresmann: Les Prolongements d'une Variété Différentiable: Transitivité des Prolongements *C. R. Acad. Sci. Paris* **233** (1951) 1081–1083

[10] C. Ehresmann: Les Prolongements d'une Variété Différentiable: éléments de Contact et élements d'Enveloppe *C. R. Acad. Sci. Paris* **234** (1952) 1028–1030

[11] C. Ehresmann: Les Prolongements d'une Variété Différentiable: Covariants Différentiels et Prolongements d'une Structure Infinitésimale *C. R. Acad. Sci. Paris* **234** (1952) 1424–1425

[12] H. Goldschmidt: Integrability criteria for systems of non-linear partial differential equations *J. Diff. Geom.* **1** (1969) 269–307

[13] D. R. Grigore: Lagrangian formalism on Grassmann manifolds *In: Handbook of Global Analysis* (Amsterdam: Elsevier, 2007) 327–373

[14] V. W. Guillemin and S. Sternberg: An algebraic model of transitive differential geometry *Bull. Amer. Math. Soc.* **70** (1964) 16–47

[15] L. Hörmander: *Linear partial differential operators* (Grundl. Math. Wissenschaft 116, Springer, 1966)

[16] M. Horák and I. Kolář: On the higher-order Poincaré-Cartan forms *Czech. Math. J.* **33** (1983) 467–475

[17] I. Kolář: Weil bundles as generalized jet spaces *In: Handbook of Global Analysis* (Amsterdam: Elsevier, 2007) 625–664

[18] I. Kolář, P. W. Michor P. W. and J. Slovák: *Natural Operations in Differential Geometry* (Berlin: Springer-Verlag, 1993)

[19] B. Kruglikov and V. Lychagin: Geometry of differential equations *In: Handbook of Global Analysis* (Amsterdam: Elsevier, 2007) 725–771

[20] D. Krupka: A map associated to the Lepagean forms of the calculus of variations in fibered manifolds *Czech. Math. J.* **27** (1977) 114–118

[21] D. Krupka: Global variational theory in fibered spaces *In: Handbook of Global Analysis* (Amsterdam: Elsevier, 2007) 773–836

[22] O. Krupková and G. E. Prince : Second Order Ordinary Differential Equations in Jet Bundles and the Inverse Problem of the Calculus of Variations *In: Handbook of Global Analysis* (Amsterdam: Elsevier, 2007) 837–904

[23] M. Kuranishi: On E. Cartan's prolongation theorem of exterior differential systems *Amer. J. Math.* **79** (1957) 1–47

[24] L. A. Lambe and W. M. Seiler: Differential equations, Spencer cohomology, and computing resolutions *Georgian Math. J.* **9(4)** (2002) 723–772

[25] H. Lewy: An example of a smooth linear partial differential equation without solution *Ann. Math.* **66(1)** (1957) 155–158

[26] L. Mangiarotti and M. Modugno: Fibred Spaces, Jet Spaces and Connections *In: Proceedings of the Meeting "Geometry and Physics" October 12–15 1982* (Bologna: Pitagora Editrice, 1982) 135–165

[27] G. Manno and R. Vitolo: Variational sequences on finite order jets of submanifolds *In: Proceedings of the 8th International Conference on Differential Geometry and its Applications* (Silesian University at Opava, 2001) 435–447

[28] N. Newns and A. Walker: Tangent planes to a differentiable manifold *J. London Math. Soc.* **31** (1956) 400–407

[29] A. Nijenhuis: Natural bundles and their general properties *In: Differential Geometry in Honor of K. Yano* (Tokyo: Kinokuniya, 1972) 317–334

[30] P. J. Olver: Equivalence and the Cartan form *Acta Applicandae Math* **31** (1993) 99–136

[31] R. S. Palais: Natural operations on differential forms *Trans. Amer. Math. Soc.* **92** (1959) 125–141

[32] R. S. Palais and C. L. Terng: Natural bundles have finite order *Topology* **16** (1977) 271–277

[33] J. Peetre: Une caractérisation abstraite des opérateurs différentiels *Math. Scand.* **7** (1959) 211–218

[34] J. Peetre: Rectifications à l'article "Une caractérisation abstraite des opérateurs différentiels" *Math. Scand.* **8** (1960) 116–120

[35] J. F. Pommaret: *Systems of partial differential equations and Lie pseudogroups* (New York: Gordon & Breach, 1978)

[36] D. J. Saunders: *The geometry of jet bundles* (Lecture Notes in Mathematics 142: Cambridge University Press, 1989)

[37] D. J. Saunders: A note on Legendre transformations *Diff. Geom. Appl.* **1** (1991) 109–122

[38] D. J. Saunders: Homogeneous variational complexes and bicomplexes *Preprint* `arXiv:math.DG/0512383` (2005)

[39] J. Slovák: Peetre theorem for nonlinear operators *Ann. Global. Anal. Geom.* **6** (1988) 273–283

[40] D. C. Spencer: Overdetermined systems of linear partial differential equations *Bull. Amer. Math. Soc.* **75** (1969) 179–239

[41] J. Szilasi and R. L. Lovas: Some aspects of differential theories *In: Handbook of Global Analysis* (Amsterdam: Elsevier, 2007) 1069–1114

[42] F. Treves: *Topological vector spaces, distributions and kernels* (Academic Press, 1967)

[43] R. Vitolo: Variational sequences *In: Handbook of Global Analysis* (Amsterdam: Elsevier, 2007) 1115–1163

[44] A. Weil: Théorie des points proches sur les variétes différentielles *In: Colloque de topol. et géom. diff.* (Strasbourg, 1953) 111–117

D. J. SAUNDERS

Department of Algebra and Geometry, Palacký University,
Tomkova 40, 779 00 Olomouc, Czech Republic
Address for correspondence:
30 Little Horwood Road, Great Horwood, Milton Keynes, MK17 0QE, UK

E-mail: `david@symplectic.demon.co.uk`

Handbook of Global Analysis
Demeter Krupka and David Saunders

Some aspects of differential theories

Respectfully dedicated to the memory of Serge Lang

József Szilasi and Rezső L. Lovas

Contents

Introduction

Global analysis is a particular amalgamation of topology, geometry and analysis, with strong physical motivations in its roots, its development and its perspectives. To formulate a problem in global analysis exactly, we need

(1) a base manifold,

(2) suitable fibre bundles over the base manifold,

(3) a differential operator between the topological vector spaces consisting of the sections of the chosen bundles.

In quantum physical applications, the base manifold plays the role of space-time; its points represent the location of the particles. The particles obey the laws of quantum physics, which are encoded in the vector space structure attached to the particles in the mathematical model. A particle carries this vector space structure with itself as it moves. Thus we arrive at the intuitive notion of vector bundle, which had arisen as the 'repère mobile' in Élie Cartan's works a few years *before* quantum theory was discovered.

To describe a system (e.g. in quantum physics) in the geometric framework of vector bundles effectively, we need a suitably flexible differential calculus. We have to differentiate vectors which change smoothly together with the vector space carrying them. The primitive idea of differentiating the coordinates of the vector in a fixed basis is obviously not satisfactory, since there is no intrinsic coordinate system which could guarantee the invariance of the results. These difficulties were traditionally solved by classical tensor calculus, whose most sophisticated version can be found in Schouten's 'Ricci-Calculus'. From there, the theory could only develop to the global direction, from the debauch of indices to totally index-free calculus. This may be well illustrated by A. Nijenhuis' activity.

It had become essentially clear by the second decade of the last century that, for a coordinate-invariant tensorial differential calculus, we need a structure establishing an isomorphism between vector spaces at different points. Such a structure is called a *connection*. A connection was first constructed by Levi-Civita in the framework of Riemannian geometry by defining a *parallel transport* between the tangent spaces at two points of the base manifold along a smooth arc connecting the given points. This makes it possible to form a difference quotient and to differentiate vector and tensor fields along the curve. The differentiation procedure so defined is *covariant differentiation*. Thus it also becomes clear that

connection, *parallel transport*, *covariant differentiation*

are essentially equivalent notions: these are the same object, from different points of view.

The history sketched in the foregoing is of course well-known, and technical details are nowadays available in dozens of excellent monographs and textbooks. Somewhat paradoxically, *purely historical* features (the exact original sources of main ideas, the evolution of main streamlines) are not clear in every detail, and they would be worth of a more profound study. In our present work, we would like to sketch some aspects of the rich theory of differentiation on manifolds which are less traditional and less known, but which definitely seem to be progressive. One of the powerful trends nowadays is the globalization of calculus on infinite-dimensional non-Banach topological vector spaces, i.e., its transplantation onto manifolds, and a formulation of a corresponding Lie theory. The spectrum becomes more colourful (and the theory less transparent) by the circumstance that we see contesting calculi even on a local level. One corner stone in this direction is definitely A. Kriegl and P. W. Michor's truly monumental monograph [18], which establishes the theory of calculus in so-called convenient vector spaces. Differentiation theory in convenient vector spaces is outlined in the contribution of J. Margalef-Roig and E. Outerelo Domínguez in this volume. Another promising approach is Michal – Bastiani differential calculus, which became known mainly due to J. Milnor and R. Hamilton. We have chosen this way, drawing much from H. Glöckner and K.-H. Neeb's monograph is preparation, which will contain a thorough and exhaustive account of this calculus.

We would like to emphasize that the fundamental notions and techniques of 'infinite dimensional analysis' can be spared neither by those who study analysis 'only' on finite-dimensional manifolds. The reason for this is very simple: even the vector space of smooth functions on an open subset of $\mathbb{R}^n$ is infinite-dimensional, and the most natural structure with which it may be endowed is a suitable multinorm which makes it a Fréchet space. Accordingly, we begin our treatment by a review of some basic notions and facts in connection with topological vector spaces. The presentation is organized in such a way that we may provide a non-trivial application by a detailed proof of Peetre's theorem. This famous theorem characterizes linear differential operators as support-decreasing $\mathbb{R}$-linear maps, and it is of course well-known, together with its various proofs. The source of one of the standard proofs is Narasimhan's book [29]; Helgason takes over essentially his proof [15]. Although Narasimhan's proof is very clear in its main features, our experience is that understanding it in every fine detail requires serious intellectual efforts. Therefore we thought that a detailed treatment of Narasimhan's line of thought could be useful by filling in the wider logical gaps. Thus, besides presenting fundamental techniques, we shall also have a possibility to demonstrate hidden subtleties of these sophisticated (at first sight mysterious) constructions.

We discuss Peetre's theorem in a local framework, its transplantation to vector bundles, however, does not raise any difficulty. After climbing this first peak, we sketch the main steps towards the globalization of Michal – Bastiani differential calculus, on the level of basic notions and constructions, on a manifold modeled on a locally convex topological vector space.

In the last section, for simplicity's sake, we return to finite dimension, and we discuss a special problem about covariant derivatives. Due to Élie Cartan's activity, it has been known since the 1930s that a covariant derivative compatible with a Finsler structure cannot be constructed on the base manifold (more precisely, on its tangent bundle), and the velocity-dependent character of the objects makes it necessary to start from a so-called line element bundle. In a contemporary language: the introduction of a covariant derivative, analogous to Levi-Civita's, metrical with respect to the Finsler structure, is only possible in the pull-back of the tangent bundle over itself, or in some 'equivalent' fibre bundle. We have to note that it was a long and tedious way from Élie Cartan's intuitively very clear but conceptually rather obscure construction to today's strict formulations, and this way was paved, to a great extent, by the demand for understanding Élie Cartan. In the meantime, in 1943, another 'half-metrical' covariant derivative in Finsler geometry was discovered by S. S. Chern (rediscovered by Hanno Rund in 1951). This covariant derivative was essentially in a state of suspended animation until the 1990s. Then, however, came a turning point. Due to Chern's renewed activity and D. Bao and Z. Shen's work, Chern's covariant derivative became one of the most important tools of those working in this field. We wanted to understand which were the properties of Chern's derivative which could give priority to it over other covariant derivatives used in Finsler geometry (if there are such properties). As we have mentioned, Chern's connection is only half-metrical: covariant derivatives of the metric tensor arising from the Finsler structure in vertical directions do not vanish in general. 'In return', however, the derivative is 'vertically natural': it induces the natural parallelism of vector spaces on the fibres. This property makes it possible to interpret Chern's derivative as a covariant derivative *given on the base manifold*, parametrized locally by a nowhere vanishing vector field. We think that this possibility of interpretation does distinguish Chern's derivative in some sense. As for genuine applications of Chern's connection, we refer to T. Aikou and L. Kozma's study in this volume.

Notation

As usual, $\mathbb{R}$ and $\mathbb{C}$ will denote the fields of real and complex numbers, respectively. $\mathbb{N}$ stands for the 'half-ring' of natural numbers (integers $\geqq 0$). If $\mathbb{K}$ if one of these number systems, then $\mathbb{K}^* := \mathbb{K} \setminus \{0\}$. $\mathbb{R}_+ := \{z \in \mathbb{R} | z \geqq 0\}$, $\mathbb{R}_+^* := \mathbb{R}_+ \cap \mathbb{R}^*$. A mapping from a set into $\mathbb{R}$ or $\mathbb{C}$ will be called a *function*.

Discussing functions defined on a subset of $\mathbb{R}^n$, it will be convenient to use the multi-index notation $\alpha = (\alpha_1, \dots, \alpha_n)$, where $\alpha_i \in \mathbb{N}$, $1 \leqq i \leqq n$. We agree that

$$|\alpha| := \alpha_1 + \ldots + \alpha_n, \quad \alpha! := \alpha_1! \ldots \alpha_n!.$$

If $\beta := (\beta_1, \dots, \beta_n) \in \mathbb{N}^n$, and $\beta_j \leqq \alpha_j$ for all $j \in \{1, \dots, n\}$, we write $\beta \leqq \alpha$ and define

$$\alpha - \beta := (\alpha_1 - \beta_1, \dots, \alpha_n - \beta_n), \qquad \binom{\alpha}{\beta} := \frac{\alpha!}{(\alpha - \beta)!\beta!}.$$

If V and W are vector spaces over a field $\mathbb{F}$, then $\mathcal{L}_{\mathbb{F}}(V, W)$ or simply $\mathcal{L}(V, W)$ denotes the vector space of all linear mappings from V into W, and $V^* := \mathcal{L}(V, \mathbb{F})$ is the (algebraic) dual of V. If $k \in \mathbb{N}^*$, $\mathcal{L}^k(V, W)$ is the vector space of all k-multilinear mappings $\varphi : V \times \dots \times V \to W$. We note that $\mathcal{L}\left(V, \mathcal{L}^k(V, W)\right)$ is canonically isomorphic to $\mathcal{L}^{k+1}(V, W)$. We use an analogous notation if, more generally, V and W are modules over the same commutative ring.

1 Background

Topology

We assume the reader is familiar with the rudiments of point set topology, so the meaning of such elementary terms as open and closed set, neighbourhood, connectedness, Hausdorff topology, (open) covering, first and second countability, compactness and local compactness, continuity, homeomorphism, ... does not demand an explanation. For the sake of definiteness, we are going to follow, as closely as feasible, the convention of Dugundji's Topology [7]. Thus by a *neighbourhood* of a point or a set in a topological space we shall always mean an *open* subset containing the point or subset, and we include in the definition of second countability, compactness and local compactness the requirement that the topology is Hausdorff. This subsection serves to fix basic terminology and notation, as well as to collect some more subtle topological ideas which may be beyond the usual knowledge of non-specialists.

We denote by $\mathcal{N}(p)$ the set of all neighbourhoods of a point p in a topological space. A subset $\mathcal{F} \subset \mathcal{N}(p)$ is said to be a *fundamental system* of p if for every $\mathcal{V} \in \mathcal{N}(p)$ there exists $\mathcal{U} \in \mathcal{F}$ such that $\mathcal{U} \subset \mathcal{V}$. If S is a topological space, and $A \subset S$, then $\overset{\circ}{A}$, $\bar{A}$ and ∂A denote the interior, the closure and the boundary of A, resp. If G is an Abelian group, and $f : S \to G$ is a mapping, then

$$\operatorname{supp}(f) := \overline{\{p \in S | f(p) \neq 0\}}$$

is the *support* of f.

Let M be a metric space with distance function $\varrho : M \times M \to \mathbb{R}$. If $a \in M$, and $r \in \mathbb{R}_+^*$, the set

$$B_r^{\varrho}(a) := \{p \in M | \varrho(a, p) < r\}$$

is called the *open ϱ-ball* of centre a and radius r. (We shall omit the distinguishing ϱ whenever the distance function is clear from the context.) By declaring a subset of M to be open if it is a (possibly empty) union of open balls, a topology is obtained on M, called the *metric topology* of M, or the topology *induced by the distance function* ϱ. Unless otherwise stated, a metric space will always be topologized by its metric topology. Conversely, if a topology $\mathcal{T}$ on a set S is induced by a distance function $\varrho : S \times S \to \mathbb{R}$, then $\mathcal{T}$ and ϱ are called *compatible*, and $\mathcal{T}$ is said to be *metrizable*.

By far the most important metric space is the *Euclidean n-space* $\mathbb{R}^n$, the set of all n-tuples $v = (\nu^1, \dots, \nu^n)$ endowed with the *Euclidean distance* defined by

$$\varrho_E(a,b) := \|a-b\| = \langle a-b, a-b\rangle^{1/2} \quad (a, b \in \mathbb{R}^n),$$

where $\langle\ ,\ \rangle$ is the canonical scalar product in $\mathbb{R}^n$, and $\|\cdot\|$ is *Euclidean norm* arising from $\langle\ ,\ \rangle$. The metric topology of $\mathbb{R}^n$ is called the *Euclidean topology* of $\mathbb{R}^n$. We shall assume that $\mathbb{R}^n$ (in particular, $\mathbb{R} = \mathbb{R}^1$) is topologized with the Euclidean topology.

Let $\mathcal{U}$ be an open subset of $\mathbb{R}^n$. A sequence $(K_i)_{i\in\mathbb{N}^*}$ of compact subsets of $\mathcal{U}$ is said to be a *compact exhaustion* of $\mathcal{U}$ if

$$K_i \subset \overset{\circ}{K}_{i+1} \quad (i \in \mathbb{N}^*) \quad \text{and} \quad \mathcal{U} = \bigcup_{i\in\mathbb{N}^*} K_i.$$

Lemma 1.1 *There does exist a compact exhaustion for any open subset of* $\mathbb{R}^n$.

Proof. Let a (nonempty) open subset $\mathcal{U}$ of $\mathbb{R}^n$ be given. For every positive integer m, define a subset K_m of $\mathcal{U}$ as follows:

$$K_m := \left\{ p \in \mathcal{U} \,\middle|\, \varrho_E(p, \partial\mathcal{U}) \geqq \frac{1}{m} \right\} \cap \overline{B_m(0)}$$

(by convention, $\varrho_E(p, \partial\mathcal{U}) := \infty$ if $\partial\mathcal{U} = \emptyset$). It is then clear that each set K_m is compact, $K_m \subset \overset{\circ}{K}_{m+1}$ $(m \in \mathbb{N}^*)$, and $\cup_{i\in\mathbb{N}^*} K_i = \mathcal{U}$. □

Now we recall some more delicate concepts and facts of point set topology.

A Hausdorff space is said to be a *Lindelöf space* if each open covering of the space contains a *countable* covering. By a theorem of Lindelöf [7, Ch. VIII, 6.3] *all second countable spaces are Lindelöf* (but the converse is not true!).

A locally compact space is called *σ-compact* if it can be expressed as the union of a sequence of its compact subsets. It can be shown (see [7, Ch. XI, 7.2]) that *a topological space is σ-compact, if and only if, it is a locally compact Lindelöf space*.

Let S be a topological space. An open covering $(\mathcal{U}_\alpha)_{\alpha\in A}$ of S is said to be *locally finite* (or *nbd-finite*) if each point of S has a nbd $\mathcal{U}$ such that $\mathcal{U} \cap \mathcal{U}_\alpha \neq \emptyset$ for at most finitely many indices α. If $(\mathcal{U}_\alpha)_{\alpha\in A}$ and $(\mathcal{V}_\beta)_{\beta\in B}$ are two coverings of S, then $(\mathcal{U}_\alpha)$ is a *refinement* of $(\mathcal{V}_\beta)$ if for each $\alpha \in A$ there is some $\beta \in B$ such that $\mathcal{U}_\alpha \subset \mathcal{V}_\beta$. A Hausdorff space is said to be *paracompact* if each open covering of the space has an open locally finite refinement.

The following result has important applications in analysis.

Lemma 1.2 *Any σ-compact topological space, in particular any second countable locally compact topological space, is paracompact.*

Proof. The statement is a fairly immediate consequence of some general topological facts. Namely, the σ-compact spaces, as we have remarked above, are just the locally compact Lindelöf spaces. Locally compact spaces are regular: each point and closed set not containing the point have disjoint nbds. Since by a theorem of K. Morita [7, VIII, 6.5] *in Lindelöf spaces regularity and paracompactness are equivalent concepts*, we get the result. □

It is well-known that the classical concept of convergence of sequences is not sufficient for the purposes of analysis. Nets, which are generalizations of sequences, and their convergence provide an efficient tool to handle a wider class of problems. To define nets, let us first recall that a *directed set* is a set A endowed with a partial ordering $\leqq$ such that for any two elements $\alpha, \beta \in A$ there is an element $\gamma \in A$ such that $\gamma \geqq \alpha$ and $\gamma \geqq \beta$. A *net* in a set S is a family $(s_\alpha)_{\alpha \in A}$ of elements of S, i.e., a mapping $s : A \to S$, $\alpha \mapsto s(\alpha) =: s_\alpha$, where A is a directed set. Obviously, any sequence $s : n \in \mathbb{N} \mapsto s(n) =: s_n \in S$ is a net, since $\mathbb{N}$ is a directed set with its usual ordering.

A net $(s_\alpha)_{\alpha \in A}$ in a *topological space* S is said to *converge* (or to *tend*) to a point $p \in S$ if for every $\mathcal{U} \in \mathcal{N}(p)$ there is a $\beta \in A$ such that

$$s_\alpha \in \mathcal{U} \textit{ whenever } \alpha \geqq \beta.$$

Then we use the standard notation $p = \lim_{\alpha \in A} s_\alpha$, and we say that $(s_\alpha)_{\alpha \in A}$ is *convergent* in S and has a *limit* $p \in S$.

Lemma 1.3 *A topological space is Hausdorff if and only if any two limits of any convergent net are equal.*

Indication of proof. Let S be a topological space. It can immediately be seen that if S is Hausdorff, then any convergent net in S has a unique limit. Conversely, suppose that S is not Hausdorff. Let $p, q \in S$ be two points which cannot be separated by open sets. Consider the directed set A whose elements are ordered pairs $\alpha = (\mathcal{U}, \mathcal{V})$ where $\mathcal{U} \in \mathcal{N}(p)$, $\mathcal{V} \in \mathcal{N}(q)$ with the partial ordering

$$(\mathcal{U}, \mathcal{V}) \geqq (\mathcal{U}_1, \mathcal{V}_1) :\Longleftrightarrow (\mathcal{U} \subset \mathcal{U}_1 \text{ and } \mathcal{V} \subset \mathcal{V}_1).$$

For any $\alpha = (\mathcal{U}, \mathcal{V})$ let s_α be some point of $\mathcal{U} \cap \mathcal{V}$. Then the mapping $s : \alpha \in A \mapsto s_\alpha \in S$ is a net, and it is easy to check that s converges to both p and q. □

The next result shows that nets are sufficient to control continuity.

Lemma 1.4 *A mapping $\varphi : S \to T$ between two topological spaces is continuous if and only if for every net $s : A \to S$ converging to $p \in S$, the net $\varphi \circ s : A \to T$ converges to $\varphi(p)$.*

For a (quite immediate) proof see e.g. [5, Ch. I, 6.6].

If S and T are topological spaces, then the set of continuous mappings of S into T will be denoted by $C(S, T)$. In particular, $C(S) = C^0(S) := C(S, \mathbb{R})$, and

$$C_c(S) := \{f \in C(S) | \operatorname{supp}(f) \text{ is compact}\}.$$

Topological vector spaces

When Banach published his famous book '*Théorie des operations linéaires*' in 1932, it was the opinion that normed spaces provide a sufficiently wide framework to comprehend all interesting concrete problems of analysis. It turned out, however, in a short time that this is an illusion: a number of (non-artificial) problems of analysis lead to infinite-dimensional vector spaces whose topology cannot be derived from a norm. In this subsection we recall the most basic definitions and facts concerning topological vector spaces, especially locally convex spaces. We restrict ourselves to real vector spaces, although vector spaces over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ can also be treated without any extra difficulties.

If V is a (real) vector space, then a subset H of V is said to be *convex* if for each $t \in [0,1]$, $tH + (1-t)H \subset H$;

balanced if $\alpha H \subset H$ whenever $\alpha \in [-1,1]$;

absorbing if $\cup_{\lambda \in \mathbb{R}_+^*} \lambda H = V$, i.e., for all $v \in V$ there is a positive real number λ (depending on v) such that $v \in \lambda H$.

By a *topological vector space* (TVS) we mean a real vector space V endowed with a Hausdorff topology compatible with the vector space structure of V in the sense that the addition map $V \times V \to V$, $(u,v) \mapsto u+v$ and the scalar multiplication $\mathbb{R} \times V \to V$, $(\lambda, v) \mapsto \lambda v$ are continuous. (The product spaces are equipped with the product topology.) Such a topology is called a *linear topology* or *vector topology* on V.

Let V be a topological vector space. For each $a \in V$ the *translation* $T_a : v \in V \mapsto T_a(v) := a + v \in V$, and for each $\lambda \in \mathbb{R}^*$ the *homothety* $h_\lambda : v \in V \mapsto \lambda v \in V$ are homeomorphisms. This simple observation is very important: it implies, roughly speaking, that the linear topology *looks like the same* at any point. Thus, practically, in a TVS it is enough to define a local concept or to prove a local property only in a neighbourhood of the origin. We agree that in TVS context *a fundamental system will always mean a fundamental system of the origin*. Thus $\mathcal{F}$ is a fundamental system in V if $\mathcal{F} \subset \mathcal{N}(0)$, and every neighbourhood of 0 (briefly 0-neighbourhood) contains a member of $\mathcal{F}$.

In the category TVS the isomorphisms are the *toplinear isomorphisms*: linear isomorphisms which are homeomorphisms at the same time. As the following classical result shows, the structure of finite dimensional topological vector spaces is the simplest possible.

Lemma 1.5 (A. Tychonoff) *If V is an n-dimensional topological vector space, then there is a toplinear isomorphism from V onto the Euclidean n-space $\mathbb{R}^n$.*

The *idea of proof* is immediate: we choose a basis $(v_1, \dots, v_n)$ of V, and we show that the linear isomorphism

$$\left(\nu^1, \dots, \nu^n\right) \in \mathbb{R}^n \mapsto \sum_{i=1}^{n} \nu^i v_i \in V,$$

which is obviously continuous, is an *open* map. For details see e.g. [7, p. 413] or [36, p. 28].

Now we return to a generic topological vector space V. A subset H of V is said to be *bounded* if for every 0-neighbourhood $\mathcal{U}$ there is a positive real number ε such that $\varepsilon H \subset \mathcal{U}$. V is called *locally convex* if it has a fundamental system whose members are convex sets, i.e., every 0-neighbourhood contains a convex 0-neighbourhood. V has the *Heine – Borel property* if every closed and bounded subset of V is compact.

Lemma 1.6 *If a topological vector space V has a countable fundamental system, then V is metrizable. More precisely, there is a distance function ϱ on V such that*

(i) ϱ is compatible with the linear topology of V;

(ii) ϱ is translation invariant, i.e., $\varrho(T_a(u), T_a(v)) = \varrho(u,v)$ for all $a, u, v \in V$;

(iii) the open ϱ-balls centred at the origin are balanced.

If, in addition, V *is locally convex, then* ϱ *can be chosen so as to satisfy* (i)–(iii)*, and also*

(iv) all open ϱ*-balls are convex.*

For a proof we refer to [33, I, 1.24].

A net $(v_\alpha)_{\alpha\in A}$ in a locally convex space V is said to be a *Cauchy net* if for every 0-neighbourhood $\mathcal{U}$ in V there is an $\alpha_0 \in A$ such that

$$v_\alpha - v_\beta \in \mathcal{U} \textit{ whenever } \alpha, \beta \geqq \alpha_0.$$

In particular, a sequence in V is called a *Cauchy sequence* if it is a Cauchy net. Note that if the topology of V is compatible with a translation invariant distance function $\varrho : V \times V \to \mathbb{R}$, then a sequence $(v_n)_{n\in\mathbb{N}^*}$ in V is a Cauchy sequence if and only if it is a Cauchy sequence in metrical sense, i.e., for every $\varepsilon \in \mathbb{R}_+^*$ there is an integer n_0 such that $\varrho(v_m, v_n) < \varepsilon$ whenever $m > n_0$ and $n > n_0$. A locally convex space V is said to be *complete*, resp. *sequentially complete* if any Cauchy net, resp. Cauchy sequence is convergent in V. These completeness concepts coincide if V is metrizable, i.e., we have

Lemma 1.7 *A metrizable locally convex space is complete if and only if it is sequentially complete.*

As for the *proof*, the only technical difficulty is to check that sequential completeness implies completeness, i.e., the convergence of every Cauchy net. For a detailed reasoning we refer to [8, B.6.2].

It follows from our preceding remarks that if the topology of a locally convex space is induced by a translation invariant distance function, then the space is complete if and only if it is complete as a metric space. TVSs sharing these properties deserve an own name: a locally convex space is said to be a *Fréchet space* if its linear topology is induced by a complete, translation invariant distance function. We shall see in the next chapter that important examples of Fréchet spaces occur even in the context of classical analysis. In the rest of this chapter we are going to indicate how one can construct locally convex spaces, in particular Fréchet spaces, starting from a family of seminorms.

We recall that a *seminorm* on a real vector space V is a function $\nu : V \to \mathbb{R}$, satisfying the following axioms:

$\nu(u + v) \leqq \nu(u) + \nu(v)$ for all $u, v \in V$ (*subadditivity*);

$\nu(\lambda v) = |\lambda|\nu(v)$ for all $\lambda \in \mathbb{R}$, $v \in V$ (*absolute homogeneity*).

Then it follows that $\nu(0) = 0$, and $\nu(v) \geqq 0$ for all $v \in V$. If, in addition, $\nu(v) = 0$ implies $v = 0$, then ν is a *norm* on V. As in the case of metric spaces, given a point $a \in V$ and a positive real number r, we use the notation

$$B_r^\nu(a) := \{v \in V | \nu(v - a) < r\}$$

and the term 'open ν-ball with centre a and radius r'. It can be seen immediately that the 'open unit ν-ball'

$$B := B_1^\nu(0) = \{v \in V | \nu(v) < 1\}$$

is convex, balanced and absorbing.

A family $\mathcal{P} = (\nu_\alpha)_{\alpha \in A}$ of seminorms on V is said to be *separating* if for any point $v \in V \setminus \{0\}$ there is an index $\alpha \in A$ such that $\nu_\alpha(v) \neq 0$. A separating family of seminorms on a vector space is also called a *multinorm*. A *multinormed vector space* is a vector space endowed with a multinorm.

Lemma 1.8 *Suppose $\mathcal{P} = (\nu_\alpha)_{\alpha \in A}$ is a separating family of seminorms on a vector space V. For each $\alpha \in A$ and $n \in \mathbb{N}^*$, let*

$$\mathcal{V}(\alpha, n) := \left\{ v \in V \,\middle|\, \nu_\alpha(v) < \frac{1}{n} \right\}.$$

If $\mathcal{F}$ is the family of all finite intersections of the sets $\mathcal{V}(\alpha, n)$, then $\mathcal{F}$ is a fundamental system for a topology on V, which makes V into a locally convex TVS such that

(i) the members of $\mathcal{F}$ are convex balanced sets;

(ii) for each $\alpha \in A$, the function $\nu_\alpha : V \to \mathbb{R}$ is continuous;

(iii) a subset H of V is bounded if and only if every member of $\mathcal{P}$ is bounded on H.

For a proof we refer to Rudin's text [33].

Some comments to this important result seem to be appropriate.

Remark 1.9 Suppose V is a locally convex space, and let $\mathcal{T}$ be its linear topology. It may be shown that if $\mathcal{F}$ is a fundamental system of V consisting of convex balanced sets, then $\mathcal{F}$ generates a separating family $\mathcal{P}$ of seminorms on V. According to 1.8, $\mathcal{P}$ induces a topology $\mathcal{T}_1$ on V. Now it can easily be checked that $\mathcal{T}_1 = \mathcal{T}$.

Remark 1.10 Suppose that in Lemma 1.8 a *countable* family $\mathcal{P} = (\nu_n)_{n \in \mathbb{N}^*}$ of seminorms is given on V. Then the fundamental system arising from $\mathcal{P}$ is also countable, therefore, by Lemma 1.6, the induced topology is compatible with a translation invariant distance function. In addition, such a distance function can explicitly be constructed in terms of $\mathcal{P}$; for example, the formula

$$\varrho(u, v) := \sum_{n=1}^{\infty} \frac{2^{-n} \nu_n(u - v)}{1 + \nu_n(u - v)}; \quad (u, v) \in V \times V$$

defines a distance function which satisfies the desired properties.

As a consequence of our preceding discussion, we have the following useful characterization of Fréchet spaces.

Corollary 1.11 *A vector space is a Fréchet space, if and only if, it is a complete multinormed vector space $(V, (\nu_\alpha)_{\alpha \in A})$, where the set A is countable. Completeness is understood with respect to the* multinorm topology *whose subbasis is the family*

$$\left\{ B_r^{\nu_\alpha}(v) \subset V \,|\, \alpha \in A, r \in \mathbb{R}_+^*, v \in V \right\}.$$

The distance function described in 1.10 is compatible with the multinorm topology. □

2 Calculus in topological vector spaces and beyond

Local analysis in the context of topological vector spaces

In infinite-dimensional analysis there is a deep breaking between the case of (real or complex) Banach spaces and that of more general locally convex topological vector spaces which are not normable: depending on the type of derivatives used (Fréchet derivative, Gâteaux derivative, ...) one obtains non-equivalent calculi. As a consequence, there are several theories of infinite-dimensional manifolds, Lie groups and differential geometric structures. Changing the real or complex ground field to a more general topological field or ring, even more general differential calculus, Lie theory and differential geometry may be constructed [3, 4]. In this subsection we briefly explain the approach to differential calculus originated by A. D. Michal [25] and A. Bastiani [2], and popularized by J. Milnor [27] and R. Hamilton [13]. For an accurately elaborated, detailed recent account we refer to [8].

For simplicity, and in harmony with the applications we are going to present, *we restrict ourselves to the real case*. We begin with some remarks concerning the differentiability of curves with values in a locally convex topological vector space. This is the simplest situation, without any difficulty in principle. However, some interesting new phenomena occur.

Definition 2.1 Let V be a locally convex vector space and $I \subset \mathbb{R}$ an interval containing more than one point. By a C^0*-curve* on I with values in V we mean a continuous map $\gamma : I \to V$.

(1) Let $\alpha, \beta \in \mathbb{R}$, $\alpha < \beta$. A C^0-curve $\gamma : [\alpha, \beta] \to V$ is called a *Lipschitz curve* if the set

$$\left\{ \frac{1}{s-t}(\gamma(s) - \gamma(t)) \in V \middle| s, t \in [\alpha, \beta], s \neq t \right\}$$

is bounded in V.

(2) Suppose I is an open interval. A C^0-curve $\gamma : I \to V$ is said to be a C^1*-curve* if the limit

$$\gamma'(t) := \lim_{s \to 0} \frac{1}{s}(\gamma(t+s) - \gamma(t))$$

exists for all $t \in I$, and the map

$$\gamma' : t \in I \mapsto \gamma'(t) \in V$$

is continuous. Given $k \in \mathbb{N}^*$, γ is called *of class* C^k if all its iterated derivatives up to order k exist and are continuous. γ is *smooth* if it is of class C^k for all $k \in \mathbb{N}$.

Remark 2.2 It may be shown (but not at this stage of the theory) that if $\gamma : I \to V$ is a C^1-curve and $[\alpha, \beta] \subset I$ is a compact interval, then $\gamma \restriction [\alpha, \beta]$ is Lipschitz.

Definition 2.3 Suppose V is a locally convex vector space, and let $\gamma : [\alpha, \beta] \to V$ be a C^0-curve. If there exists a vector $v \in V$ such that for every continuous linear form $\lambda \in V^*$ we have

$$\lambda(v) = \int_\alpha^\beta \lambda \circ \gamma \quad \text{(Riemann integral)},$$

then $v \in V$ is called the *weak integral* of γ from α to β, and the notation

$$v =: \int_\alpha^\beta \gamma = \int_\alpha^\beta \gamma(t) dt$$

is applied. V is said to be *Mackey-complete* if the weak integral $\int_\alpha^\beta \gamma$ exists for each *Lipschitz-curve* $\gamma : [\alpha, \beta] \to V$.

Remark 2.4 (1) If the weak integral of a C^0-curve $\gamma : [\alpha, \beta] \to V$ exists, then it is unique, since by a version of the Hahn – Banach theorem V^* separates points on V (see e.g. [33, 3.4, Corollary]).

(2) If V is a sequentially complete locally convex vector space, then it is also Mackey-complete: it may be shown that the weak integral of any C^0-curve $\gamma : [\alpha, \beta] \to V$ can be obtained as the limit of a sequence of Riemann sums. For details see [8]. The importance of the concept of Mackey-completeness lies in the fact that a number of important constructions of the theory depends only on the existence of some weak integrals. For an alternative definition and an exhaustive description of Mackey-completeness we refer to the monograph of A. Kriegl and P. Michor [18].

(3) Let V and W be locally convex spaces, $\varphi : V \to W$ a continuous linear mapping, and $\gamma : I \to V$ a C^1-curve. One can check by an immediate application of the definition that $\varphi \circ \gamma : I \to W$ is also a C^1-curve, and

$$(\varphi \circ \gamma)' = \varphi \circ \gamma'.$$

Lemma 2.5 (Fundamental theorem of calculus) *Let V be a locally convex vector space, and $I \subset \mathbb{R}$ be an open interval.*

(1) If $\gamma : I \to V$ is a C^1-curve and $\alpha, \beta \in I$, then

$$\gamma(\beta) - \gamma(\alpha) = \int_\alpha^\beta \gamma'(t)dt.$$

(2) If $\gamma : I \to V$ is a C^0-curve, $\alpha \in I$, and the weak integral

$$\eta(t) := \int_\alpha^t \gamma(s)ds$$

exists for all $t \in I$, then $\eta : I \to V$ is a C^1-curve, and $\eta' = \gamma$.

Proof of part 1. Let $\lambda \in V^*$ be a continuous linear form. By Remark 2.4(3) and the classical 'Fundamental theorem of calculus', it follows that

$$\lambda(\gamma(\beta) - \gamma(\alpha)) = \lambda\gamma(\beta) - \lambda\gamma(\alpha) = \int_\alpha^\beta (\lambda \circ \gamma)' = \int_\alpha^\beta \lambda \circ \gamma'.$$

It means that $\gamma(\beta) - \gamma(\alpha)$ is the weak integral of γ' from α to β. □

Remark 2.6 The second part of the lemma is proved in [13]; it needs a more sophisticated argument, and hence additional preparations.

Definition 2.7 Let V and W be locally convex topological vector spaces, $\mathcal{U} \subset V$ an open set, and $f : \mathcal{U} \to W$ a mapping.

(1) By the *derivative* of f at a point $p \in \mathcal{U}$ *in the direction* $v \in V$ we mean the limit

$$D_v f(p) = df(p, v) := \lim_{t \to 0} \frac{1}{t}(f(p + tv) - f(p))$$

whenever it exists. f is called *differentiable at* p if $df(p, v)$ exists for all $v \in V$.

(2) The map f is said to be *continuously differentiable* or *of class* C^1 (briefly C^1) on $\mathcal{U}$ if it is differentiable at every point of $\mathcal{U}$ and the map

$$df : \mathcal{U} \times V \to W, \quad (p, v) \mapsto df(p, v)$$

is continuous.

(3) Let $k \in \mathbb{N}$, $k \geqq 2$. f is called a C^k*-map* (or briefly C^k) if the iterated directional derivatives

$$d^j f(p, v_1, \ldots, v_j) := (D_{v_j} \ldots D_{v_1} f)(p)$$

exist for all $j \in \mathbb{N}^*$ such that $j \leqq k$, $p \in \mathcal{U}$ and $v_1, \ldots, v_j \in V$, and the mappings $d^j f : \mathcal{U} \times V^j \to W$ are continuous. $d^j f$ is called the j*th differential* of f. If f is C^k for all $k \in \mathbb{N}$, then f is said to be C^∞ or *smooth*.

Notation We write $C^k(\mathcal{U}, W)$ for the set (in fact a vector space) of C^k-maps from $\mathcal{U}$ into W. When W is 1-dimensional, and hence $W \cong \mathbb{R}$, we usually just write $C^k(\mathcal{U})$.

Remark 2.8 (1) It is obvious from the definition that the derivative of a *linear mapping* exists at every point. Since there are linear mappings which are not continuous, it follows that *differentiability does not imply continuity*. However, if $\varphi : V \to W$ is a *continuous linear mapping*, then φ is smooth, and at every $(p, v) \in V \times V$ we have $d\varphi(p, v) = \varphi(v)$, while $d^k\varphi = 0$ for $k \geqq 2$.

(2) $d^j f$ in our notation is not the same as $d^j f$ in [8], rather it is the same as $d^{(j)} f$ in [8].

(3) The concept of C^k-differentiability introduced here will occasionally be mentioned as the *Michal – Bastiani differentiability*. Observe that in the formulation of the definition the local convexity of the underlying vector spaces does not play any role. However, if one wants to build a 'reasonable' theory of differentiation (with 'expectable' rules for calculation), the requirement of local convexity is indispensable.

Lemma 2.9 *Let V and W be locally convex vector spaces, $\mathcal{U} \subset V$ an open subset, and $f : \mathcal{U} \to W$ a C^1-map.*

(1) The map

$$f'(p) : V \to W, \quad v \mapsto f'(p)(v) := df(p, v)$$

is a continuous linear map for each $p \in \mathcal{U}$, and f is continuous.

(2) Let $p \in \mathcal{U}$, $v \in V$, and suppose that $p + tv \in \mathcal{U}$ for all $t \in [0, 1]$. Define the C^0-curve $c : [0, 1] \to W$ by

$$c(t) := df(p + tv, v) = f'(p + tv)(v).$$

Then

$$f(p+v) = f(p) + \int_0^1 c,$$

therefore f is locally constant, if and only if, $df = 0$.

(3) (Chain rule) *Suppose Z is another locally convex vector space, $\mathcal{V} \subset W$ is an open subset, and $h : \mathcal{V} \to Z$ is a C^1-map. If $f(\mathcal{U}) \subset \mathcal{V}$, then $h \circ f : \mathcal{U} \to Z$ is also a C^1-map, and for all $p \in \mathcal{U}$ we have*

$$(h \circ f)'(p) = h'(f(p)) \circ f'(p).$$

(4) (Schwarz's theorem) *If f is of class C^k $(k \geqq 2)$, then*

$$f^{(k)}(p) : (v_1, \dots, v_k) \in V^k \mapsto f^{(k)}(p)(v_1, \dots, v_k) := d^k f(p, v_1, \dots, v_k)$$

is a continuous, symmetric k-linear map for all $p \in \mathcal{U}$.

(5) (Taylor's formula) *Suppose f is of class C^k $(k \geqq 2)$. Then, if $p \in \mathcal{U}$, $v \in V$ and the segment joining p and $p+v$ is in $\mathcal{U}$, we have*

$$f(p+v) = f(p) + f'(p)(v) + \dots + \frac{1}{(k-1)!} f^{(k-1)}(p)(v, \dots, v) \\ + \frac{1}{(k-1)!} \int_0^1 c_k,$$

where $c_k : [0,1] \to W$ is a C^0-curve given by

$$c_k(t) := (1-t)^{k-1} f^{(k)}(p+tv)(v, \dots, v), \quad t \in [0,1].$$

Indication of proof. The continuity of $f'(p)$ is obvious, since $f'(p) = df(p, \cdot)$, and df is continuous. An immediate application of the definition of differentiability leads to the homogeneity of $f'(p)$. To check the additivity of $f'(p)$ some further (but not difficult) preparation is necessary, see [8, 1.2.13, 1.2.14]. To prove the integral representation in (2), let

$$\gamma(t) := f(p+tv), \quad t \in [0,1].$$

Then γ is differentiable at each $t \in [0,1]$ in the sense of 2.1(2), namely

$$\gamma'(t) := \lim_{s \to 0} \frac{\gamma(t+s) - \gamma(t)}{s} = \lim_{s \to 0} \frac{1}{s}(f(p+tv+sv) - f(p+tv)) \\ =: df(p+tv, v) = f'(p+tv)(v) =: c(t).$$

Thus $\gamma' = c$, and the fundamental theorem of calculus (2.5(1)) gives

$$f(p+v) - f(p) = \gamma(1) - \gamma(0) = \int_0^1 \gamma' = \int_0^1 c.$$

To see that f is continuous, choose a continuous seminorm $\nu : W \to \mathbb{R}$, and let ε be an arbitrary positive real number. Then there exists a balanced neighbourhood $\mathcal{U}_0$ of the origin in V such that $p + \mathcal{U}_0 \subset \mathcal{U}$, and for all $t \in [0, 1]$, $v \in \mathcal{U}_0$ we have

$$\nu(c(t)) = \nu(f'(p + tv)(v)) \leqq \varepsilon.$$

Now it may be shown [8, 1.1.8] that

$$\nu\left(\int_0^1 c\right) \leqq \sup\{\nu(c(t)) \in \mathbb{R} | t \in [0, 1]\},$$

therefore

$$\nu(f(p + v) - f(p)) = \nu\left(\int_0^1 c\right) \leqq \varepsilon,$$

and hence f is continuous. This concludes the sketchy proof of (1) and (2).

For a proof of the chain rule and Schwarz's theorem we refer to [13] and [8]. The latter reference also contains a detailed treatment of Taylor's formula. □

Remark 2.10 We keep the hypotheses and notations of the Lemma.

(1) The continuous linear map $f'(p) : V \to W$ introduced in 2.9(1) is said to be the *derivative* of f at p. Note that the symbol $f'(t)$ carries double meaning if $f : I \to W$ is a C^1-curve: by 2.1(2) $f'(t) \in W$, while by 2.9(1) $f'(t) \in \mathcal{L}(\mathbb{R}, W)$. Fortunately, this abuse of notation leads to no serious conflict since the vector spaces W and $\mathcal{L}(\mathbb{R}, W)$ can be canonically identified via the linear isomorphism

$$\gamma \in \mathcal{L}(\mathbb{R}, W) \mapsto \gamma(1) \in W.$$

(2) In Lemma 2.9 we have listed only the most elementary facts concerning Michal–Bastiani differentiation. It is a more subtle problem, for example, to obtain inverse (or implicit) function theorems in this (or a more general) context. The idea of generalization of the classical *inverse function theorem* for mappings between some types of Fréchet spaces is due to John Nash. Nash's inverse function theorem played an important role in his famous paper on isometric embeddings of Riemannian manifolds [30]. It was F. Sergeraert who stated the theorem explicitly in terms of a category of maps between Fréchet spaces [34]. In Moser's formulation [28] the theorem became an abstract theorem in functional analysis of wide applicability. Further generalizations have been given by Hamilton [14], Kuranishi [19], Zehnder [37], and more recently Leslie [22] and Ma [24]. In reference [10], inspired by Hiltunen's results [16], Glöckner proves *implicit function theorems* for mappings defined on topological vector spaces over valued fields. In particular, in the real and complex cases he obtains implicit function theorems for mappings from not necessarily locally convex topological vector spaces to Banach spaces.

To emphasize that the results of calculus in Banach spaces cannot be transplanted into the wider framework of topological vector spaces in general, finally we mention a quite typical pathology: the uniqueness and existence of solutions to ordinary differential equations are not guaranteed beyond Banach spaces. For a simple illustration of this phenomenon in Fréchet spaces see [13, 5.6.1].

(3) We briefly discuss the relation between the concept of Michal – Bastiani differentiability and the classical concept of Fréchet differentiability for mappings between Banach spaces. Recall that a continuous map from an open subset $\mathcal{U}$ of a Banach space V into a Banach space W is called *continuously Fréchet differentiable* or FC^1, if for all $p \in \mathcal{U}$ there exists a (necessarily unique) continuous linear map $f'(p) : V \to W$ such that

$$\lim_{v \to 0} \frac{f(p+v) - f(p) - f'(p)(v)}{\|v\|} = 0,$$

and the map $f' : \mathcal{U} \to \mathcal{L}(V,W)$, $p \mapsto f'(p)$ is continuous (with respect to the operator norm in $\mathcal{L}(V,W)$). Inductively, we define f to be FC^k ($k \geqq 2$) if f' is FC^{k-1}. Now it may be shown that *every FC^k-map is C^k (in the sense of Michal – Bastiani), and every C^{k+1}-map between open subsets of Banach spaces is FC^k*, so the two concepts coincide in the C^∞ case. For a proof we refer to [26] or [11].

(4) As an equally important approach to non-Banach infinite-dimensional calculus we have to mention the so-called *convenient calculus* elaborated in detail by A. Kriegl and P. W. Michor [18]. Let V and W be locally convex vector spaces and $\mathcal{U} \subset V$ an open subset. A mapping $f : \mathcal{U} \to W$ is said to be *conveniently smooth* if $f \circ \gamma : I \to W$ is a smooth curve for each smooth curve $\gamma : I \to \mathcal{U}$. By the chain rule 2.9(3) it is clear that if a mapping is Michal – Bastiani smooth, then it is conveniently smooth as well. The converse of this statement is definitely false: a conveniently smooth mapping need not even be continuous. However, *if V is a Fréchet space, then $f : \mathcal{U} \to W$ is conveniently smooth if and only if it is Michal – Bastiani smooth*. For a sketchy proof see [31, II, 2.10].

Remark 2.11 It should be noticed that all the difficulties arising in our preceding discussion disappear if the underlying vector spaces are finite dimensional. The first reason for this lies in the fact that Tychonoff's theorem (1.5) guarantees, roughly speaking, that the Euclidean topology of $\mathbb{R}^n$ is the only linear topology that an n-dimensional real vector space can have. This canonical topology is locally convex, locally compact and metrizable. In particular, if V and W are finite dimensional vector spaces, then $\mathcal{L}(V,W)$ also carries the canonical linear topology. This also leads to a significant difference between the calculus in finite or infinite dimensional Banach spaces on the one hand, and in non-Banachable spaces on the other hand. For example, if V and W are (non-Banach) Fréchet spaces, then the vector space of continuous linear maps between V and W is not necessarily a Fréchet space. In the following *we assume that all finite dimensional vector spaces are endowed with the canonical topology assured by Tychonoff's theorem.*

In the finite dimensional case the differentiability concepts introduced above result in the same class of C^k mappings. For lack of norms we shall always use the Michal – Bastiani definition. So if V and W are finite dimensional (real) vector spaces, $\mathcal{U} \subset V$ is an open set, then a mapping $f : \mathcal{U} \to W$ is C^1 if there is a continuous mapping $f' : \mathcal{U} \to \mathcal{L}(V,W)$ such that at each point $p \in \mathcal{U}$ and for all $v \in V$ we have

$$f'(p)(v) = \lim_{t \to 0} \frac{1}{t}(f(p+tv) - f(p)).$$

Higher derivatives and smoothness can be defined as in 2.7(3). Notice that if $f \in C^k(\mathcal{U}, W)$ ($k \geqq 2$), then its kth derivative at a point $p \in \mathcal{U}$ is denoted by $f^{(k)}(p)$, and it is an element of

$$\mathcal{L}\left(V, \mathcal{L}^{k-1}(V,W)\right) \cong \mathcal{L}^k(V,W)$$

given by

$$f^{(k)}(p)(v_1,\ldots,v_k) := \lim_{t\to 0}\frac{1}{t}\left(f^{(k-1)}(p+tv_1)(v_2,\ldots,v_k) - f^{(k-1)}(p)(v_2,\ldots,v_k)\right)$$

for each $(v_1,\ldots,v_k) \in V^k$.

Smooth functions and differential operators on $\mathbb{R}^n$

In this subsection we have a closer look at the most important special case, when the domain of the considered functions is a subset of the Euclidean n-space $\mathbb{R}^n$, and we describe the linear differential operators acting on the spaces of these functions.

First we recall a basic existence result.

Lemma 2.12 *Let $\mathcal{U}$ be a nonempty open subset of $\mathbb{R}^n$, and $(\mathcal{U}_i)_{i\in I}$ an open covering of $\mathcal{U}$. There exists a family $(f_i)_{i\in I}$ of smooth functions on $\mathcal{U}$ such that*

(i) $0 \leqq f_i(p) \leqq 1$ *for all $i \in I$ and $p \in \mathcal{U}$;*

(ii) $\operatorname{supp}(f_i) \subset \mathcal{U}_i$ *for all $i \in I$;*

(iii) $(\operatorname{supp}(f_i))_{i\in I}$ *is locally finite;*

(iv) for each point $p \in \mathcal{U}$ we have $\sum_{i\in I} f_i(p) = 1$. □

The family $(f_i)_{i\in I}$ in the Lemma is said to be a *partition of unity* subordinate to the covering $(\mathcal{U}_i)_{i\in I}$. The heart of the proof consists of an application of the purely topological Lemma 1.2 and the construction of a smooth function $f : \mathbb{R}^n \to \mathbb{R}$, called a smooth bump function, with the following properties:

$$0 \leqq f(p) \leqq 1 \text{ for all } p \in \mathbb{R}^n; \quad f(q) = 1 \text{ if } q \in \overline{B_1(0)}; \quad \operatorname{supp}(f) \subset \overline{B_2(0)}$$

(the balls are taken with respect to the Euclidean distance).

More generally, let $\mathcal{U}$ and $\mathcal{V}$ be open subsets, K a closed subset of $\mathbb{R}^n$, and suppose that $K \subset \mathcal{V} \subset \mathcal{U}$. A smooth function $f : \mathcal{U} \to \mathbb{R}$ is said to be a *bump function for K supported in $\mathcal{V}$* if $0 \leqq f(p) \leqq 1$ for each $p \in \mathcal{U}$, $f(q) = 1$ if $q \in K$, and $\operatorname{supp}(f) \subset \mathcal{V}$. As an immediate consequence of Lemma 2.12, we have

Corollary 2.13 *If $\mathcal{U}$ and $\mathcal{V}$ are open subsets, K is a closed subset of $\mathbb{R}^n$, and $K \subset \mathcal{V} \subset \mathcal{U}$, then there exists a bump function for K supported in $\mathcal{V}$.*

Indeed, if $\mathcal{U}_0 := \mathcal{V}, \mathcal{U}_1 := \mathcal{U} \setminus K$, then $(\mathcal{U}_0, \mathcal{U}_1)$ is an open covering of $\mathcal{U}$, so by 2.12 there exists a partition of unity (f_0, f_1) subordinate to $(\mathcal{U}_0, \mathcal{U}_1)$. Then $f_1 \restriction K = 0$, and for each $q \in K$ we have $f_0(q) = (f_0 + f_1)(q) = 1$, therefore the function $f := f_0$ has the desired properties.

Remark 2.14 In the following the canonical basis of $\mathbb{R}^n$ will be denoted by $(e_i)_{i=1}^n$, and $(e^i)_{i=1}^n$ will stand for its dual. The family $(e^i)_{i=1}^n$ will also be mentioned as the canonical coordinate system for $\mathbb{R}^n$. If $\mathcal{U} \subset \mathbb{R}^n$ is an open subset, and $f \in C^\infty(\mathcal{U})$, then

$$D_i f : \mathcal{U} \to \mathbb{R}, \quad p \mapsto D_i f(p) := f'(p)(e_i) \quad (i \in \{1, \dots, n\})$$

is the *ith partial derivative* of f (with respect to the canonical coordinate system). For each multi-index $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ we write

$$D^\alpha f := D_1^{\alpha_1} \dots D_n^{\alpha_n} f; \quad D_i^{\alpha_i} f := \underbrace{D_i \dots D_i}_{\alpha_i \text{times}} f \quad (i \in \{1, \dots, n\})$$

with the convention $D^\alpha f := f$ if $|\alpha| = 0$. We say that D^α is an *elementary partial differential operator of order* $|\alpha|$. A *linear differential operator* is a linear combination $D = \sum_{\|\alpha\| \leqq m} a_\alpha D^\alpha$, where $m \in \mathbb{N}$, $a_\alpha \in C^\infty(\mathcal{U})$. Clearly, D maps $C^\infty(\mathcal{U})$ linearly into $C^\infty(\mathcal{U})$. It is not difficult to show (see e.g. [6, (8.13.1)], or [20, Ch. XI, §1]) that *if a linear differential operator is identically* 0 *on* $C^\infty(\mathcal{U})$*, then each of its coefficients is identically* 0 *on* $\mathcal{U}$. From this it follows that the coefficients of a linear diferential operator $D = \sum_{\|\alpha\| \leqq m} a_\alpha D^\alpha$ are uniquely determined; the highest value of $|\alpha|$ such that $a_\alpha \neq 0$ is called the *order* of D.

Lemma 2.15 (generalized Leibniz rule) *Let* $\mathcal{U} \neq \emptyset$ *be an open subset of* $\mathbb{R}^n$*. If* $f, h \in C^\infty(\mathcal{U})$ *and* $\alpha \in \mathbb{N}^n$ *is a multi-index, then*

$$D^\alpha(fh) = \sum_{\mu + \nu = \alpha} \binom{\alpha}{\nu} (D^\nu f)(D^\mu h).$$

□

Remark 2.16 We need some further notations. We shall denote by $C_c^\infty(\mathbb{R}^n)$ the subspace of $C^\infty(\mathbb{R}^n)$ consisting of smooth functions on $\mathbb{R}^n$ which have compact support. For any nonempty subset S of $\mathbb{R}^n$, $C_c^\infty(S)$ will stand for the space of functions in $C_c^\infty(\mathbb{R}^n)$ whose support lies in S. A function in $C_c^\infty(S)$ will be identified with its restriction to S. (Notice that L. Schwartz's notation $\mathcal{E}(\mathcal{U}) := C^\infty(\mathcal{U})$, $\mathcal{D}(\mathcal{U}) := C_c^\infty(\mathcal{U})$, $\mathcal{D}(K) := C_c^\infty(K)$ is widely used in distribution theory.)

Proposition 2.17 *Let a (nonempty) open subset* $\mathcal{U}$ *of* $\mathbb{R}^n$ *be given. For any compact set* K *contained in* $\mathcal{U}$ *and any multi-index* $\alpha \in \mathbb{N}^n$*, define the function* $\| \ \|_\alpha^K : C^\infty(\mathcal{U}) \to \mathbb{R}$ *by setting*

$$\|f\|_\alpha^K := \sup_{p \in K} |D^\alpha f(p)|, \quad f \in C^\infty(\mathcal{U}).$$

Then the family $\left(\| \ \|_\alpha^K\right)$ *is a multinorm on* $C^\infty(\mathcal{U})$ *which makes it into a Fréchet space having the Heine – Borel property, such that* $C_c^\infty(K)$ *is a closed subspace of* $C^\infty(\mathcal{U})$ *whenever* $K \subset \mathcal{U}$ *is compact.*

Proof. It is clear that the family $\left(\| \ \|_\alpha^K\right)$ is a multinorm, so it defines a locally convex topology on $C^\infty(\mathcal{U})$ according to 1.8. We claim that the TVS so obtained is metrizable and complete. In order to show this we construct another *countable* family of seminorms on $C^\infty(\mathcal{U})$ which is equivalent to the given one in the sense that the two families generate the same topology.

Let $(K_n)_{n\in\mathbb{N}^*}$ be the compact exhaustion of $\mathcal{U}$ described in the proof of 1.1. If for each $n \in \mathbb{N}^*$

$$\nu_n(f) := \sup_{p\in K_n, |\alpha|\leqq n} |D^\alpha f(p)|, \quad f \in C^\infty(\mathcal{U}),$$

then $(\nu_n)_{n\in\mathbb{N}^*}$ is a multinorm on $C^\infty(\mathcal{U})$, which defines a metrizable locally convex topology on $C^\infty(\mathcal{U})$ by 1.8 and 1.10. To prove the equivalence of the two seminorm-families, it is enough to check that each member of the first family is majorized by a finite linear combination of the second family, and conversely.

Now, on the one hand, it is obvious that for every positive integer n and smooth function f in $C^\infty(\mathcal{U})$ we have

$$\nu_n(f) \leqq \sum_{|\alpha|\leqq n} \|f\|_\alpha^{K_n} \left(\overset{2.18}{=} \|f\|_n^{K_n}\right),$$

thus ν_n is majorized by $\left(\| \ \|_\alpha^{K_n}\right)_{|\alpha|\leqq n}$. To prove the converse, choose a compact set $K \subset \mathcal{U}$, and define the following functions:

$$\delta : K \to \left]0, \infty\right], \quad p \mapsto \delta(p) := \varrho_E(p, \partial\mathcal{U});$$
$$\Delta : K \to \mathbb{R}_+, \quad p \mapsto \Delta(p) := \varrho_E(p, 0) = \|p\|.$$

Then both δ and Δ are continuous; δ attains its minimum $c \in \mathbb{R}_+^*$, and Δ attains its maximum $C \in \mathbb{R}_+$ on K. Now choose $n \in \mathbb{N}^*$ such that

$$\frac{1}{n} < c \quad \text{and} \quad n > C.$$

Then the member K_n of the compact exhaustion of $\mathcal{U}$ contains the compact set K. If, in addition, $n \geqq |\alpha|$, the seminorm ν_n is a majorant of the seminorm $\| \ \|_\alpha^K$.

For a proof of the remaining claims we refer to [33, 1.46]. □

Remark 2.18 Let $\mathcal{U}$ be an open subset of $\mathbb{R}^n$, S any subset of $\mathcal{U}$, and $m \in \mathbb{N}$. It is easy to check that the mapping

$$\| \ \|_m^S : f \in C_c^\infty(\mathcal{U}) \mapsto \|f\|_m^S := \sum_{|\alpha|\leqq m} \sup_{p\in S} |D^\alpha f(p)|$$

is also a seminorm on $C_c^\infty(\mathcal{U})$. Using this seminorm, in case of $S = \mathcal{U}$ we shall omit the superscript in the notation.

Now, following Narasimhan [29, 1.5.1], we introduce a purely technical term. We say that a function $f \in C^\infty(\mathcal{U})$ is *m-flat* at a point $p \in \mathcal{U}$ if $D^\alpha f(p) = 0$, whenever $|\alpha| \leqq m$. We shall find the following observation useful:

Lemma 2.19 *If a function $f \in C^\infty(\mathbb{R}^n)$ is m-flat at the origin, then for every $\varepsilon \in \mathbb{R}_+^*$ there exists a function $h \in C^\infty(\mathbb{R}^n)$, vanishing on a 0-neighbourhood, such that $\|h - f\|_m < \varepsilon$.*

Proof. Corollary 2.13 assures the existence of a function $\psi \in C^\infty(\mathbb{R}^n)$ such that

$$\forall p \in \mathbb{R}^n : 0 \leqq \psi(p) \leqq 1, \quad \psi(p) = 0 \text{ if } p \in \overline{B_{\frac{1}{2}}(0)}, \quad \psi(p) = 1 \text{ if } p \in \mathbb{R}^n \smallsetminus B_1(0).$$

Given a positive real number δ, consider the homothety h_δ of $\mathbb{R}^n$, and let

$$h := \left(\psi \circ h_{1/\delta}\right) f.$$

Then, obviously, $h \in C^\infty(\mathbb{R}^n)$, and h vanishes on a 0-neighbourhood. So it is sufficient to show that if $|\alpha| \leqq m$,

$$\sup_{p \in \mathbb{R}^n} |D^\alpha h(p) - D^\alpha f(p)| \to 0 \quad \text{as } \delta \to 0. \tag{$*$}$$

Since $h(p) = f(p)$ if $\|p\| > \delta$, it follows that

$$\begin{aligned} \sup_{p \in \mathbb{R}^n} |D^\alpha h(p) - D^\alpha f(p)| &= \sup_{\|p\| \leqq \delta} |D^\alpha h(p) - D^\alpha f(p)| \\ &\leqq \sup_{\|p\| \leqq \delta} |D^\alpha h(p)| + \sup_{\|p\| \leqq \delta} |D^\alpha f(p)|. \end{aligned}$$

By our assumption, we have $D^\alpha f(0) = 0$ for $|\alpha| \leqq m$; thus

$$\sup_{\|p\| \leqq \delta} |D^\alpha f(p)| \to 0 \quad \text{as } \delta \to 0. \tag{$**$}$$

Now we consider the function $D^\alpha h$. Using 2.15, we obtain

$$\begin{aligned} D^\alpha h = D^\alpha \left(\left(\psi \circ h_{1/\delta}\right) f\right) &= \sum_{\mu+\nu=\alpha} \binom{\alpha}{\nu} D^\nu \left(\psi \circ h_{1/\delta}\right) D^\mu f \\ &= \sum_{\mu+\nu=\alpha} \binom{\alpha}{\nu} \delta^{-|\nu|} \left(\left(D^\nu \psi\right) \circ h_{1/\delta}\right) D^\mu f. \end{aligned}$$

Since ψ is constant outside $B_1(0)$, the function $D^\nu \psi$ is bounded. If

$$C_\nu := \sup_{p \in \mathbb{R}^n} |D^\nu \psi(p)|, \quad C := \max_\nu \binom{\alpha}{\nu} C_\nu,$$

then we get the following estimation:

$$|D^\alpha h| \leqq C \sum_{\mu+\nu=\alpha} \delta^{-|\nu|} |D^\mu f|.$$

As f is m-flat at 0, $D^\mu f$ is $(m - |\mu|)$-flat at the origin. Thus, using Landau's symbol $o(\cdot)$,

$$\sup_{\|p\| \leqq \delta} |D^\mu f(p)| = o\left(\delta^{m-|\mu|}\right),$$

therefore

$$\sup_{\|p\| \leqq \delta} |D^\alpha h(p)| = o\left(\textstyle\sum_{\mu+\nu=\alpha} \delta^{m-|\mu|-|\nu|}\right) = o\left(\delta^{m-|\alpha|}\right),$$

and hence

$$\sup_{p\in\mathbb{R}^n} |D^\alpha h(p) - D^\alpha f(p)| \leqq o\left(\delta^{m-|\alpha|}\right) + \sup_{\|p\|\leqq\delta} |D^\alpha f(p)|.$$

In view of $(**)$, this implies the desired relation $(*)$. □

Lemma 2.20 *Let $\mathcal{U} \subset \mathbb{R}^n$ be a (nonempty) open set. Suppose $D : C^\infty(\mathcal{U}) \to C^\infty(\mathcal{U})$, $f \mapsto Df$ is a linear mapping which decreases supports, that is*

$$\operatorname{supp}(Df) \subset \operatorname{supp}(f)$$

for all functions $f \in C^\infty(\mathcal{U})$. Then for each point $p \in \mathcal{U}$ there exists a relatively compact neighbourhood $\mathcal{V}$ of p $\left(\bar{\mathcal{V}} \subset \mathcal{U}\right)$, a positive integer m and a positive real number C such that

$$\|Du\|_0 \leqq C\|u\|_m$$

holds for any function $u \in C_c^\infty(\mathcal{V} \smallsetminus \{p\})$ ($\| \ \|_k$ $(k \in \mathbb{N})$ is the seminorm introduced in 2.18).

Note Recall that any function $f \in C_c^\infty(\mathcal{V} \smallsetminus \{p\})$ is identified with a function $\tilde{f} \in C^\infty(\mathcal{U})$ according to 2.16. Since D is support-decreasing, we may apply it to f by the formula $Df := D\tilde{f}$.

Proof of the lemma. Suppose the contrary: there is a point $p \in \mathcal{U}$ such that for any relatively compact nbd $\mathcal{V} \subset \mathcal{U}$ of p, any positive integer m and positive real number C, there is a function $u \in C_c^\infty(\mathcal{V} \smallsetminus \{p\})$ such that

$$\|Du\|_0 > C\|u\|_m.$$

1st step Choose a relatively compact set $\mathcal{U}_0 \subset \mathcal{U}$ in $\mathcal{N}(p)$, and let $m := 1$, $C := 2^2$. By our assumption, there is a function $u_1 \in C_c^\infty(\mathcal{U}_0 \smallsetminus \{p\})$ such that

$$\|Du_1\|_0 > 2^2\|u_1\|_1.$$

Let $\mathcal{U}_1 := \{q \in \mathcal{U}_0 | u_1(q) \neq 0\}$. Then $\mathcal{U}_0 \smallsetminus \overline{\mathcal{U}_1} \in \mathcal{N}(p)$, so, using our assumption again, there is a function $u_2 \in C_c^\infty\left(\mathcal{U}_0 \smallsetminus \overline{\mathcal{U}_1} \smallsetminus \{p\}\right)$ such that

$$\|Du_2\|_0 > 2^4\|u_2\|_2.$$

Now we define $\mathcal{U}_2 := \left\{q \in \mathcal{U}_0 \smallsetminus \overline{\mathcal{U}_1} \middle| u_2(q) \neq 0\right\}$, and we repeat our argument. Thus, by induction, we obtain a sequence $(\mathcal{U}_k)_{k\in\mathbb{N}^*}$ of open sets and $(u_k)_{k\in\mathbb{N}^*}$ of functions such that

$$\overline{\mathcal{U}_k} \subset \mathcal{U}_0 \smallsetminus \{p\} \quad (k \in \mathbb{N}^*); \quad \overline{\mathcal{U}_k} \cap \overline{\mathcal{U}_\ell} = \emptyset \text{ if } k \neq \ell;$$
$$u_k \in C_c^\infty\left(\mathcal{U}_0 \smallsetminus \overline{\mathcal{U}_1} \smallsetminus \ldots \smallsetminus \overline{\mathcal{U}_{k-1}} \smallsetminus \{p\}\right) \subset C_c^\infty(\mathcal{U});$$
$$\mathcal{U}_k = \left\{q \in \mathcal{U}_0 \smallsetminus \overline{\mathcal{U}_1} \smallsetminus \ldots \smallsetminus \overline{\mathcal{U}_{k-1}} \middle| u_k(q) \neq 0\right\} \quad (k \in \mathbb{N}^*),$$

and, finally,

$$\|Du_k\|_0 > 2^{2k}\|u_k\|_k. \tag{$*$}$$

2nd step Consider the function

$$u := \sum_{k\in\mathbb{N}^*} \frac{2^{-k}}{\|u_k\|_k} u_k.$$

Since at each point of $\mathcal{U}_0$, at most one member of the right-hand side differs from zero, u is well-defined. We claim that u *is smooth on its domain*. To prove this, we have to show that each point $p \in \mathcal{U}_0$ has a neighbourhood on which u is smooth. We divide the points of $\mathcal{U}_0$ into three disjoint classes in the following way. First, let

$$\mathcal{V}_1 := \overset{\infty}{\underset{k=1}{\cup}} \mathcal{U}_k.$$

Then any point $p \in \mathcal{V}_1$ is contained in some $\mathcal{U}_k$, on which u is obviously smooth. Next, let $\mathcal{V}_2$ be the set of points in $\mathcal{U}_0$ which have a neighbourhood $\mathcal{V}$ intersecting with at most one of the $\mathcal{U}_k$'s. Then u is smooth on $\mathcal{V}$ as well. Finally, let $\mathcal{V}_3$ be the set of points in $\mathcal{U}_0$ whose every neighbourhood intersects with infinitely many of the $\mathcal{U}_k$'s. The only difficulty is to verify the smoothness of u in a neighbourhood of such points. By the Bolzano – Weierstraß theorem, $\mathcal{V}_3$ cannot be empty.

Let $p \in \mathcal{V}_3$. First we show that u is continuous at p. Since $p \notin \mathcal{U}_k$ $(k \in \mathbb{N}^*)$, $u(p) = 0$. Let $\varepsilon > 0$ be arbitrary, and

$$k := \left\lceil \log_2 \frac{1}{\varepsilon} \right\rceil$$

(where the symbol $\lceil\ \rceil$ denotes upper integer part). Now we define

$$\delta := \min_{1 \leqq i \leqq k} d\left(p, \overline{\mathcal{U}_i}\right).$$

Then $\delta > 0$, otherwise $p \in \mathcal{V}_2$ would follow. If $q \in B_\delta(p)$, we have either $u(q) = 0$ or $q \in \mathcal{U}_\ell$ for some $\ell > k$, thus

$$|u(q)| = \left| \frac{2^{-\ell}}{\|u_\ell\|_\ell} u_\ell(q) \right| \leqq \left| \frac{2^{-\ell}}{\|u_\ell\|_0} u_\ell(q) \right| \leqq 2^{-\ell} < 2^{-k} \leqq 2^{-\log_2(1/\varepsilon)} = \varepsilon,$$

therefore u is continuous at p.

Now we proceed by induction. Let m be a fixed positive integer, and suppose that all partial derivatives of u up to order $m-1$ exist and are continuous, and they all vanish at p. Let $\alpha \in \mathbb{N}^n$ be a multi-index such that $|\alpha| = m-1$, and $i \in \{1, \ldots, n\}$. To show that $D_i D^\alpha u(p)$ exists, we have to consider the following limit:

$$\lim_{t\to 0} \frac{D^\alpha u(p + te_i) - D^\alpha u(p)}{t} = \lim_{t\to 0} \frac{D^\alpha u(p + te_i)}{t}.$$

Let $\varepsilon > 0$ be arbitrary. Now we define k *almost* in the same way as in the previous part of the proof:

$$k := \max \left\{ m, \left\lceil \log_2 \frac{1}{\varepsilon} \right\rceil \right\}.$$

Let $t \in \,]0, \delta[$, where δ is defined *exactly* in the same way as above. Then either $D^\alpha u(p + te_i) = 0$, or $p + te_i \in \mathcal{U}_\ell$ for some $\ell > k$. Let

$$t_0 := \sup\{s \in \,]0, t[\,|p + se_i \notin \mathcal{U}_\ell\}.$$

Then the function $s \mapsto D^\alpha u(p + se_i)$ is continuous on $[t_0, t]$ and differentiable on $]t_0, t[$, so, by Lagrange's mean value theorem, there is some $\xi \in \,]t_0, t[$ such that

$$\begin{aligned}(s \mapsto D^\alpha u(p + se_i))'(\xi) &= D_i D^\alpha u(p + \xi e_i)\\ &= \frac{D^\alpha u(p + te_i) - D^\alpha u(p + t_0 e_i)}{t - t_0} = \frac{D^\alpha u(p + te_i)}{t - t_0}.\end{aligned}$$

Thus we have the following estimation:

$$\begin{aligned}&\left|\frac{D^\alpha u(p + te_i)}{t}\right| \leqq \left|\frac{D^\alpha u(p + te_i)}{t - t_0}\right| = |D_i D^\alpha u(p + \xi e_i)|\\ &= \left|\frac{2^{-\ell}}{\|u_\ell\|_\ell} D_i D^\alpha u_\ell(p + \xi e_i)\right| \leqq \left|\frac{2^{-\ell}}{\|D_i D^\alpha u_\ell\|_0} D_i D^\alpha u_\ell(q)\right|\\ &\leqq 2^{-\ell} < 2^{-k} \leqq 2^{-\log_2(1/\varepsilon)} = \varepsilon.\end{aligned}$$

If $t \in \,]-\delta, 0[$, then we proceed in the same way to obtain

$$D_i D^\alpha u(p) = \lim_{t \to 0} \frac{D^\alpha u(p + te_i) - D^\alpha u(p)}{t} = 0.$$

Therefore $D_i D^\alpha u$ exists at every point of $\mathcal{U}_0$. Its continuity is shown in the same way as that of u.

3rd step We obviously have

$$u \restriction \mathcal{U}_k = 2^{-k}(\|u_k\|_k)^{-1} u_k. \tag{$**$}$$

Since D is linear and support-decreasing, it follows that

$$(Du) \restriction \mathcal{U}_k = 2^{-k}(\|u_k\|_k)^{-1}(Du_k) \restriction u_k.$$

Thus, taking into account relations $(*)$ and $(**)$, we conclude that there is a point $p_k \in \mathcal{U}_k$ such that

$$|Du(p_k)| = 2^{-k}(\|u_k\|_k)^{-1}|Du_k(p_k)| > 2^{-k}(\|u_k\|_k)^{-1} \cdot 2^{2k}\|u_k\|_k = 2^k.$$

On the other hand, the function Du is continuous, and its support is compact, hence it is bounded. This contradicts the above assertion. □

Lemma 2.21 *Suppose $\mathcal{U} \subset \mathbb{R}^n$ is a nonempty open set, and let $\mathcal{V}$ be a relatively compact open set contained in $\mathcal{U}$. Consider a support-decreasing linear mapping $D : C^\infty(\mathcal{U}) \to C^\infty(\mathcal{U})$. Assume there is a positive integer m and a positive real number C such that*

$$\|Du\|_0 \leqq C\|u\|_m \tag{$*$}$$

for all $u \in C_c^\infty(\mathcal{V})$. Then *(i)* $Du(p) = 0$ *whenever* u *is* m*-flat at* p,

(ii) there exist smooth functions a_α ($\alpha \in \mathbb{N}^n$, $|\alpha| \leqq m$) *in* $C^\infty(\mathcal{V})$ *such that for each* $u \in C_c^\infty(\mathcal{V})$, $p \in \mathcal{V}$ *we have*

$$Du(p) = \sum_{|\alpha| \leqq m} a_\alpha(p)(D^\alpha u)(p).$$

Proof. (i) By Lemma 2.19, there is a sequence $(u_n)_{n\in\mathbb{N}}$ of functions in $C_c^\infty(\mathcal{V})$ such that u_n vanishes in a neighbourhood of p for each $n \in \mathbb{N}$, and

$$\lim_{n\to\infty} \|u_n - u\|_m = 0.$$

Taking into account our condition $(*)$, this implies that $(Du_n)_{n\in\mathbb{N}}$ converges uniformly to Du on $\mathcal{V}$. Since $\operatorname{supp}(Du_n) \subset \operatorname{supp}(u_n)$, and u_n vanishes near p, we have $Du_n(p) = 0$ for each $n \in \mathbb{N}$. Hence

$$Du(p) = \lim_{n\to\infty} (Du_n)(p) = 0,$$

as we claimed.

(ii) To prove the second statement, consider for an arbitrarily chosen point $p \in \mathcal{U}$ and multi-index $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ the polynomial

$$\mu_{\alpha,p}(\xi) := \left(\xi^1 - p^1\right)^{\alpha_1} \cdot \ldots \cdot (\xi^n - p^n)^{\alpha_n}$$

($\xi = \left(\xi^1, \dots, \xi^n\right)$ is a symbol). Then $\mu_{\alpha,p}$ can be viewed as a smooth function in $C^\infty(\mathcal{U})$, and the functions

$$\eta_\alpha : p \in \mathcal{U} \mapsto \eta_\alpha(p) := (D\mu_{\alpha,p})(p) \quad (\alpha \in \mathbb{N}^n)$$

also belong to $C^\infty(\mathcal{U})$. Now let $u \in C_c^\infty(\mathcal{V})$, and define

$$f := u - \sum_{|\alpha| \leqq m} \frac{1}{\alpha!}(D^\alpha u)(p)\mu_{\alpha,p}.$$

Then for each $\beta \in \mathbb{N}^n$, $|\beta| \leqq m$ we have

$$D^\beta f = D^\beta u - \sum_{|\alpha| \leqq m, \alpha \geqq \beta} \frac{1}{\alpha!}(D^\alpha u)(p)\frac{\alpha!}{(\alpha-\beta)!}\mu_{\alpha-\beta,p}.$$

Since

$$\mu_{\alpha-\beta,p}(p) = \begin{cases} 0 & \text{if } \beta \neq \alpha, \\ 1 & \text{if } \beta = \alpha, \end{cases}$$

it follows that

$$D^\beta f(p) = 0, \quad |\beta| \leqq m;$$

i.e., the function f is m-flat at p. Thus, by part (i), $Df(p) = 0$, therefore

$$Du(p) = \sum_{|\alpha| \leqq m} \frac{1}{\alpha!} (D^\alpha u)(p) \eta_\alpha(p);$$

so with the help of the smooth functions $a_\alpha := \frac{1}{\alpha!}\eta_\alpha$ ($\alpha \in \mathbb{N}^n$, $|\alpha| \leqq m$) Du can be represented in the desired form.

□

Now we introduce a more sophisticated version of the concept of a linear differential operator mentioned in 2.14.

Definition 2.22 Let $\mathcal{U} \subset \mathbb{R}^n$ be a nonempty open subset. A *differential operator* on $\mathcal{U}$ is a linear mapping $D : C^\infty(\mathcal{U}) \to C^\infty(\mathcal{U})$ with the following property: for each relatively compact open set $\mathcal{V}$ whose closure in contained in $\mathcal{U}$ there exists a finite family of functions $a_\alpha \in C^\infty(\mathcal{V})$ ($\alpha \in \mathbb{N}^n$) such that for each $u \in C^\infty(\mathcal{V})$,

$$Du = \sum_\alpha a_\alpha (D^\alpha u).$$

Note Differential operators in this more general sense also have a well-defined *order* locally, according to 2.14.

Having this concept, we are ready to formulate and prove the main result of this subsection.

Theorem 2.23 (local Peetre theorem) *Let $\mathcal{U} \subset \mathbb{R}^n$ be a nonempty open subset. If $D : C^\infty(\mathcal{U}) \to C^\infty(\mathcal{U})$ is a support-decreasing linear mapping, then D is a differential operator on $\mathcal{U}$. Conversely, any differential operator is support decreasing.*

Proof. The converse statement is clearly true. To prove the direct statement, consider a linear mapping $D : C^\infty(\mathcal{U}) \to C^\infty(\mathcal{U})$ with the property

$$\operatorname{supp}(Du) \subset \operatorname{supp}(u), \quad u \in C^\infty(\mathcal{U}).$$

Let $\mathcal{V} \subset \mathcal{U}$ be a relatively compact open set such that $\bar{\mathcal{V}} \subset \mathcal{U}$. According to 2.20, for each point $p \in \mathcal{V}$ there is a neighbourhood $\mathcal{U}_p \subset \mathcal{U}$ of p, a positive integer m_p and a positive real number C_p such that

$$\|Du\|_0 \leqq C_p \|u\|_{m_p}$$

holds for any function $u \in C_c^\infty(\mathcal{U}_p \setminus \{p\})$. The family $(\mathcal{U}_p)_{p \in \mathcal{V}}$ is an open covering of $\bar{\mathcal{V}}$, so by its compactness, there are finitely many points $p_1, \dots, p_k$ in $\mathcal{V}$ such that the corresponding open sets $\mathcal{U}_1, \dots, \mathcal{U}_k$ still cover $\bar{\mathcal{V}}$. Let

$$m := \max\{m_{p_i} | i \in \{1, \dots, k\}\} \quad \text{and} \quad C := \max\{C_{p_i} | i \in \{1, \dots, k\}\}.$$

Then for each $i \in \{1, \dots, k\}$ and $u \in C_c^\infty(\mathcal{U}_i \setminus \{p_i\})$ we have

$$\|Du\|_0 \leqq C \|u\|_m. \qquad (*)$$

We show that there is also a positive constant $\tilde{C}$ such that

$$\|Du\|_0 \leqq \tilde{C} \|u\|_m$$

holds for any function $u \in C_c^\infty(\mathcal{V} \smallsetminus \{p_1, \dots, p_k\})$. Using Lemma 2.12, let $(f_i)_{i=1}^{k+1}$ be a partition of unity subordinate to the open covering $\big(\mathcal{U}_1, \dots, \mathcal{U}_k, \mathcal{U} \smallsetminus \bar{\mathcal{V}}\big)$ of $\mathcal{U}$. Then each function $u \in C_c^\infty(\mathcal{V} \smallsetminus \{p_1, \dots, p_n\})$ can be written in the form

$$u = \sum_{i=1}^{k+1} f_i u = \sum_{i=1}^{k} f_i u,$$

and $(*)$ holds for every member of the sum, therefore

$$\begin{aligned}\|Du\|_0 = \left\| D\left(\textstyle\sum_{i=1}^k f_i u\right)\right\|_0 = \left\|\textstyle\sum_{i=1}^k D(f_i u)\right\|_0 \\ \leqq \sum_{i=1}^k \|D(f_i u)\|_0 \leqq \sum_{i=1}^k C\|f_i u\|_m = C\sum_{i=1}^k \|f_i u\|_m.\end{aligned}$$

Thus it is enough to show that there are positive constants C_i such that

$$\|f_i u\|_m \leqq C_i \|u\|_m \quad (i = 1, \dots, k)$$

which are independent of u (and may depend on the partition of unity chosen, however). By the generalized Leibniz rule (2.15), we have

$$\begin{aligned}\|f_i u\|_m &= \sum_{|\alpha\| \leqq m} \sup_{p \in \mathcal{V}} |D^\alpha (f_i u)(p)| \\ &= \sum_{|\alpha\| \leqq m} \sup_{p \in \mathcal{V}} \left| \textstyle\sum_{\mu+\nu=\alpha} \binom{\alpha}{\nu} (D^\nu f_i)(p)(D^\mu u)(p)\right| \\ &\leqq \sum_{|\alpha\| \leqq m} \sup_{p \in \mathcal{V}} \sum_{\mu+\nu=\alpha} \binom{\alpha}{\nu} |D^\nu f_i(p)||D^\mu u(p)| \\ &\leqq \sum_{|\alpha\| \leqq m} \sum_{\mu+\nu=\alpha} \binom{\alpha}{\nu} \sup_{p \in \mathcal{V}} |D^\nu f_i(p)||D^\mu u(p)| \\ &\leqq \sum_{|\alpha\| \leqq m} \sum_{\mu+\nu=\alpha} \binom{\alpha}{\nu} \big(\sup\nolimits_{p \in \mathcal{V}} |D^\nu f_i(p)|\big)\big(\sup\nolimits_{p \in \mathcal{V}} |D^\mu u(p)|\big) \\ &= \sum_{|\mu| \leqq m} \left[\textstyle\sum_{|\mu+\nu| \leqq m} \binom{\mu+\nu}{\nu} \sup_{p \in \mathcal{V}} |D^\nu f_i(p)|\right] \sup_{p \in \mathcal{V}} |D^\mu u(p)|.\end{aligned}$$

Letting

$$C_i := \sum_{|\mu+\nu| \leqq m} \binom{\mu+\nu}{\nu} \sup_{p \in \mathcal{V}} |D^\nu f_i(p)|,$$

we obtain

$$\|f_i u\|_m \leqq C_i \sum_{|\mu| \leqq m} \sup_{p \in \mathcal{V}} |D^\mu u(p)| = C_i \|u\|_m,$$

thus the desired estimation is valid for each term. Finally, if

$$\tilde{C} := C \sum_{i=1}^{k} C_i,$$

then we conclude

$$\|Du\|_0 \leqq C \sum_{i=1}^{k} \|f_i u\|_m \leqq C \sum_{i=1}^{k} C_i \|u\|_m = \tilde{C}\|u\|_m.$$

Lemma 2.21 then implies that for each $p \in \mathcal{V} \setminus \{p_1, \dots, p_n\}$,

$$Du(p) = \sum_{|\alpha| \leqq m} a_\alpha(p)(D^\alpha u)(p), \quad a_\alpha \in C^\infty(\mathcal{V}).$$

However, both functions Du and $\sum_{|\alpha|\leqq m} a_\alpha(D^\alpha u)$ are continuous, so the desired relation is valid at every point of $\mathcal{V}$. □

Transition from local to global

1st step: manifolds and bundles Since we have a chain rule for C^1-mappings between locally convex TVSs, the concept of a manifold modeled on such a vector space can be introduced without any difficulty: we may follow the well paved path of the finite-dimensional theory.

Let M be a Hausdorff space and V a locally convex TVS. By a *V-chart* on M we mean a pair $(\mathcal{U}, x)$, where $\mathcal{U} \subset M$ is an open subset of M, and x is a homeomorphism of $\mathcal{U}$ onto an open subset of V. Two charts, $(\mathcal{U}, x)$ and $(\mathcal{V}, y)$ are said to be *smoothly compatible* if the transition mappings

$$y \circ x^{-1} : x(\mathcal{U} \cap \mathcal{V}) \to y(\mathcal{U} \cap \mathcal{V}) \text{ and } x \circ y^{-1} : y(\mathcal{U} \cap \mathcal{V}) \to x(\mathcal{U} \cap \mathcal{V})$$

are smooth mappings between open subsets of V in Michal – Bastiani's sense, or $\mathcal{U} \cap \mathcal{V} = \emptyset$. A *$V$-atlas* on M is a family $\mathcal{A} = (\mathcal{U}_i, x_i)_{i\in I}$ of pairwise compatible V-charts of M such that the sets $\mathcal{U}_i$ form an open covering of M. A *smooth V-structure on M* is a maximal V-atlas, and a *smooth V-manifold* is a Hausdorff space endowed with a smooth V-structure. A smooth V-manifold is said to be a *Fréchet manifold*, a *Banach manifold* and an *n-manifold*, resp., if the model space V is a Fréchet space, a Banach space and an n-dimensional (real) vector space, resp. In the latter case, without loss of generality, the Euclidean n-space $\mathbb{R}^n$ can be chosen as a model space.

The concept of smoothness of a mapping between a V-manifold M and a W-manifold N can formally be defined in the same way as in the finite-dimensional case: a mapping $\varphi : M \to N$ is said to be *smooth* if it is continuous and, for every chart $(\mathcal{U}, x)$ on M and $(\mathcal{V}, y)$ on N, the mapping

$$y \circ \varphi \circ x^{-1} : x\left(\varphi^{-1}(\mathcal{V}) \cap \mathcal{U}\right) \subset V \to W$$

is Michal – Bastiani smooth. If, in addition, φ has a smooth inverse, then it is called a *diffeomorphism*. The space of all smooth mappings from M to N will be denoted by $C^\infty(M, N)$. In particular, $C^\infty(M)$ stands for the algebra of smooth functions on M, and

$$C_c^\infty(M) := \{f \in C^\infty(M) |\, \mathrm{supp}(f) \text{ is compact}\}.$$

Let $\pi : E \to M$ be a smooth mapping between smooth manifolds modeled on some locally convex TVSs, and let V be a fixed locally convex space. $\pi : E \to M$ is said to be a *vector bundle* with *typical fibre* V if the following conditions are satisfied:

(i) every *fibre* $E_p := \pi^{-1}(p)$, $p \in M$, is a locally convex TVS;

(ii) for each point $p \in M$ there is a neighbourhood $\mathcal{U}$ of p and a diffeomorphism $\Phi : \mathcal{U} \times V \to \pi^{-1}(\mathcal{U})$ such that $\pi \circ \Phi = \mathrm{pr}_1$, where $\mathrm{pr}_1 : \mathcal{U} \times V \to \mathcal{U}$ is the projection to the first factor, and the mappings

$$\Phi_q : V \to E_q, \quad v \mapsto \Phi_q(v) := \Phi(q, v), \quad q \in \mathcal{U}$$

are toplinear isomorphisms.

Further terminology: M is the *base manifold*, E is the *total manifold*, and π is the *projection* of the bundle. Φ is a *trivialization* of $\pi^{-1}(\mathcal{U})$ (or a *local trivialization* of E). For a vector bundle $\pi : E \to M$, we shall use both the abbreviation π (which is unambiguous) and the shorthand E (which may be ambiguous), depending on what we wish to emphasize. If the base manifold is finite dimensional, and the typical fibre is a k-dimensional vector space, then we shall speak of a vector bundle of *rank* k.

Let $\pi_1 : E_1 \to M_1$ and $\pi_2 : E_2 \to M_2$ be vector bundles. A smooth mapping $\varphi : E_1 \to E_2$ is called *fibre preserving* if $\pi_1(z_1) = \pi_1(z_2)$ implies $\pi_2(\varphi((z_1)) = \pi_2(\varphi(z_2))$ for all $z_1, z_2 \in E_1$. Then φ induces a smooth mapping $\underline{\varphi} : M_1 \to M_2$ such that $\pi_2 \circ \varphi = \underline{\varphi} \circ \pi_1$. A fibre preserving mapping $\varphi : E_1 \to E_2$ is said to be a *bundle map* if it restricts to continuous linear mappings $\varphi_1 : (E_1)_p \to (E_2)_{\underline{\varphi}(p)}$, $p \in M$. If, in addition, $M_1 = M_2 =: M$, and $\underline{\varphi} = 1_M$, then φ is called a *strong bundle map*.

Suppose $\pi : E \to M$ is a vector bundle with typical fibre V and $f : N \to M$ is a smooth mapping. For each $q \in N$, let $(f^*E)_q := \{q\} \times E_{f(q)}$ be endowed with the vector space structure inherited from $E_{f(q)}$:

$$(q, z_1) + (q, z_2) := (q, z_1 + z_2), \quad \lambda(q, z) := (q, \lambda z)$$

$\big(z_1, z_2, z \in E_{f(q)},\ \lambda \in \mathbb{R}\big)$. If

$$f^*E := \bigcup_{q \in N} (f^*E)_q = \{(q, z) \in N \times E | f(q) = \pi(z)\} =: N \times_M E,$$

then f^*E carries a unique smooth structure which makes it a vector bundle with base manifold N, projection $\pi_1 := \mathrm{pr}_1 \restriction N \times_M E$ and typical fibre V. If $\Phi : \mathcal{U} \times V \to \pi^{-1}(\mathcal{U})$ is a local trivialization of E, then the mapping

$$f^*\Phi : f^{-1}(\mathcal{U}) \times V \to \pi_1^{-1}(f^{-1}(\mathcal{U})), \quad (q, v) \mapsto (q, \Phi(f(q), v))$$

is a local trivialization for f^*E. The vector bundle $\pi_1 : f^*E \to N$ so obtained is said to be the *pull-back* of π over f, and it is also denoted by $f^*\pi$. Note that the mapping $\pi_2 := \mathrm{pr}_2 \restriction N \times_M E$ is a bundle map from f^*E to E which induces the given mapping f between the base manifolds.

For our next remarks, let us fix a vector bundle $\pi : E \to M$. A *section* of π is a smooth mapping $\sigma : M \to E$ with $\pi \circ \sigma = 1_M$; thus $\sigma(p) \in E_p$ for all $p \in M$. Similarly, a section of π over an open subset $\mathcal{U} \subset M$ is a smooth mapping $\sigma : \mathcal{U} \to E$ with $\pi \circ \sigma = 1_{\mathcal{U}}$. The *support* of a section $\sigma : \mathcal{U} \to E$ is $\mathrm{supp}(\sigma) := \overline{\{p \in \mathcal{U} | \sigma(p) \neq 0\}}$ (the closure is meant in $\mathcal{U}$). We denote by $\Gamma(\pi)$ (or $\Gamma(E)$) the set of sections of π. $\Gamma(\mathcal{U}, E)$ stands for the set of sections of E over $\mathcal{U}$; $\Gamma_c(\mathcal{U}, E) = \{\sigma \in \Gamma(\mathcal{U}, E) | \,\mathrm{supp}(\sigma) \text{ is compact}\}$. $\Gamma(\mathcal{U}, E)$, in particular, $\Gamma(\pi)$ is surely nonempty: we have the *zero section* $p \in \mathcal{U} \mapsto o(p) := 0_p :=$ *the zero vector of* E_p. $\overset{\circ}{E} := \cup_{p \in M}(E_p \setminus \{0_p\})$ will denote the *deleted bundle* for E, and $\overset{\circ}{\pi} := \pi \restriction \overset{\circ}{E}$. Sections (local sections with common domain) can be added to each other and multiplied by smooth functions on M using the standard pointwise definitions. These two operations make $\Gamma(\pi)$ and $\Gamma(\mathcal{U}, E)$ a $C^\infty(M)$-module and a $C^\infty(\mathcal{U})$-module, resp. In particular, $\Gamma(\pi)$ and $\Gamma(\mathcal{U}, E)$ (as well as $\Gamma_c(\mathcal{U}, E)$) are real vector spaces. Under some assumptions on the topological and the smooth structure of the base manifold and the TVS structure of the typical fibre, the spaces of sections can be endowed with 'similarly nice' TVS structure. For more information on this subtle problem see Kriegl – Michor's monograph [18, section 30]. In the next subsection we shall briefly discuss the finite dimensional case.

Now suppose that a smooth mapping $f : N \to M$ is also given, and consider the pullback bundle $\pi_1 : N \times_M E \to N$. A smooth mapping $S : N \to N \times_M E$ is a section of π_1 if and only if there is a smooth mapping $\underline{S} : N \to E$ such that $\pi \circ \underline{S} = f$ and $S(q) = (q, \underline{S}(q))$ for all $q \in N$; $\underline{S}$ is mentioned as the *principal part* of the section S. A smooth mapping $\underline{S} : N \to E$ satisfying $\pi \circ \underline{S} = f$ is called a *section of* E *along* f. Identifying a section of f^*E with its principal part, we get a canonical module isomorphism between $\Gamma(f^*E)$ and the module $\Gamma_f(E) = \Gamma_f(\pi)$ of sections of E along f.

2nd step: tangent bundle The crucial step in transporting of calculus from the model space V to a smooth V-manifold M is the construction of tangent vectors. Choose a point p of M and consider all triples of the form $(\mathcal{U}, x, a)$, where $(\mathcal{U}, x)$ is a chart around p and $a \in V$. The relation $\sim$ defined by

$$(\mathcal{U}, x, a) \sim (\mathcal{V}, y, b) :\Longleftrightarrow \left(y \circ x^{-1}\right)'(x(p))(a) = b$$

is an equivalence relation. The $\sim$-equivalence class $[(\mathcal{U}, x, a)]$ of a triple $(\mathcal{U}, x, a)$ is said to be a *tangent vector to* M *at* p. The set of tangent vectors to M at p is the *tangent space* to M at p; it will be denoted by T_pM. A chart $(\mathcal{U}, x)$ around p determines a bijection $\vartheta_p : T_pM \to V$ by the rule

$$v \in T_pM \mapsto \vartheta_p(v) := a \in V \quad \text{if } (\mathcal{U}, x, a) \in v.$$

There is a unique locally convex TVS structure on T_pM which makes the bijection ϑ_p a toplinear isomorphism. To be explicit, the linear structure of T_pM is given by

$$\lambda v + \mu w := \vartheta_p^{-1}(\lambda\vartheta_p(v) + \mu\vartheta_p(w)); \quad v, w \in T_pM; \quad \lambda, \mu \in \mathbb{R}.$$

The TVS structure so obtained on T_pM does not depend on the choice of $(\mathcal{U}, x)$, since if $(\mathcal{V}, y)$ is another chart around p, and $\eta_p : T_pM \to V$ is associated with $(\mathcal{V}, y)$, then $\eta_p \circ \vartheta_p^{-1} = \left(y \circ x^{-1}\right)'(x(p))$, which is a toplinear isomorphism.

Let $TM := \cup_{p \in M} T_pM$ (disjoint union), and define the mapping $\tau : TM \to M$ by $\tau(v) := p$ if $v \in T_pM$. *There is a unique smooth structure on* TM *which makes* $\tau : TM \to M$ *into a vector bundle with fibres* $(TM)_p = T_pM$ *and typical fibre* V. To indicate the construction of this smooth structure, let $(\mathcal{U}, x)$ be a chart on M, and assign to each $v \in \tau^{-1}(\mathcal{U}) \subset TM$ the pair $\big(x(\tau(v)), \vartheta_{\tau(v)}(v)\big) \in V \times V$. Thus we get a bijective mapping

$$\tau_{\mathcal{U}} := \big(x \circ \tau, \vartheta_{\tau(\cdot)}\big) : \tau^{-1}(\mathcal{U}) \to x(\mathcal{U}) \times V \subset V \times V.$$

As $(\mathcal{U}, x)$ runs over all charts of an atlas of M, the pairs $\big(\tau^{-1}(\mathcal{U}), \tau_{\mathcal{U}}\big)$ form an atlas for TM and make it into a smooth manifold modeled on $V \times V$ such that the topology of TM is the finest topology for which each mapping $\tau_{\mathcal{U}}$ is a homeomorphism. (For details see [8, 2.3].) The vector bundle so obtained is called the *tangent bundle* of M, and it is denoted by τ, TM or τ_M. If, in particular, $\mathcal{U}$ is an open subset of the model space V, then $\mathcal{U}$ (as well as V) can be regarded as a smooth manifold modeled on V. In this case there is a canonical identification $T\mathcal{U} \cong \mathcal{U} \times V$, which will be frequently used, without further mention.

Now take two manifolds M and N, modeled on V and W, respectively. Let $f : M \to N$ be a smooth mapping and $p \in M$. Choose charts $(\mathcal{U}, x)$ around p and $(\mathcal{V}, y)$ around $f(p)$ such that $f(\mathcal{U}) \subset \mathcal{V}$. Let $\vartheta_p : T_pM \to V$ and $\eta_{f(p)} : T_{f(p)}N \to W$ be the isomorphisms associated with $(\mathcal{U}, x)$ and $(\mathcal{V}, y)$. Then

$$(f_*)_p := \eta_{f(p)}^{-1} \circ \big(y \circ f \circ x^{-1}\big)' (x(p)) \circ \vartheta_p : T_pM \to T_{f(p)}N$$

is a well-defined linear mapping, called the *tangent map* of f *at* p. ('Well-defined' means that $(f_*)_p$ does not depend on the choice of $(\mathcal{U}, x)$ and $(\mathcal{V}, y)$.) Having the fibrewise tangent maps, we associate to $f : M \to N$ the bundle map

$$f_* : TM \to TN, \quad f_* \restriction T_pM := (f_*)_p.$$

Let, in particular, $f : M \to W$ be a smooth mapping. Then the mapping

$$df := \mathrm{pr}_2 \circ f_* : TM \to W \times W \to W$$

is called the *differential* of f. By restriction, it leads to a continuous linear mapping

$$df(p) = d_pf := df \restriction T_pM \to W,$$

for each $p \in M$.

The sections of the tangent bundle $\tau : TM \to M$ are said to be *vector fields* on M. We denote their $C^\infty(M)$-module by $\mathfrak{X}(M)$ rather than $\Gamma(\tau)$ or $\Gamma(TM)$. Any vector field X on M induces a derivation ϑ_X of the real algebra $C^\infty(M)$ by the rule

$$f \in C^\infty(M) \mapsto \vartheta_X(f) = X.f := df \circ X.$$

It may be shown that if $X, Y \in \mathfrak{X}(M)$, then there exists a unique vector field $[X, Y]$ on M such that on each open subset $\mathcal{U}$ of M we have

$$\vartheta_{[X,Y]}(f) = \vartheta_X(\vartheta_Y(f)) - \vartheta_Y(\vartheta_X(f))$$

for all $f \in C^\infty(\mathcal{U})$, and the mapping $[\cdot, \cdot] : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ makes $\mathfrak{X}(M)$ a (real) Lie algebra. For an accurate proof of these claims we refer to [8]. Lang's monograph [20] treats vector fields on Banach manifolds, while Klingenberg's book [17] deals with the Hilbertian case. Notice that if $\dim M = \infty$, then not all derivations of $C^\infty(M)$ can be described as above by vector fields.

Back to the finite dimension

In the subsequent concluding part of this section we shall consider *vector bundles of finite rank* over a common base manifold M. Then, by our convention, M is also finite dimensional; let $n := \dim M$, $n \in \mathbb{N}^*$. We assume, furthermore, that *the topology of M is second countable*. Then, as it is well-known, *M admits a (smooth) partition of unity*; in particular, for each neighbourhood of every point of M there is a bump function supported in the given neighbourhood. Under these conditions we have the following useful technical result:

Lemma 2.24 *Let $\pi : E \to M$ be a vector bundle of rank k over the n-dimensional second countable base manifold M, and let p be a point of M. Then for each $z \in E_p$ there is a section $\sigma \in \Gamma(\pi)$ such that $\sigma(p) = z$. If $\mathcal{U} \subset M$ is an open set containing p, and $\sigma : \mathcal{U} \to E$ is a section of π over $\mathcal{U}$, then there is a section $\tilde{\sigma} \in \Gamma(\pi)$ which coincides with σ in a neighbourhood of p.*

Proof. Choose a local trivialization $\Phi : \mathcal{V} \times \mathbb{R}^k \to \pi^{-1}(\mathcal{V})$ of E with $p \in \mathcal{V}$ and a bump function $f \in C^\infty(M)$ supported in $\mathcal{V}$ such that $f(p) = 1$. Given $z \in E_p$, there is a unique vector $v \in \mathbb{R}^k$ for which $\Phi(p, v) = z$. Define a mapping $\sigma : M \to E$ by

$$\sigma(q) := \begin{cases} \Phi(q, f(q)v) & \text{if } q \in \mathcal{V}, \\ 0 & \text{if } q \notin \mathcal{V}. \end{cases}$$

Then σ is clearly a (smooth) section of π with the desired property $\sigma(p) = z$, so our first claim is true. The second statement can be verified similarly. □

By a *frame* of $\pi : E \to M$ over an open subset $\mathcal{U}$ of M we mean a sequence $(\sigma_i)_{i=1}^k$ of sections of E over $\mathcal{U}$ such that $(\sigma_i(p))_{i=1}^k$ is a basis of E_p for all $p \in \mathcal{U}$. If the domain $\mathcal{U}$ of the sections σ_i is not specified, we shall speak of a *local frame*. *Local frames are actually the same objects as local trivializations.* Indeed, if $(\sigma_1, \dots, \sigma_k)$ is a frame of E over $\mathcal{U}$, then the mapping

$$\Phi : \mathcal{U} \times \mathbb{R}^k \to E, \quad \left(p, \textstyle\sum_{i=1}^k \nu^i e_i\right) \mapsto \sum_{i=1}^k \nu^i \sigma_i(p)$$

is a trivialization of $\pi^{-1}(\mathcal{U})$. Conversely, if $\Phi : \mathcal{U} \times \mathbb{R}^k \to E$ is a local trivialization of E, then the mappings

$$\sigma_i : p \in \mathcal{U} \mapsto \sigma_i(p) := \Phi(p, e_i) \quad (1 \leqq i \leqq k)$$

form a local frame of E.

Now consider two vector bundles, $\pi_1 : E_1 \to M$ and $\pi_2 : E_2 \to M$ of finite rank. A mapping $\Gamma(E_1) \to \Gamma(E_2)$ will be called *tensorial* if it is $C^\infty(M)$-linear. The following simple result is of basic importance and frequently (in general, tacitly) used.

Lemma 2.25 (the fundamental lemma of strong bundle maps) *A mapping $\mathcal{F} : \Gamma(\pi_1) \to \Gamma(\pi_2)$ is tensorial if and only if there is a strong bundle map $F : E_1 \to E_2$ such that $\mathcal{F}(\sigma) = F \circ \sigma$ for all $\sigma \in \Gamma(\pi_1)$.*

For a proof see e.g. John M. Lee's text [21]. In the following we shall usually identify a tensorial mapping $\mathcal{F} : \Gamma(\pi_1) \to \Gamma(\pi_2)$ with the corresponding bundle map F, and write $F\sigma$ rather than $\mathcal{F}(\sigma)$ or $F \circ \sigma$.

The concept of a linear differential operator introduced in 2.14 can immediately be generalized to the context of vector bundles. Our initial definition is strongly motivated by the property formulated in the local Peetre theorem 2.23.

Definition 2.26 An $\mathbb{R}$-linear mapping $D : \Gamma(\pi_1) \to \Gamma(\pi_2)$, $\sigma \mapsto D(\sigma) =: D\sigma$ is said to be a *linear differential operator* if it is support-decreasing, i.e., $\operatorname{supp}(D\sigma) \subset \operatorname{supp}(\sigma)$ for any section $\sigma \in \Gamma(\pi_1)$.

Linear differential operators are natural with respect to restrictions: if $\mathcal{U}$ is an open subset of M, and $D : \Gamma(\pi_1) \to \Gamma(\pi_2)$ is a linear differential operator, then (using the abbreviation $(\pi_i)_{\mathcal{U}} := \pi_i \restriction \pi_i^{-1}(\mathcal{U})$, $i \in \{1,2\}$) there is a unique differential operator $D_{\mathcal{U}} : \Gamma((\pi_1)_{\mathcal{U}}) \to \Gamma((\pi_2)_{\mathcal{U}})$ such that $D\sigma \restriction \mathcal{U} = D_{\mathcal{U}}(\sigma \restriction \mathcal{U})$ for every section $\sigma \in \Gamma(\pi_1)$. Indeed, let $p \in \mathcal{U}$ be an arbitrary point. By Lemma 2.24, for any section $\sigma_{\mathcal{U}} \in \Gamma((\pi_1)_{\mathcal{U}})$ there is a section σ of π_1 such that $\sigma = \sigma_{\mathcal{U}}$ in a neighbourhood of p. If $(D_{\mathcal{U}}\sigma_{\mathcal{U}})(p) := (D\sigma)(p)$, then $D_{\mathcal{U}} : \Gamma((\pi_1)_{\mathcal{U}}) \to \Gamma((\pi_2)_{\mathcal{U}})$ is a well-defined linear differential operator with the desired naturality property. An equivalent formulation: D is a *local operator* in the sense that *for each open subset $\mathcal{U}$ of M and each section $\sigma \in \Gamma(\pi_1)$ such that $\sigma \restriction \mathcal{U} = 0$, we have $(D\sigma) \restriction \mathcal{U} = 0$.*

Let $(\mathcal{U}, x)$ be a chart on M. Suppose that π_1 and π_2 are trivializable over $\mathcal{U}$, i.e., there exist trivializations $\Phi_1 : \mathcal{U} \times \mathbb{R}^k \to \pi_1^{-1}(\mathcal{U})$ and $\Phi_2 : \mathcal{U} \times \mathbb{R}^\ell \to \pi_2^{-1}(\mathcal{U})$. If $\sigma \in \Gamma(\pi_1)$, then $\tilde{\sigma} := \operatorname{pr}_2 \circ \Phi_1^{-1} \circ \sigma \circ x^{-1}$ is a smooth mapping from $x(\mathcal{U}) \subset \mathbb{R}^n$ to $\mathbb{R}^k$, and the mapping $\sigma \in \Gamma(\pi_1) \mapsto \tilde{\sigma} \in C^\infty\left(x(\mathcal{U}), \mathbb{R}^k\right)$ is a linear isomorphism. Let $D : \Gamma(\pi_1) \to \Gamma(\pi_2)$ be a linear differential operator, and consider the induced operator $D_{\mathcal{U}} : \Gamma((\pi_1)_{\mathcal{U}}) \to \Gamma((\pi_2)_{\mathcal{U}})$. There exists a well-defined continuous linear mapping $\tilde{D}$ from the Fréchet space $C^\infty\left(x(\mathcal{U}), \mathbb{R}^k\right)$ into the Fréchet space $C^\infty\left(x(\mathcal{U}), \mathbb{R}^\ell\right)$ such that

$$\operatorname{pr}_2 \circ \Phi_2^{-1} \circ D_{\mathcal{U}}(\sigma \restriction \mathcal{U}) \circ x^{-1} = \tilde{D}\left(\operatorname{pr}_2 \circ \Phi_1^{-1} \circ (\sigma \restriction \mathcal{U}) \circ x^{-1}\right);$$

briefly $\widetilde{D_{\mathcal{U}}\sigma} = \tilde{D}(\tilde{\sigma})$. $\tilde{D}$ is said to be the *local expression* of D with respect to the chart $(\mathcal{U}, x)$ and the trivializations Φ_1, Φ_2.

Now we are in a position to transpose Peetre's local theorem (2.23) to the context of vector bundles.

Theorem 2.27 (Peetre's theorem) *Let π_1 and π_2 be vector bundles over M of rank k and ℓ, respectively. If $D : \Gamma(\pi_1) \to \Gamma(\pi_2)$ is a linear differential operator, then there exists a chart $(\mathcal{U}, x)$ at each point of M, such that π_1 and π_2 are trivializable over $\mathcal{U}$, and the corresponding local expression of D is of the form*

$$f \in C^\infty\left(x(\mathcal{U}), \mathbb{R}^k\right) \mapsto \sum_{|\alpha| \leqq m} A_\alpha \circ D^\alpha f,$$

where for each multi-index α such that $|\alpha| \leqq m$, A_α is a smooth mapping from $x(\mathcal{U}) \subset \mathbb{R}^n$ to $\mathcal{L}\left(\mathbb{R}^k, \mathbb{R}^\ell\right)$. □

Peetre's theorem makes it possible to define the *order* of a linear differential operator $D : \Gamma(\pi_1) \to \Gamma(\pi_2)$ *at a point* $p \in M$ as the largest $m \in \mathbb{N}$ for which there exists a multi-index α such that $|\alpha| = m$, and $A_\alpha(p) \neq 0$ in a local expression of D in a neighbourhood of p. It follows at once that if D is *of order* 0, then it may be identified with a strong bundle map $E_1 \to E_2$, therefore it acts by the rule $\sigma \in \Gamma(\pi_1) \mapsto D \circ \sigma \in \Gamma(\pi_2)$. Equivalently, in view of Lemma 2.25, D can be considered as a tensorial mapping from $\Gamma(\pi_1)$ to $\Gamma(\pi_2)$.

For the rest of this section, we let $\pi : E \to M$ be a vector bundle of rank k over an n-dimensional base manifold M. We continue to assume that M admits a partition of unity.

An important class of first order differential operators from $\Gamma(\pi)$ to $\Gamma(\pi)$ may be specified by introducing the concept of covariant differentiation. The notion is well-known, but fundamental, so we briefly recall that a mapping

$$D : \mathfrak{X}(M) \times \Gamma(\pi) \to \Gamma(\pi), \quad (X, \sigma) \mapsto D_X\sigma$$

is said to be a *covariant derivative* on E if it is tensorial in X and derivation in σ. $D_X\sigma$ is called the covariant derivative of σ in the direction of X, while for any section $\sigma \in \Gamma(\pi)$, the mapping

$$D\sigma : \mathfrak{X}(M) \to \Gamma(\pi), \quad \sigma \mapsto D_X\sigma$$

is the *covariant differential* of σ. Then $D\sigma \in \mathcal{L}_{C^\infty(M)}(\mathfrak{X}(M), \Gamma(\pi))$. For every fixed vector field X on M, the mapping

$$D_X : \Gamma(\pi) \to \Gamma(\pi), \quad \sigma \mapsto D_X\sigma$$

is obviously $\mathbb{R}$-linear. Moreover, *D_X is a local operator.* Indeed, let $\mathcal{U} \subset M$ be an open set, and choose a point $p \in M$. The existence of bump functions supported in $\mathcal{U}$ also guarantees that there is a smooth function f on M such that $f(p) = 0$ and $f = 1$ outside $\mathcal{U}$. If $\sigma \in \Gamma(\pi)$, and $\sigma \restriction \mathcal{U} = 0$, then $\sigma = f\sigma$, and

$$(D_X\sigma)(p) = (D_X(f\sigma))(p) = (Xf)(p)\sigma(p) + f(p)(D_X\sigma)(p) = 0,$$

therefore $D_X\sigma \restriction \mathcal{U} = 0$. $\mathbb{R}$-linearity and locality imply that D_X is a linear differential operator. To see that it is of first order, let $\left(\mathcal{U}, \left(u^i\right)_{i=1}^n\right)$ be a chart on M, and $(\sigma_j)_{j=1}^k$ a frame over $\mathcal{U}$. If $X \in \mathfrak{X}(M)$, $\sigma \in \Gamma(\pi)$, then $X \restriction \mathcal{U} = \sum_{i=1}^n X^i \frac{\partial}{\partial u^i}$, $\sigma \restriction \mathcal{U} = \sum_{r=1}^k f^r \sigma_r \left(X^i, f^r \in C^\infty(\mathcal{U});\ 1 \leqq i \leqq n, 1 \leqq r \leqq k\right)$, and by the local character of D_X,

$$(D_X\sigma) \restriction \mathcal{U} = D_{X \restriction \mathcal{U}}(\sigma \restriction \mathcal{U}) = \sum_{i=1}^n \sum_{r=1}^k X^i \left(\frac{\partial f^r}{\partial u^i} \sigma_r + f^r D_{\frac{\partial}{\partial u^i}} \sigma_r \right).$$

The local sections $D_{\frac{\partial}{\partial u^i}}\sigma_r$ can be combined from the frame $(\sigma_j)_{j=1}^k$ in the form

$$D_{\frac{\partial}{\partial u^i}}\sigma_r = \sum_{s=1}^k \Gamma_{ir}^s \sigma_s, \quad \Gamma_{ir}^s \in C^\infty(\mathcal{U});$$

so we get

$$(D_X\sigma) \restriction \mathcal{U} = \sum_{i=1}^n \sum_{s=1}^k X^i \left(\frac{\partial f^s}{\partial u^i} + \textstyle\sum_{r=1}^k \Gamma_{ir}^s f^r \right) \sigma_s.$$

Since $\frac{\partial f^s}{\partial u^i} := D_i\left(f^s \circ u^{-1}\right) \circ u$, D_X is indeed of first order. Notice that the functions $\Gamma_{ir}^s \in C^\infty(\mathcal{U})$ $(1 \leqq i \leqq n;\ 1 \leqq r, s \leqq k)$ are called the *Christoffel symbols* of D with respect to the chart $\left(\mathcal{U}, \left(u^i\right)_{i=1}^n\right)$ and local frame $(\sigma_i)_{i=1}^k$.

3 The Chern – Rund derivative

Conventions

Throughout this section we shall work over an n-dimensional ($n \in \mathbb{N}^*$) smooth manifold M admitting a partition of unity. More precisely, the main scene of our next considerations will be the tangent bundle $\tau : TM \to M$ of M and the pull-back bundle $\tau_1 : \tau^*TM = TM \times_M TM \to TM$. We shall also need the tangent bundle $\tau_{TM} : TTM \to TM$ of TM, the deleted bundle $\overset{\circ}{\tau} : \overset{\circ}{T}M \to M$ $\left(\overset{\circ}{T}M := \{v \in TM | v \neq 0\}, \overset{\circ}{\tau} := \tau \restriction \overset{\circ}{T}M\right)$ for τ and the pull-back $\overset{\circ}{\tau}{}^*TM = \overset{\circ}{T}M \times_M TM$ of TM via $\overset{\circ}{\tau}$. For the $C^\infty(TM)$-modules $\Gamma(\tau^*TM) \cong \Gamma_\tau(TM)$ and $\Gamma\left(\overset{\circ}{\tau}{}^*TM\right) \cong \Gamma_{\overset{\circ}{\tau}}(TM)$ we use the convenient notations $\mathfrak{X}(\tau)$ and $\mathfrak{X}\left(\overset{\circ}{\tau}\right)$, resp. Any vector field X on M induces a vector field $\hat{X}$ along τ and $\overset{\circ}{\tau}$ with principal parts $X \circ \tau$ and $X \circ \overset{\circ}{\tau}$, resp. $\hat{X}$ is called a *basic vector field* along τ (or $\overset{\circ}{\tau}$). Locally, the basic vector fields generate the modules $\mathfrak{X}(\tau)$ and $\mathfrak{X}\left(\overset{\circ}{\tau}\right)$. A distinguished vector field along τ is the *canonical vector field*

$$\delta : v \in TM \mapsto \delta(v) := (v, v) \in TM \times_M TM.$$

Generic vector fields along τ (or $\overset{\circ}{\tau}$) will be denoted by $\tilde{X}, \tilde{Y}, \dots$. From the module $\mathfrak{X}(\tau)$ (or $\mathfrak{X}\left(\overset{\circ}{\tau}\right)$) one can build the spaces of type (r, s) tensors along τ (or $\overset{\circ}{\tau}$). These $C^\infty(TM)$-multilinear machines can also be interpreted as 'fields'. For example, a type $(0, 2)$ tensor field $g : \mathfrak{X}\left(\overset{\circ}{\tau}\right) \times \mathfrak{X}\left(\overset{\circ}{\tau}\right) \to C^\infty\left(\overset{\circ}{T}M\right)$ can be regarded as a mapping

$$v \in \overset{\circ}{T}M \mapsto g_v \in \mathcal{L}^2\left(T_{\overset{\circ}{\tau}(v)}M, \mathbb{R}\right)$$

which has the following smoothness property: the function

$$g\left(\tilde{X}, \tilde{Y}\right) : v \in \overset{\circ}{T}M \mapsto g\left(\tilde{X}, \tilde{Y}\right)(v) := g_v\left(\tilde{X}(v), \tilde{Y}(v)\right)$$

is smooth for any two vector fields $\tilde{X}, \tilde{Y}$ along $\overset{\circ}{\tau}$.

For coordinate calculations we choose a chart $\left(\mathcal{U}, \left(u^i\right)_{i=1}^n\right)$ on M, and employ the induced chart

$$\left(\tau^{-1}(\mathcal{U}), \left(x^i, y^i\right)\right); \quad x^i := u^i \circ \tau, \quad y^i : v \in \tau^{-1}(\mathcal{U}) \mapsto y^i(v) := v\left(u^i\right)$$

($1 \leqq i \leqq n$) on TM. The coordinate vector fields

$$\frac{\partial}{\partial u^i} : f \in C^\infty(\mathcal{U}) \mapsto \frac{\partial f}{\partial u^i} := D_i\left(f \circ u^{-1}\right) \circ u \in C^\infty(\mathcal{U}) \quad (1 \leqq i \leqq n)$$

form a frame of $\tau : TM \to M$ over $\mathcal{U}$. Similarly, $\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial y^i}\right)_{i=1}^n$ is a local frame of $\tau_{TM} : TTM \to TM$. The basic vector fields

$$\widehat{\frac{\partial}{\partial u^i}} : v \in \tau^{-1}(\mathcal{U}) \mapsto \left(v, \left(\frac{\partial}{\partial u^i}\right)_{\tau(v)}\right) \quad (1 \leqq i \leqq n)$$

provide a local frame for τ^*TM. Using this frame, over $\tau^{-1}(\mathcal{U})$ we have

$$\hat{X} = \sum_{i=1}^{n} \left(X^i \circ \tau\right) \widehat{\frac{\partial}{\partial u^i}} \quad \text{if } X \in \mathfrak{X}(M),\ X \restriction \mathcal{U} = \sum_{i=1}^{n} X^i \frac{\partial}{\partial u^i};$$
$$\delta = \sum_{i=1}^{n} y^i \widehat{\frac{\partial}{\partial u^i}}.$$

The coordinate expression of a generic section $\tilde{X}$ of τ^*TM is

$$\tilde{X} \restriction \tau^{-1}(\mathcal{U}) = \sum_{i=1}^{n} \tilde{X}^i \widehat{\frac{\partial}{\partial u^i}}, \quad \tilde{X}^i \in C^\infty\left(\tau^{-1}(\mathcal{U})\right) \quad (1 \leqq i \leqq n).$$

Canonical constructions on TM and τ^*TM

We begin with a frequently used, simple observation. Let V be an n-dimensional real vector space, endowed with the canonical smooth structure determined by a linear isomorphism of V onto $\mathbb{R}^n$. *For any* $p \in V$, V *may be naturally identified with its tangent space* T_pV via the linear isomorphism

$$\imath_p : v \in V \mapsto \imath_p(v) := \dot{\varrho}(0) \in T_pV, \quad \varrho(t) := p + tv \quad (t \in \mathbb{R})$$

($\dot{\varrho}(0)$ is the tangent vector of ϱ at 0 in the sense of classical manifold theory). If $(e_i)_{i=1}^n$ is a basis of V, and $\left(e^i\right)_{i=1}^n$ is its dual, then

$$\imath_p(v) = \sum_{i=1}^{n} e^i(v) \left(\frac{\partial}{\partial e^i}\right)_p.$$

By means of these identifications, we get an injective strong bundle map

$$\mathbf{i} : TM \times_M TM \to TTM,\ (v, w) \in \{v\} \times T_{\tau(v)}M \mapsto \mathbf{i}(v, w) := \imath_v(w) \in T_vT_{\tau(v)}M.$$

$\mathrm{Im}(\mathbf{i}) =: VTM$ is said to be the *vertical bundle* of $\tau_{TM} : TTM \to TM$. It is easy to check that $VTM = \mathrm{Ker}(\tau_*)$. By Lemma 2.25, $\mathbf{i}$ may be interpreted as a tensorial mapping from $\mathfrak{X}(\tau)$ to $\mathfrak{X}(TM)$ denoted by the same symbol. $\mathfrak{X}^v(TM) := \mathbf{i}(\mathfrak{X}(\tau))$ is the module of *vertical vector fields* on TM (in fact, $\mathfrak{X}^v(TM)$ is a Lie-subalgebra of $\mathfrak{X}(TM)$). In particular, $X^v := \mathbf{i}\hat{X}$ is said to be the *vertical lift* of $X \in \mathfrak{X}(M)$; $C := \mathbf{i}\delta$ is the *Liouville vector field* on TM. It is easy to check that

$$[X^v, Y^v] = 0, \quad [C, X^v] = -X^v; \quad X, Y \in \mathfrak{X}(M).$$

In terms of local coordinates,

$$\mathbf{i}(v, w) = \sum_{i=1}^{n} y^i(w) \left(\frac{\partial}{\partial y^i}\right)_v \quad (\tau(v) = \tau(w));$$
$$\mathbf{i}\left(\widehat{\frac{\partial}{\partial u^i}}\right)_v = \mathbf{i}\left(v, \left(\frac{\partial}{\partial u^i}\right)_{\tau(v)}\right) = \left(\frac{\partial}{\partial y^i}\right)_v,$$

$$\text{hence } \left(\frac{\partial}{\partial u^i}\right)^v = \frac{\partial}{\partial y^i} \quad (1 \leqq i \leqq n);$$

$$X^v = \sum_{i=1}^n \left(X^i \circ \tau\right) \frac{\partial}{\partial y^i} \quad \text{if } X \restriction \mathcal{U} = \sum_{i=1}^n X^i \frac{\partial}{\partial u^i}.$$

A further and surjective strong bundle map is

$$\mathbf{j} := (\tau_{TM}, \tau_*) : TTM \to TM \times_M TM, \quad z \in T_vTM \mapsto \mathbf{j}(z) := (v, \tau_*(z)),$$

which can also be regarded as a tensorial mapping from $\mathfrak{X}(TM)$ to $\mathfrak{X}(\tau)$. $\mathbf{j}$ acts on the local frame $\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial y^i}\right)_{i=1}^n$ by

$$\mathbf{j}\left(\frac{\partial}{\partial x^i}\right) = \widehat{\frac{\partial}{\partial u^i}}, \quad \mathbf{j}\left(\frac{\partial}{\partial y^i}\right) = 0 \quad (1 \leqq i \leqq n),$$

therefore $\mathrm{Ker}(\mathbf{j}) = \mathrm{Im}(\mathbf{i}) = \mathfrak{X}^v(TM)$. The composition $J := \mathbf{i} \circ \mathbf{j}$ is said to be the *vertical endomorphism* of $\mathfrak{X}(TM)$ (or TTM). It follows that

$$\mathrm{Im}(J) = \mathrm{Ker}(J) = \mathfrak{X}^v(TM), \quad J^2 = 0.$$

By the *complete lift* of a smooth function f on M we mean the function $f^c : v \in TM \mapsto f^c(v) := v(f) \in \mathbb{R}$. Then, obviously, $f^c \in C^\infty(TM)$. It can be shown that *for any vector field X on M there exists a unique vector field X^c on TM such that $X^c f^c = (Xf)^c$ for all $f \in C^\infty(M)$* [35]. X^c is said to be the *complete lift* of X. A great deal of calculations may be simplified by the fact that *if $(X_i)_{i=1}^n$ is a local frame of TM, then $(X_i^v, X_i^c)_{i=1}^n$ is a local frame of TTM.*

It follows immediately that $\mathbf{j}X^c = \hat{X}$, or, equivalently, $JX^c = X^v$. Concerning the vertical and the complete lifts, we have

$$[X^c, Y^c] = [X,Y]^c, \quad [X^v, Y^c] = [X,Y]^v, \quad [C, X^c] = 0; \quad X, Y \in \mathfrak{X}(M).$$

We shall need, furthermore, the following

Lemma 3.1 *If X and Z are vector fields on M, and F is a smooth function on TM, then*

$$X(F \circ Z) = (X^c F + [X,Z]^v F) \circ Z.$$

The simplest, but not too aesthetic way to prove this relation is to express everything in terms of local coordinates.

On τ^*TM there exists a canonical differential operator of first order which makes it possible to differentiate tensors along τ in vertical directions. We call this operator the *canonical v-covariant derivative*, and we denote it by ∇^v. It can explicitly be given as follows: for each section $\tilde{X}$ in $\mathfrak{X}(\tau)$,

$$\nabla^v_{\tilde{X}} F := \left(\mathbf{i}\tilde{X}\right) F \qquad \text{if } F \in C^\infty(TM);$$

$$\nabla^v_{\tilde{X}} \tilde{Y} := \mathbf{j}\left[\mathbf{i}\tilde{X}, \eta\right] \qquad \text{if } \tilde{Y} \in \mathfrak{X}(\tau) \text{ and } \eta \in \mathfrak{X}(TM) \text{ such that } \mathbf{j}\eta = \tilde{Y}.$$

Using the frame $\left(\widehat{\frac{\partial}{\partial u^i}}\right)_{i=1}^n$ over $\tau^{-1}(\mathcal{U})$, if $\tilde{X} \restriction \tau^{-1}(\mathcal{U}) = \sum_{i=1}^n \tilde{X}^i \widehat{\frac{\partial}{\partial u^i}}$ and $\tilde{Y} \restriction \tau^{-1}(\mathcal{U}) = \sum_{i=1}^n \tilde{Y}^i \widehat{\frac{\partial}{\partial u^i}}$, then

$$\nabla^v_{\tilde{X}} F \restriction \tau^{-1}(\mathcal{U}) = \sum_{i=1}^n \tilde{X}^i \frac{\partial F}{\partial y^i}, \quad \nabla^v_{\tilde{X}} \tilde{Y} \restriction \tau^{-1}(\mathcal{U}) = \sum_{i,j=1}^n \tilde{X}^i \frac{\partial \tilde{Y}^j}{\partial y^i} \widehat{\frac{\partial}{\partial u^j}}.$$

From the last expression it is clear that $\nabla^v_{\tilde{X}} \tilde{Y}$ is well-defined: it does not depend on the choice of η. The mapping $\nabla^v : \left(\tilde{X}, \tilde{Y}\right) \in \mathfrak{X}(\tau) \times \mathfrak{X}(\tau) \to \nabla^v_{\tilde{X}} \tilde{Y} \in \mathfrak{X}(\tau)$ has the formal properties of a covariant derivative operator: it is tensorial in $\tilde{X}$ and satifies the derivation rule

$$\nabla^v_{\tilde{X}} F\tilde{Y} = \left(\nabla^v_{\tilde{X}} F\right) \tilde{Y} + F \nabla^v_{\tilde{X}} \tilde{Y}, \quad F \in C^\infty(TM).$$

From the very definition, or using the coordinate expression, it can easily be deduced that

$$\nabla^v_{\tilde{X}} \hat{Y} = 0, \quad \nabla^v_{\tilde{X}} \delta = \tilde{X}; \quad \tilde{X} \in \mathfrak{X}(\tau),\ Y \in \mathfrak{X}(M).$$

Using an appropriate version of Willmore's theorem on tensor derivations (see e.g. [35, 1.32]), ∇^v can uniquely be extended to a tensor derivation of the tensor algebra of $\mathfrak{X}(\tau)$. For any tensor $\tilde{A}$ along τ we may also consider the (canonical) v-covariant differentials of $\tilde{A}$ by the rule $i_{\tilde{X}} \nabla^v \tilde{A} := \nabla^v_{\tilde{X}} \tilde{A}$ ($i_{\tilde{X}}$ denotes the substitution operator associated to $\tilde{X}$). If, in particular, $F \in C^\infty(TM)$, then $\nabla^v F$ is a one-form, $\nabla^v \nabla^v F := \nabla^v(\nabla^v F)$ is a type $(0,2)$ tensor along τ. $\nabla^v \nabla^v F$ is called the (vertical) *Hessian* of F. For any two vector fields X, Y on M we have

$$\begin{aligned}\nabla^v \nabla^v F\left(\hat{X}, \hat{Y}\right) &= X^v(Y^v F) = [X^v, Y^v]F + Y^v(X^v F) \\ &= Y^v(X^v F) = \nabla^v \nabla^v F\left(\hat{Y}, \hat{X}\right),\end{aligned}$$

so the tensor $\nabla^v \nabla^v F$ is *symmetric*.

Ehresmann connections and Berwald derivatives

Definition 3.2 By an *Ehresmann connection* on TM we mean a mapping $\mathcal{H} : TM \times_M TM \to TTM$ satisfying the following conditions:

(C$_1$) $\mathcal{H} \restriction \overset{\circ}{T}M \times_M TM$ is smooth;

(C$_2$) for all $v \in \overset{\circ}{T}M$, $\mathcal{H} \restriction \{v\} \times T_{\tau(v)}M$ is a linear mapping to T_vTM;

(C$_3$) $\mathbf{j} \circ \mathcal{H} = 1_{TM \times_M TM}$;

(C$_4$) if $o : M \to TM$ is the zero section of TM, then $\mathcal{H}(o(p), v) = (o_*)_p(v)$ for all $p \in M, v \in T_pM$.

Then (C_3) and (C_4) are clearly consistent, however, *the smoothness of $\mathcal{H}$ (and the objects derived from $\mathcal{H}$) is not guaranteed on its whole domain.* (This weakening of smoothness allows more flexibility in applications.) If $HTM := \mathrm{Im}(\mathcal{H})$, then $TTM = HTM \oplus VTM$ (Whitney sum); HTM is said to be a *horizontal subbundle* of TTM.

Given an Ehresmann connection $\mathcal{H}$ on TM, there exists a unique strong bundle map $\mathcal{V} : TTM \to TM \times_M TM$, smooth in general only on $T\overset{\circ}{T}M$, such that

$$\mathcal{V} \circ \mathbf{i} = 1_{TM \times_M TM} \quad \text{and} \quad \mathrm{Ker}(\mathcal{V}) = \mathrm{Im}(\mathcal{H}).$$

$\mathcal{V}$ is called the *vertical map* associated to $\mathcal{H}$. $\mathbf{h} := \mathcal{H} \circ \mathbf{j}$ and $\mathbf{v} := \mathbf{i} \circ \mathcal{V}$ are projectors on TTM, the *horizontal* and *vertical projector* belonging to $\mathcal{H}$, respectively. $\mathcal{H}$ (as well as $\mathcal{V}$, $\mathbf{h}$ and $\mathbf{v}$) induce tensorial mappings at the level of sections. $\mathfrak{X}^h(TM) := \mathcal{H}(\mathfrak{X}(\tau)) = \mathbf{h}(\mathfrak{X}(TM))$ is called the space of *horizontal vector fields* on TM. In particular, $X^h := \mathcal{H}(\hat{X}) = \mathbf{h}X^c$ is the *horizontal lift* of $X \in \mathfrak{X}(M)$. We have: $\mathfrak{X}(TM) = \mathfrak{X}^h(TM) \oplus \mathfrak{X}^v(TM)$ (direct sum of modules). In terms of the local frames $\left(\widehat{\frac{\partial}{\partial u^i}}\right)_{i=1}^n$ of τ^*TM and $\left(\left(\frac{\partial}{\partial x^i}\right)_{i=1}^n, \left(\frac{\partial}{\partial y^i}\right)_{i=1}^n\right)$ of TTM, we get the following coordinate expressions:

$$\mathcal{H}\left(\widehat{\frac{\partial}{\partial u^i}}\right) = \frac{\partial}{\partial x^i} - \sum_{j=1}^n N_i^j \frac{\partial}{\partial y^j}; \quad \mathcal{V}\left(\frac{\partial}{\partial x^i}\right) = \sum_{j=1}^n N_i^j \widehat{\frac{\partial}{\partial u^j}}, \quad \mathcal{V}\left(\frac{\partial}{\partial y^i}\right) = \widehat{\frac{\partial}{\partial u^i}};$$

$$\mathbf{h}\left(\frac{\partial}{\partial x^i}\right) = \frac{\partial}{\partial x^i} - \sum_{j=1}^n N_i^j \frac{\partial}{\partial y^j}, \quad \mathbf{h}\left(\frac{\partial}{\partial y^i}\right) = 0;$$

$$\mathbf{v}\left(\frac{\partial}{\partial x^i}\right) = \sum_{j=1}^n N_i^j \frac{\partial}{\partial y^j}, \quad \mathbf{v}\left(\frac{\partial}{\partial y^i}\right) = \frac{\partial}{\partial y^i} \quad (1 \leqq i \leqq n).$$

The functions N_i^j, defined on $\tau^{-1}(\mathcal{U})$ and smooth on $\overset{\circ}{\tau}{}^{-1}(\mathcal{U})$, are called the *Christoffel symbols* of $\mathcal{H}$ with respect to the given local frames. (The minus sign in the first formula is more or less traditional.)

Via linearization, any Ehresmann connection $\mathcal{H}$ leads to a covariant derivative $\nabla : \mathfrak{X}\left(\overset{\circ}{T}M\right) \times \mathfrak{X}(\overset{\circ}{\tau}) \to \mathfrak{X}(\overset{\circ}{\tau})$ on τ^*TM, called the *Berwald derivative* induced by $\mathcal{H}$. The explicit rules of calculation are

$$\nabla_{\mathcal{H}\tilde{X}}\tilde{Y} := \mathcal{V}\left[\mathcal{H}\tilde{X}, \mathbf{i}\tilde{Y}\right] \quad \text{and} \quad \nabla_{\mathbf{i}\tilde{X}}\tilde{Y} := \nabla^v_{\tilde{X}}\tilde{Y}; \quad \tilde{X}, \tilde{Y} \in \mathfrak{X}(\overset{\circ}{\tau}).$$

The *h-part* of ∇, given by $\nabla^h_{\tilde{X}}\tilde{Y} := \nabla_{\mathcal{H}\tilde{X}}\tilde{Y}$, has the coordinate expression

$$\nabla^h_{\widehat{\frac{\partial}{\partial u^i}}} \widehat{\frac{\partial}{\partial u^j}} = \nabla_{\left(\frac{\partial}{\partial u^i}\right)^h} \widehat{\frac{\partial}{\partial u^j}} = \sum_{k=1}^n \frac{\partial N_j^k}{\partial y^i} \widehat{\frac{\partial}{\partial u^k}} \quad (1 \leqq i, j \leqq n),$$

where the functions N_j^k are the Christoffel symbols of $\mathcal{H}$. If $\tilde{A}$ is any tensor along $\overset{\circ}{\tau}$, we may also consider the *h-Berwald differential* $\nabla^h\tilde{A}$ of $\tilde{A}$ given by $\nabla^h\tilde{A}(\tilde{X}) := \nabla^h_{\tilde{X}}\tilde{A}$. $\mathbf{t} := \nabla^h\delta$ is said to be the *tension* of $\mathcal{H}$. Then for any section $\tilde{X}$ along $\overset{\circ}{\tau}$ we have

$$\mathbf{t}(\tilde{X}) = \nabla^h\delta(\tilde{X}) = \nabla_{\mathcal{H}\tilde{X}}\delta = \mathcal{V}\left[\mathcal{H}\tilde{X}, C\right].$$

$\mathcal{H}$ is called *homogeneous* if $\mathbf{t} = 0$. If a homogeneous Ehresmann connection is of class C^1 at the zeros, then there exists a covariant derivative D on M (more precisely, on TM) such that for all $X, Y \in \mathfrak{X}(M)$,

$$(D_X Y)^v = \left[X^h, Y^v\right].$$

So *under homogeneity and C^1-differentiability Ehresmann connections lead to classical covariant derivatives on the base manifold.*

For covariant derivatives given on a generic vector bundle there is no reasonable concept of 'torsion'. However, if D is a covariant derivative on the pull-back bundle $\tau^* TM$, and an Ehresmann connection $\mathcal{H}$ is also specified on TM, we have useful generalizations of the classical torsion: the *vertical torsion* $T^v(D)$ and the *horizontal torsion* $T^h(D)$ given by

$$T^v(D)\left(\tilde{X}, \tilde{Y}\right) := D_{\mathbf{i}\tilde{X}}\tilde{Y} - D_{\mathbf{i}\tilde{Y}}\tilde{X} - \mathbf{i}^{-1}\left[\mathbf{i}\tilde{X}, \mathbf{i}\tilde{Y}\right]$$

and

$$T^h(D)\left(\tilde{X}, \tilde{Y}\right) := D_{\mathcal{H}\tilde{X}}\tilde{Y} - D_{\mathcal{H}\tilde{Y}}\tilde{X} - \mathbf{j}\left[\mathcal{H}\tilde{X}, \mathcal{H}\tilde{Y}\right],$$

respectively. (It is easy to check that both $T^v(D)$ and $T^h(D)$ are tensorial.)

The vertical torsion of any Berwald derivative ∇ vanishes. Indeed, for all $\xi, \eta \in \mathfrak{X}(TM)$ we have

$$\begin{aligned}\mathbf{i}(T^v(\nabla)(\mathbf{j}\xi, \mathbf{j}\eta)) &= \mathbf{i}\nabla_{J\xi}\mathbf{j}\eta - \mathbf{i}\nabla_{J\eta}\mathbf{j}\xi - [J\xi, J\eta] \\ &= J[J\xi, \eta] - J[J\eta, \xi] - [J\xi, J\eta] = -N_J(\xi, \eta),\end{aligned}$$

where N_J is the *Nijenhuis torsion* of J. However, as it is well-known, $N_J = 0$, therefore $T^v(\nabla) = 0$.

The horizontal torsion of the Berwald derivative induced by an Ehresmann connection $\mathcal{H}$ is said to be the *torsion of* $\mathcal{H}$. Denoting this tensor by $\mathbf{T}$, we get

$$\mathbf{i}\mathbf{T}\left(\hat{X}, \hat{Y}\right) = \left[X^h, Y^v\right] - \left[Y^h, X^v\right] - [X, Y]^v; \quad X, Y \in \mathfrak{X}(M).$$

If $\mathcal{H}$ is of class C^1 and homogeneous, then there is a covariant derivative D on M such that $(D_X Y)^v = \left[X^h, Y^v\right]$ $(X, Y \in \mathfrak{X}(M))$; hence

$$\mathbf{i}\mathbf{T}\left(\hat{X}, \hat{Y}\right) = (D_X Y - D_Y X - [X, Y])^v =: (T(D))^v,$$

so $\mathbf{T}$ reduces to the usual torsion of D.

Definition 3.3 By a *semispray* on TM we mean a mapping $S : TM \to TTM$ such that $\tau_{TM} \circ S = 1_{TM}$, $JS = C$ (or, equivalently, $\mathbf{j}S = \delta$) and $S \restriction \overset{\circ}{T}M$ is smooth. A semispray is said to be a *spray* if it is of class C^1 on TM, and has the homogeneity property $[C, S] = S$. A spray is called *affine* if it is of class C^2 (and hence smooth) on TM.

Any semispray S on TM induces an Ehresmann connection $\mathcal{H}_S$ on TM such that for all $X \in \mathfrak{X}(M)$ we have

$$\mathcal{H}_S\big(\hat{X}\big) = \frac{1}{2}(X^c + [X^v, S]).$$

The torsion of $\mathcal{H}_S$ vanishes. Conversely, *if an Ehresmann connection $\mathcal{H}$ has vanishing torsion, then there exists a semispray S on TM such that $\mathcal{H}_S = \mathcal{H}$*, i.e., '$\mathcal{H}$ is generated by a semispray'. These important results (at least in an intrinsic formulation) are due to M. Crampin and J. Grifone (independently). For details we refer to [35].

Parametric Lagrangians and Finsler manifolds

First we recall that a function $f : TM \to \mathbb{R}$ is called *positive-homogeneous of degree r* ($r \in \mathbb{R}$), briefly *r^+-homogeneous* if for each $\lambda \in \mathbb{R}_+^*$ and $v \in TM$ we have $f(\lambda v) = \lambda^r f(v)$. If f is smooth on $\overset{\circ}{T}M$, then $Cf = rf$, or, equivalently, $\nabla_\delta^v f = rf$ (Euler's relation). Conversely, this property implies that f is r^+-homogeneous on $\overset{\circ}{T}M$.

Definition 3.4 By a *parametric Lagrangian* we mean a 1^+-homogeneous function $F : TM \to \mathbb{R}$ which is smooth on $\overset{\circ}{T}M$. Then $Q := \frac{1}{2}F^2$ is called the *quadratic Lagrangian* or *energy function* associated to F. The symmetric type (0,2) tensor

$$g_F := \nabla^v\nabla^v Q = \frac{1}{2}\nabla^v\nabla^v F^2$$

along $\overset{\circ}{\tau}$ is the *metric tensor* determined by F. If, in addition,

$$F(v) > 0 \quad \text{whenever } v \in \overset{\circ}{T}M,$$

then F is called *positive definite*.

Lemma 3.5 *Let $F : TM \to \mathbb{R}$ be a parametric Lagrangian. Then*

(i) $\nabla^v\nabla^v F\left(\delta, \hat{X}\right) = 0$ *for every vector field X on M, therefore the vertical Hessian of F is degenerate.*

(ii) The quadratic Lagrangian $Q = \frac{1}{2}F^2$ is 2^+-homogeneous; C^1 on TM, smooth on $\overset{\circ}{T}M$.

(iii) The metric tensor $g_F = \nabla^v\nabla^v Q$ is 0^+-homogeneous in the sense that $\nabla_\delta^v g_F = 0$.

(iv) g and Q are related by $g_F(\delta, \delta) = 2Q$.

(v) If $\tilde{\vartheta}_F\big(\tilde{X}\big) := g\left(\tilde{X}, \delta\right)$, then $\tilde{\vartheta}_F$ is a one-form along $\overset{\circ}{\tau}$, and $\vartheta_F := \tilde{\vartheta}_F \circ \mathbf{j}$ is a one-form on $\overset{\circ}{T}M$. We have $\tilde{\vartheta}_F = \nabla^v Q = F\nabla^v F$.

Except the second statement of (ii), each claim can be verified by immediate calculations. For example, taking into account that $\nabla_{\tilde{X}}^v \hat{Y} = 0$ for all $\tilde{X} \in \mathfrak{X}(\tau)$, $Y \in \mathfrak{X}(M)$,

$$\begin{aligned}\nabla^v\nabla^v F\left(\delta, \hat{X}\right) &= \nabla_\delta^v(\nabla^v F)\big(\hat{X}\big) = C\left(\nabla_{\hat{X}}^v F\right) - \nabla^v F\left(\nabla_\delta^v \hat{X}\right) = C(X^v F)\\ &= [C, X^v]F + X^v(CF) = -X^v F + X^v F = 0,\end{aligned}$$

whence (i). As for the (not difficult) proof of the fact that Q is C^1 on TM, see [35, p. 1378, Observation].

Remark 3.6 Both $\tilde{\vartheta}_F$ and ϑ_F are called the *Lagrange one-form* associated to F; $\omega_F := d\vartheta_F$ is the *Lagrange two-form* (d is the classical exterior derivative on TM). ω_F and the metric tensor g_F are related by

$$\omega_F(J\xi, \eta) = g_F(\mathbf{j}\xi, \mathbf{j}\eta); \quad \xi, \eta \in \mathfrak{X}\left(\overset{\circ}{T}M\right)$$

(to check this it is enough to evaluate both sides on a pair (X^c, Y^c) with $X, Y \in \mathfrak{X}(M)$). We conclude that *the Lagrange two-form and the metric tensor associated to a parametric Lagrangian are non-degenerate at the same time*. (Non-degeneracy is meant pointwise. This implies a corresponding property at the level of vector fields, but not vice versa.)

Definition 3.7 By a *Finsler function* we mean a positive definite parametric Lagrangian whose associated metric tensor is non-degenerate (and hence is a pseudo-Riemannian metric on $\overset{\circ}{\tau}{}^*TM$). A manifold is said to be a *Finsler manifold* if its tangent bundle is endowed with a Finsler function.

Proposition 3.8 *If (M, F) is a Finsler manifold, then the metric tensor g_F is positive definite, i.e., g_F is a Riemannian metric in $\overset{\circ}{\tau}{}^*TM$.*

Proof. The problem can be reduced to prove the following: if a smooth function $Q : \mathbb{R}^n \setminus \{0\} \to [0, \infty[$ is 2^+-homogeneous, and the second derivative $Q''(p) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a non-degenerate symmetric bilinear form at each point $p \in \mathbb{R}^n \setminus \{0\}$, then $Q''(p)$ is positive definite. A further reduction is provided by the fact that the index of $Q''(p)$ does not depend on the point of $p \in \mathbb{R}^n \setminus \{0\}$ (this can be seen, e.g., by an immediate continuity argument). Thus it is enough to show that $Q''(p)$ is positive definite at a suitably chosen point p.

The continuity of Q and the compactness of the Euclidean unit sphere $\partial B_1(0)$ implies the existence of a point $e \in \partial B_1(0)$ such that $Q(e) \leqq Q(a)$ for all $a \in \partial B_1(0)$. Let H be the orthogonal complement of the linear span of e. If $v \in H$, then $Q'(e)(v) = 0$, and $Q''(e)(v, v) \geqq 0$. On the other hand, applying (v) of 3.5, it follows that

$$Q''(e)(e, v) = Q'(e)(v) = 0, \quad v \in H.$$

Now let $u \in \mathbb{R}^n \setminus \{0\}$ be any vector. It can uniquely be decomposed in the form $u = \alpha e + v$; $\alpha \in \mathbb{R}$, $v \in H$. Since by Euler's relation $Q''(e)(e, e) = 2Q(e)$, we get

$$Q''(e)(u, u) = 2\alpha^2 Q(e) + Q''(e)(v, v) \geqq 0.$$

This inequality is in fact strict, because $Q''(e)$ is non-degenerate. □

Note The above proof is due to P. Varjú, a student of University of Szeged (Hungary). Another argument can be found in [23].

If (M, F) is a Finsler manifold, then by 3.6 the Lagrange two-form ω_F is non-degenerate, so there exists a unique semispray S on TM such that $i_S\omega_F = -dQ$. Due to the 2^+-homogeneity of Q, S is in fact a spray, called the *canonical spray* of the Finsler manifold.

Lemma 3.9 (fundamental lemma of Finsler geometry) *Let (M,F) be a Finsler manifold. There exists a unique Ehresmann connection $\mathcal{H}$ on TM such that $\mathcal{H}$ is homogeneous, the torsion of $\mathcal{H}$ vanishes, and $dF \circ \mathcal{H} = 0$.*

Proof. We are going to show only the existence statement. As for the uniqueness, which needs more preparation and takes about one page, we refer to [35, p. 1384].

Let S be the canonical spray of (M,F), and consider the Ehresmann connection $\mathcal{H}_S$ induced by S according to the Crampin – Grifone construction. Then, as we have already pointed out, the torsion of $\mathcal{H}_S$ vanishes. Since

$$\mathcal{H}_S \text{ is homogeneous} \overset{\text{def.}}{\iff} \mathbf{t} = 0 \iff \mathbf{i} \circ \mathbf{t} = 0 \iff \left[X^h, C\right] = 0,\ X \in \mathfrak{X}(M),$$

we calculate: $\left[X^h, C\right] = \left[\frac{1}{2}(X^c + [X^v, S]), C\right] = \frac{1}{2}[[X^v, S], C] = -\frac{1}{2}([[S,C],X^v] + [[C,X^v],S]) = \frac{1}{2}([S,X^v] + [X^v,S]) = 0$.

Now we show that $dF \circ \mathcal{H}_S = 0$. Since $dQ = F dF$, and F vanishes nowhere on $\overset{\circ}{T}M$, our claim is equivalent to

$$dQ \circ \mathcal{H}_S = 0 \iff X^h Q = 0 \text{ for all } X \in \mathfrak{X}(M).$$

By the definition of S, $0 = i_S \omega_F + dQ$. We evaluate both sides on a horizontal lift X^h:

$$\begin{aligned} 0 &= \omega_F\left(S, X^h\right) + X^h Q = d\vartheta_F\left(S, X^h\right) + X^h Q = S\tilde{\vartheta}_F\left(\mathbf{j}X^h\right) - X^h\tilde{\vartheta}_F(\mathbf{j}S) \\ &\quad - \tilde{\vartheta}_F\left(\mathbf{j}\left[S, X^h\right]\right) + X^h Q \overset{3.5(\text{v})}{=} S\left(\nabla^v Q\left(\mathbf{j}X^h\right)\right) - X^h(\nabla^v Q(\delta)) \\ &\quad - \nabla^v Q\left(\mathbf{j}\left[S, X^h\right]\right) + X^h Q = S(X^v Q) - 2X^h Q - J\left[S, X^h\right] Q + X^h Q \\ &= [S, X^v]Q - 2X^h Q - X^c Q + X^h Q + X^h Q, \end{aligned}$$

taking into account that $SQ = dQ(S) = -i_S\omega_F(S) = -\omega_F(S,S) = 0$, and (see [35, 3.3, Cor. 3]) $J\left[S, X^h\right] = X^c - X^h$. Finally, by the definition of $\mathcal{H}_S$ again, $[S, X^v]Q = X^c Q - 2X^h Q$, therefore $-2X^h Q = 0$.

Thus we have proved that the canonical spray of a Finsler manifold induces an Ehresmann connection with the desired properties. □

Note In a coordinate-free form, the fundamental lemma of Finsler geometry was first stated and proved by J. Grifone [12]. We call the Ehresmann connection so described the *Barthel connection* of the Finsler manifold. (Other terms, e.g., 'nonlinear Cartan connection' are also used.) Now we briefly discuss the relation between the fundamental lemma of Finsler geometry and Riemannian geometry.

Consider the type $(0,3)$ tensor $\mathcal{C}_\flat := \nabla^v g_F = \nabla^v\nabla^v\nabla^v Q$ along $\overset{\circ}{\tau}$ and the type $(1,2)$ tensor $\mathcal{C}$ defined by the musical duality given by

$$g_F\left(\mathcal{C}\left(\tilde{X}, \tilde{Y}\right), \tilde{Z}\right) = \mathcal{C}_\flat\left(\tilde{X}, \tilde{Y}, \tilde{Z}\right); \quad \tilde{X}, \tilde{Y}, \tilde{Z} \in \mathfrak{X}(\overset{\circ}{\tau}).$$

$\mathcal{C}$, as well as $\mathcal{C}_\flat$, is called Cartan tensor of the Finsler manifold (M,F). It is easy to see that $\mathcal{C}_\flat$ vanishes if and only if g_F is the lift of a Riemannian metric $g : \mathfrak{X}(M) \times \mathfrak{X}(M) \to C^\infty(M)$ on M in the sense that $g_F = g \circ \tau$. (Then $g_F\left(\hat{X}, \hat{Y}\right) = g(X,Y) \circ \tau$ for all $X, Y \in \mathfrak{X}(M)$.) In this particular case the canonical spray of (M,F) becomes an affine spray, and, as we have already remarked, the Barthel connection induces a covariant derivative D on M such that $(D_X Y)^v = \left[X^h, Y^v\right]$ for all vector fields X, Y on M. It can be shown by an immediate calculation that D is just the Levi-Civita derivative on (M, g). Thus, roughly speaking, *if a Finsler manifold reduces to a Riemannian manifold, then its Barthel connection reduces to the Levi-Civita derivative on the base manifold*.

Covariant derivatives on a Finsler manifold

For the sake of convenient exposition, first we introduce some technical terms. Let a covariant derivative operator D on $\overset{\circ}{\tau}{}^* TM$ and an Ehresmann connection $\mathcal{H}$ on TM be given. We say that D is *strongly associated* to $\mathcal{H}$ if $D\delta = \mathcal{V}$ (= the vertical map belonging to $\mathcal{H}$). The *v-part* D^v and the *h-part* D^h of D are given by $D^v_{\tilde{X}}(\cdot) := D_{\mathrm{i}\tilde{X}}(\cdot)$ and $D^h_{\tilde{X}}(\cdot) := D_{\mathcal{H}\tilde{X}}(\cdot)$, respectively. If g is a (pseudo) Riemannian metric on $\overset{\circ}{\tau}{}^* TM$, then D is called *v-metrical*, *h-metrical* or a *metric derivative*, if $D^v g = 0$, $D^h g = 0$ and $Dg = 0$, respectively. By the *h-Cartan tensor* of a Finsler manifold (M, F) we mean the type $(0, 3)$ tensor $\mathcal{C}^h_\flat := \nabla^h g_F$, or the $(1, 2)$ tensor $\mathcal{C}^h$ given by

$$g_F\left(\mathcal{C}^h\left(\tilde{X}, \tilde{Y}\right), \tilde{Z}\right) = \mathcal{C}^h_\flat\left(\tilde{X}, \tilde{Y}, \tilde{Z}\right); \quad \tilde{X}, \tilde{Y}, \tilde{Z} \in \mathfrak{X}(\overset{\circ}{\tau})$$

along $\overset{\circ}{\tau}$, where ∇ is the Berwald derivative induced by the Barthel connection. (∇ will be mentioned as the *Finslerian Berwald derivative* on (M, F) in the following.)

From a 'modern' point of view, the covariant derivative operators introduced in Finsler geometry by L. Berwald, É. Cartan, S. S. Chern, H. Rund and later by M. Hashiguchi using classical tensor calculus, can be interpreted as covariant derivatives on $\overset{\circ}{\tau}{}^* TM$, specified by some nice properties (compatibility or 'semi-compatibility' with the metric tensor and the Barthel connection, vanishing of some torsion tensors). The covariant derivative constructed by É. Cartan in 1934 is an exact analogue of the Levi-Civita derivative on a Riemannian manifold, but it lives on the pull-back bundle $\overset{\circ}{\tau}{}^* TM$. To be more precise, *Cartan's derivative* $D : \mathfrak{X}\left(\overset{\circ}{T}M\right) \times \mathfrak{X}(\overset{\circ}{\tau}) \to \mathfrak{X}(\overset{\circ}{\tau})$ *is the only covariant derivative on a Finsler manifold* (M, F) *which is metric* $(Dg_F = 0)$, *and whose vertical and horizontal torsion vanish* (the latter is taken with respect to the Barthel connection $\mathcal{H}$). Then D is strongly associated to $\mathcal{H}$, and it is related to the Finslerian Berwald derivative by

$$D^v = \nabla^v + \frac{1}{2}\mathcal{C}, \quad D^h = \nabla^h + \frac{1}{2}\mathcal{C}^h.$$

For a proof of this, as well as the next result, we refer to [35].

In 1943 S. S. Chern, using Cartan's calculus of differential forms; later, but independently, in 1951, H. Rund by means of tensor calculus, constructed a covariant derivative on a Finsler manifold which may be described in our language as follows:

Lemma 3.10 *Let* (M, F) *be a Finsler manifold, and let* $\mathcal{H}$ *denote its Barthel connection. There exists a unique covariant derivative* D *on* $\overset{\circ}{\tau}{}^* TM$ *such that* $D^v = \nabla^v$, $D^h g = 0$ *(i.e.,* D *is h-metrical with respect to* $\mathcal{H}$*), and the (*$\mathcal{H}$*-)horizontal torsion of* D *vanishes. Then* D *is strongly associated to* $\mathcal{H}$*, and it is related to the Finslerian Berwald derivative by*

$$D^v = \nabla^v, \quad D^h = \nabla^h + \frac{1}{2}\mathcal{C}^h.$$

□

It will be useful to summarize the classical covariant derivatives used in Finsler geometry in a tabular form.

$D = (D^v, D^h)$	$D^v g$	$D^h g$	$T^v(D)$	$T^h(D)$	$D\delta$
Berwald (∇^v, ∇^h)	$\mathcal{C}_\flat$	$\mathcal{C}^h_\flat$	0	$\mathbf{T} = 0$	$\mathbf{t} \circ \mathbf{j} + \mathcal{V} = \mathcal{V}$
Cartan	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathcal{V}$
Chern – Rund $D^v := \nabla^v$	$\mathcal{C}_\flat$	$\mathbf{0}$	0	$\mathbf{0}$	$\mathcal{V}$
Hashiguchi $D^h := \nabla^h$	$\mathbf{0}$	$\mathcal{C}^h_\flat$	$\mathbf{0}$	$\mathbf{T} = 0$	$\mathcal{V}$

(Data printed in bold are prescribed.)

It is a historical curiosity that Chern's covariant derivative and Rund's covariant derivative were identified only in 1996 [1]. Another construction of the Chern – Rund derivative can be found in a quite recent paper of H.-B. Rademacher [32]: it is given locally as a *covariant derivative on the base manifold* M, 'parametrized' by a nowhere vanishing vector field on an open subset of M, satisfying some Koszul-type axioms. Now we identify Rademacher's constructions with ours presented in 3.10.

Theorem 3.11 *Let (M, F) be a Finsler manifold. Suppose $\mathcal{U} \subset M$ is an open set, and U is a nowhere vanishing vector field on $\mathcal{U}$. If g_U is defined by*

$$g_U(X,Y) := g_F\left(\hat{X}, \hat{Y}\right) \circ U; \quad X, Y \in \mathfrak{X}(\mathcal{U}),$$

then g_U is a Riemannian metric on $\mathcal{U}$, and there exists a unique covariant derivative

$$D^U : \mathfrak{X}(\mathcal{U}) \times \mathfrak{X}(\mathcal{U}) \to \mathfrak{X}(\mathcal{U}), \quad (X,Y) \mapsto D^U_X Y$$

such that D^U is torsion-free and almost metric in the sense that

$$X g_U(Y,Z) = g_U\left(D^U_X Y, Z\right) + g_U\left(Y, D^U_X Z\right) + \mathcal{C}_\flat\left(\widehat{D^U_X Y}, \hat{Y}, \hat{Z}\right) \circ U$$

for any vector fields X, Y, Z on $\mathcal{U}$. D^U is related to the Chern – Rund derivative D by

$$D^U_X Y = \left(D_{X^c} \hat{Y}\right) \circ U; \quad X, Y \in \mathfrak{X}(\mathcal{U}).$$

Proof. To show the *existence*, we define the mapping D^U by the prescription

$$\left(D^U_X Y\right)(p) := D_{X^c} \hat{Y}(U(p)); \quad X, Y \in \mathfrak{X}(\mathcal{U}),\ p \in \mathcal{U};$$

where D is the Chern – Rund derivative on (M, F). It is then obvious that D^U is additive in both of its variables. To check the $C^\infty(\mathcal{U})$-linearity in X and the Leibniz rule, we choose a smooth function f on $\mathcal{U}$, and using the properties of the Chern – Rund derivative, calculate:

$$D^U_{fX} Y := \left(D_{(fX)^c} \hat{Y}\right) \circ U = \left(f^c D_{X^v} \hat{Y} + f^v D_{X^c} \hat{Y}\right) \circ U$$

$$= (f^v \circ U)\left(D_{X^c}\hat{Y} \circ U\right) = fD^U_X Y \quad \left(f^v := f \circ \overset{\circ}{\tau}\right);$$

$$D^U_X fY := \left(D_{X^c} f^v \hat{Y}\right) \circ U = \left((X^c f^v)\hat{Y} + f^v D_{X^c}\hat{Y}\right) \circ U = (Xf)Y + fD^U_X Y.$$

Taking into account that the horizontal torsion of D vanishes, and hence $0 = D_{X^h}\hat{Y} - D_{Y^h}\hat{X} - \mathbf{j}\left[X^h, Y^h\right] = D_{X^h}\hat{Y} - D_{Y^h}\hat{X} - \widehat{[X,Y]}$, it follows that

$$D^U_X Y - D^U_Y X = \left(D_{X^c}\hat{Y} - D_{Y^c}\hat{X}\right) \circ U$$
$$= \left(D_{X^h}\hat{Y} - D_{Y^h}\hat{X}\right) \circ U = \widehat{[X,Y]} \circ U = [X,Y],$$

i.e., D^U is torsion-free. It remains to show that D^U is almost metric. By Lemma 3.1 and using that D is h-metrical, we get

$$Xg_U(Y,Z) = X\left(g_F\left(\hat{Y},\hat{Z}\right) \circ U\right) = \left((X^c + [X,U]^v)g_F\left(\hat{Y},\hat{Z}\right)\right) \circ U$$
$$= \left(\left(X^h + \mathbf{v}X^c + [X,U]^v\right) g_F\left(\hat{Y},\hat{Z}\right)\right) \circ U = g_F\left(D_{X^h}\hat{Y},\hat{Z}\right) \circ U$$
$$+ g_F\left(\hat{Y}, D_{X^h}\hat{Z}\right) \circ U + \mathcal{C}_\flat\left(\mathcal{V}X^c + \widehat{[X,U]}, \hat{Y}, \hat{Z}\right) \circ U.$$

Finally, we identify the term $\mathcal{V}X^c + \widehat{[X,U]}$. Observe that at each point $p \in \mathcal{U}$, $U^h(U(p)) = \mathcal{H} \circ \hat{U}(U(p)) = \mathcal{H}(U(p), U(p)) = \mathcal{H} \circ \delta(U(p)) = S(U(p))$, where S is the canonical spray of (M,F). Using this, the torsion-freeness of the Barthel connection and the relation $\mathcal{C}^h(\cdot,\delta) = 0$ (as for the latter, see [35, 3.11]), we have

$$(\mathbf{v}X^c + [X,U]^v) \circ U = \left(\mathbf{v}X^c + \left[X^h, U^v\right] - \left[U^h, X^v\right]\right) \circ U$$
$$= \left(\mathbf{v}X^c + \mathbf{i}\nabla_{X^h}\hat{U} - \mathbf{i}\nabla_{U^h}\hat{X}\right) \circ U = \left(\mathbf{v}X^c + \mathbf{i}\nabla_{X^h}\hat{U} - \mathbf{i}\nabla_S\hat{X}\right) \circ U$$
$$= \left(\mathbf{v}(X^c - [S, X^v]) + \mathbf{i}\nabla_{X^h}\hat{U} + \frac{1}{2}\mathbf{i}\mathcal{C}^h\left(\hat{X},\delta\right)\right) \circ U = \mathbf{i}\left(D_{X^c}\hat{Y}\right) \circ U$$
$$= \left(D^U_X Y\right)^v.$$

This completes the proof of the existence statement. As for uniqueness, it can be shown (see the cited paper of Rademacher) that if D^U is an almost-metric, torsion-free covariant derivative operator on $\mathcal{U}$, then it obeys a Koszul-type formula, and hence it is uniquely determined. □

Bibliography

[1] M. Anastasiei: A historical remark on the connections of Chern and Rund *In: D. Bao, S.-S. Chern and Z. Shen (eds): Finsler Geometry* (American Mathematical Society, Providence, 1996) 171–176

[2] A. Bastiani: Applications différentiables et variétés différentiables de dimension infinie *J. Anal. Math.* **13** (1964) 1–114

[3] W. Bertram: Differential Geometry, Lie Groups and Symmetric Spaces over General Base Fields and Rings *Memoirs of the AMS* (2006) (to appear); `arXiv:math.DG/0502168`

[4] W. Bertram, H. Glöckner and K.-H. Neeb: Differential calculus over general base fields and rings *Expos. Math.* **22** (2004) 213–282; `arXiv:math.GM/0303300`

[5] Glen E. Bredon: *Topology and Geometry* (Graduate Texts in Mathematics 139, Springer Science & Business Media Inc., New York, 1993)

[6] J. Dieudonné: *Foundations of modern analysis* (Academic Press, New York, 1969)

[7] J. Dugundji: *Topology* (Allyn and Bacon, Boston, 1966)

[8] H. Glöckner and K.-H. Neeb: *Infinite-Dimensional Lie Groups, Vol. I: Basic Theory and Main Examples* (book, in preparation)

[9] H. Glöckner: Infinite-dimensional Lie groups without completeness restrictions *In: A. Strasburger, W. Wojtynski, J. Hilgert and K.-H. Neeb (eds): Geometry and Analysis on finite and infinite-dimensional Lie groups* (Banach Center Publications, 2002) 43–59

[10] H. Glöckner: Implicit functions from Topological Vector Spaces to Banach Spaces *Israel J. Math.* **155** (2006) 205–252

[11] H. Glöckner: Bundles of locally convex spaces, group actions, and hypocontinuous bilinear mappings *manuscript* (submitted) 1–22

[12] J. Grifone: Structure presque tangente et connexions I. *Ann. Inst. Fourier (Grenoble)* **22** (1972) 287–334

[13] R. Hamilton: The inverse function theorem of Nash and Moser *Bull. Amer. Math. Soc.* **7** (1982) 65–222

[14] R. S. Hamilton: Deformation of complex structures on manifolds with boundary *J. Diff. Geom.* **12** (1977) 1–45

[15] S. Helgason: *Groups and Geometric Analysis* (Academic Press, Orlando, 1984)

[16] S. Hiltunen: Implicit functions from locally convex spaces to Banach spaces *Studia Math.* **134** (1999) 235–250

[17] W. Klingenberg: *Riemannian Geometry* (Walter de Gruyter, Berlin–New York, 1982)

[18] A. Kriegl and P. W. Michor: *The Convenient Setting of Global Analysis* (Math. Surveys and Monographs 53, AMS, Providence, 1997)

[19] M. Kuranishi: Deformations of isolated singularities and $\bar{\partial}_b$ *Preprint* Columbia Univ. (1973)

[20] S. Lang: *Fundamentals of Differential Geometry* (Graduate Texts in Mathematics 191, Springer-Verlag, New York, 1999)

[21] J. M. Lee: *Introduction to Smooth Manifolds* (Graduate Texts in Mathematics 218, Springer-Verlag, New York, 2003)

[22] J. Leslie: On the group of real analytic diffeomorphisms of a compact real analytic manifold *Trans. Amer. Math. Soc.* **274** (1982) 651–669

[23] R. L. Lovas: A note on Finsler – Minkowski norms *Houston Journal of Mathematics* (to appear)

[24] T. W. Ma: Inverse mapping theorems on coordinate spaces *Bull. London Math. Soc.* **33** (2001) 473–482

[25] A. D. Michal: Differential calculus in linear topological spaces *Proc. Nat. Acad. Sci.* **24** (1938) 340–342

[26] J. Milnor: On infinite-dimensional Lie groups *Preprint* Institute of Advanced Study, Princeton (1982)

[27] J. Milnor: Remarks on infinite-dimensional Lie groups *In: B. De Witt, R. Stora (eds): Relativité, groupes et topologie II Les Houches* (Éc. d'Été Phys. Théor., session 40, 1983, 1984) 1007–1057

[28] J. Moser: A rapidly convergent iteration method and non-linear differential equations *Ann. Scuola Norm. Sup. Pisa (3)* **20** (1966) 499–535

[29] R. Narasimhan: *Analysis on real and complex manifolds* (Masson & Cie, Paris, 1968)

[30] J. Nash: The embedding problem for Riemannian manifolds *Ann. of Math. (2)* **63** (1956) 20–63

[31] K.-H. Neeb: Nancy Lectures on Infinite-Dimensional Lie Groups *Preprint* (2002)

[32] H.-B. Rademacher: Nonreversible Finsler Metrics of Positive Flag Curvature *In: Riemann – Finsler Geometry* (MSRI Publications 50, 2004) 261-302

[33] W. Rudin: *Functional Analysis* (McGraw-Hill, New York, 1973)

[34] F. Sergeraert: Une généralisation du théorème des fonctions implicites de Nash *C. R. Acad. Sci. Paris, Ser. A* **270** (1970) 861–863

[35] J. Szilasi: A Setting for Spray and Finsler Geometry *In: P. L. Antonelli (ed): Handbook of Finsler Geometry Vol. 2* (Kluwer Academic Publishers, Dordrecht, 2003) 1183–1426

[36] A. C. Thompson: *Minkowski geometry* (Encyclopedia of mathematics and its applications 63, Cambridge University Press, Cambridge, 1996)

[37] E. Zehnder: Generalized implicit function theorems with applications to small divisor problems I, II *Comm. Pure Appl. Math* **28** (1975) 91–141; **29** (1976) 49–114

József Szilasi

University of Debrecen, Institute of Mathematics,
H–4010 Debrecen, P.O. Box 12, Hungary

E-mail: `szilasi@math.klte.hu`

Rezső L. Lovas

University of Debrecen, Institute of Mathematics,
H–4010 Debrecen, P.O. Box 12, Hungary

E-mail: `lovasr@math.klte.hu`

Demeter Krupka and David Saunders

Variational sequences

R. Vitolo

Contents

Introduction

The modern differential geometric approach to mechanics and field theory inspired many scientists coming from different areas of mathematics and theoretical physics to the development of a differential geometric theory of the calculus of variations (see, for example, [40, 45, 69, 108]). Relevant objects of the calculus of variations (like Lagrangians, components of Euler–Lagrange equations) were interpreted as differential forms on jet spaces[1] of a fibred manifold.

Fibred manifolds where chosen to provide a geometric model of the space of independent and dependent coordinates. This is not the most general model, see subsection 6.2.

Soon it was realized that operations like the one of passing from a Lagrangian to its Euler–Lagrange form were part of a complex, namely, the *variational sequence*. The foundational contributions to variational sequences (and much more) can be found in the papers [5, 22, 29, 30, 42, 55, 70, 89, 102, 106, 109, 110, 115, 116, 118]. More details on the development of the subject can be found in section 7.

The variational sequence is a tool that allows to fit several important problems of the calculus of variations all at once. Let us describe two of the most important among such problems.

- Given a set of Euler–Lagrange equations, the vanishing of Helmholtz conditions is a necessary and sufficient condition for the existence of a local Lagrangian for the

[1] The reader is invited to see the paper [100] of this Handbook about jet spaces.

given equations. (see, *e.g.*, [24, 105] for more details on Helmholtz conditions). What about the domain of definition of the Lagrangian? Does there exist a global Lagrangian? This problem is said to be the *inverse problem* of the calculus of variations, despite the fact that it is not the only inverse problem that can be considered in this framework.

- It is important to be able to determine all variationally trivial Lagrangians, depending on derivatives of a prescribed order, defined on a given fibred manifold. Those are Lagrangians whose Euler–Lagrange equation identically vanish. For example, in liquid crystals theories [36] minima of the action functional can be computed by adding to the physical Lagrangian a trivial Lagrangian. Such trivial Lagrangians are known to be locally of the type of a 'total divergence' of an $n-1$-form. But what about their dependence on highest order derivatives? Moreover, another inverse problem arises: are they global total divergences or not?

Let us see what are the answers of variational sequence theories to the above problems in an intuitive way. Let us denote by C the space of currents[2], by L the space of Lagrangian forms, by E the space of Euler–Lagrange-type forms and by H the space of Helmholtz forms. Then, the variational sequence is a sequence of modules (or of sheaves, depending on the approaches) of the type

$$\cdots \longrightarrow C \xrightarrow{d_H} L \xrightarrow{\mathcal{E}} E \xrightarrow{\mathcal{H}} H \xrightarrow{\mathcal{D}} \cdots \qquad ,$$

where d_H is the operator of total divergence, $\mathcal{E}$ is the operator that takes a Lagrangian into the corresponding Euler–Lagrange form, $\mathcal{H}$ is the operator which takes an Euler–Lagrange type form into the corresponding Helmholtz form, and $\mathcal{D}$ is a further operator of the complex.

The repeated application of two consecutive operators of the sequence is identically zero: this is why the homological algebra term 'complex' is used for the above sequence. In the theory of variational sequences the following facts are proved about the previous problems.

- Let $\eta \in E$ be a Euler–Lagrange form. Then $\eta = \mathcal{E}(\lambda)$ for a locally defined Lagrangian $\lambda \in L$ if and only if $\mathcal{H}(\eta) = 0$. The space $\ker \mathcal{H} / \operatorname{Im} \mathcal{E}$ is isomorphic to the $n+1$-st de Rham cohomology of the total space of the fibred manifold. This is the solution of the so-called *global inverse problem*.

- The set of variationally trivial Lagrangians is $\ker \mathcal{E}$. The space $\ker \mathcal{E} / \operatorname{Im} d_H$ is isomorphic to the n-th cohomology class of the total space of the fibred manifold. This enables us to compute which variationally trivial Lagrangians are of the global or local form of a total divergence. More information on the structure of such Lagrangians can be found in section 6.1.

Now, let us describe the structure of the paper.

[2]Here 'currents' are $n-1$-forms, hence they can be integrated on $n-1$-dimensional submanifolds. This includes conserved quantities (or conserved currents). The term 'currents' from classical calculus of variations admits a modern generalization [44] which is not dealt with hereby.

In section 1 some basic facts on jet spaces are recalled. The interested reader may consult [1, 18, 79, 82, 91, 98] and the paper [100] in this Handbook for detailed introductions to jets.

Next section is devoted to contact forms, which are forms whose pull-back by any section of the fibred manifold identically vanishes. The horizontalization is introduced in order to be able to single out the part of a form whose pull-back by any section does not identically vanish. In order to overcome some technical difficulties, infinite order jets are introduced. In practice, this trick amount at dealing with forms which are defined on an arbitrary, but finite, order jet space.

Section 3 contains a description of one of the approaches to variational sequences on fibred manifolds: the variational bicomplex. This approach has been developed mostly in [102, 109, 110, 111]. Local exactness and global cohomological properties of the variational bicomplex are discussed.

In section 4 another important approach to variational sequences is presented: the $\mathcal{C}$-spectral sequence approach by [115, 116, 118]. Contact forms provide a differential filtration of the space of forms on jets. This filtration induces a spectral sequence, the $\mathcal{C}$-spectral sequence, in a standard way. A part of the variational bicomplex and the variational sequence is recovered as some of the differential groups in the $\mathcal{C}$-spectral sequence. The $\mathcal{C}$-spectral sequence also yields a variational sequence on manifolds without fibrings (see subsection 6.2) and on differential equations (see subsection 6.3). In particular, in the latter case, the $\mathcal{C}$-spectral sequence yields differential and topological invariants of the equation, among which there are conservation laws (this particular aspect received foundational contributions also by [22, 23, 106]).

The above approaches were formulated on infinite order jets. In [5, 34, 70] a finite order variational sequence on jets of fibred manifolds is proposed. The approach in [70] is described in section 5, together with comparisons with the above infinite order approaches.

Unfortunately, space constraints do not allow to write a complete text on variational sequences. For this reason, while foundations of the theory are exposed in the above sections in the most possible detailed way (but without detailed proofs), in section 6 there is a collection of references to many interesting theoretical and applied topics like the equivariant inverse problem, symmetries of variational forms, variational sequences on jets of submanifolds, etc.. The reader who is interested in more detailed foundational expositions of the subject could consult the following books.

[4] This book is unpublished, yet it is a good source of examples, calculations and results which never appeared elsewhere.

[14] The book is devoted to the inverse problem in mechanics (one independent variable).

[23] The book covers some geometric aspects of the calculus of variations which are quite close to those of this paper, but in the framework of exterior differential systems.

[18] It is a book on the geometry of differential equations, with one chapter devoted to the variational sequence on jets of fibrings and on differential equations, with a focus on symmetries and conservation laws.

[32] *Idem.* There is one section about the variational bicomplex.

[57] *Idem.* The formalism is quite close to that of [4].

[66] *Idem.* The formalism is the same as in [18] but with a lot more of theoretical material, like the k-lines theorem.

[91] *Idem.* There is also a section on variational multivectors, which are dual objects of variational forms (see subsection 6.5).

[98] The book deals with jets of fibrings and has a final chapter on the variational sequence on infinite order jets.

[121] It is a book on the geometry of differential equations and the $\mathcal{C}$-spectral sequence (see section 4), with a mostly theoretical exposition.

[127] *Idem*, but there are some examples and applications.

We also stress that two very good web sites for this topic are the web site at Utah State University of Logan `http://www.math.usu.edu/~fg_mp` (which seems to be no longer actively maintained) and the 'diffiety' web site `http://diffiety.ac.ru`.

The paper ends with some notes on the development of the subject and a relatively complete bibliography. Despite the fact that we did extensive bibliographical researches the subject is quite vast and it is possible that some issues have been forgotten or not properly mentioned. For this reason, we excuse ourselves in advance with the scientists whose contribution was hereby overlooked or misunderstood.

As a last comment, we observe that we had to make a synthesis from a lot of sources. For this reason we adopted notation that did not come from a single source, but has the advantage of being able to express all approaches at once.

Acknowledgements. It is a pleasure to acknowledge D. Krupka, G. Manno, J. Pohjanpelto and A. M. Verbovetsky for many stimulating comments on the subject in general and on my manuscript in particular. They helped me to improve the text in an essential way.

Many thanks are also due to I. Kolář, I. S. Krasil′shchik, M. Modugno, P. Olver, M. Palese, G. Saccomandi, D. Saunders, C. Tejero Prieto, W. M. Tulczyjew, A. M. Vinogradov for several interesting discussions.

I also would like to thank the staff of the library of the Department of Mathematics (University of Lecce), namely Antonella Toni (director), Salvatore De Giuseppe and Francesco Lucarella, for their kind and professional assistance during my bibliographical researches.

This research has been supported by PRIN 2006/2007 "Leggi di conservazione e termodinamica in meccanica dei continui e in teorie di campo", by the section GNSAGA of the Istituto Nazionale di Alta Matematica `http://www.altamatematica.it`, and by the Department of Mathematics of the University of Lecce.

1 Preliminaries

Manifolds and maps between manifolds are C^∞. All morphisms of fibred manifolds (and hence bundles) will be morphisms over the identity of the base manifold, unless otherwise specified. In particular, when speaking of 'forms' we will always mean 'C^∞ differential forms'.

Some parts of this paper deals with sheaves. A concise but fairly complete exposition of sheaf theory can be found in [126]; it covers all the needs of our exposition.

Now, we recall some basic facts on jet spaces. Our framework is a fibred manifold

$$\pi : E \to M,$$

with $\dim M = n$, $\dim E = n+m$, $n, m \geq 1$. We have the vector subbundle $VE \stackrel{\text{def}}{=} \ker T\pi$ of TE, which is made by vectors which are tangent to the fibres of E.

For $1 \leq r$, we are concerned with the r-th jet space $J^r\pi$; we also set $J^0\pi \equiv E$. For $0 \leq s < r$ we recall the natural fibrings

$$\pi_{r,s} : J^r\pi \to J^s\pi, \qquad \pi_r : J^r\pi \to M,$$

and the affine bundle $\pi_{r,r-1} : J^r\pi \to J^{r-1}\pi$ associated with the vector bundle $\odot^r T^*M \otimes_{J^{r-1}\pi} VE \to J^{r-1}\pi$.

Charts on E adapted to the fibring are denoted by (x^λ, u^i). Greek indices λ, μ,... run from 1 to n and label base coordinates, Latin indices i, j,... run from 1 to m and label fibre coordinates, unless otherwise specified. We denote by $(\partial/\partial x^\lambda, \partial/\partial u^i)$ and (dx^λ, du^i), respectively, the local bases of vector fields and 1-forms on E induced by an adapted chart.

Multiindices are needed in order to label derivative coordinates on jet spaces. It is possible to use general multiindices or symmetrized multiindices in order to label derivatives. There are advantages and disadvantages of both approaches; to the purposes of this paper we prefer to use the symmetrized multiindices because of the one-to-one correspondence between them and coordinates on jets. In particular, following the approach of [98] we denote multi-indices by boldface Greek letters such as $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n) \in \mathbb{N}^n$. We also set $|\boldsymbol{\sigma}| \stackrel{\text{def}}{=} \sum_i \sigma_i$ and $\boldsymbol{\sigma}! \stackrel{\text{def}}{=} \sigma_1! \cdots \sigma_n!$. Multiindices can be summed in an obvious way; the sum of a multiindex with a Greek index $\boldsymbol{\sigma} + \lambda$ will denote the sum of $\boldsymbol{\sigma}$ and the multiindex $(0, \ldots, 1, 0, \ldots, 0)$, where 1 is at the λ-th entry.

The charts induced on $J^r\pi$ are denoted by $(x^\lambda, u^i_{\boldsymbol{\sigma}})$, where $0 \leq |\boldsymbol{\sigma}| \leq r$ and $u^i_0 \stackrel{\text{def}}{=} u^i$. The local vector fields and forms of $J^r\pi$ induced by the fibre coordinates are denoted by $(\partial/\partial u^i_{\boldsymbol{\sigma}})$ and $(du^i_{\boldsymbol{\sigma}})$, $0 \leq |\boldsymbol{\sigma}| \leq r$, $1 \leq i \leq m$, respectively.

An r-th order (ordinary or partial) *differential equation* is, by definition, a submanifold $Y \subset J^r\pi$.

We denote by $j_r s \colon M \to J^r\pi$ the jet prolongation of a section $s \colon M \to E$ and by $J^r f \colon J^r\pi \to J^r\pi$ the jet prolongation of a bundle automorphism $f \colon E \to E$ over the identity. Any vector field $X \colon E \to TE$ which projects onto a vector field $X \colon M \to TM$ can be prolonged to a vector field $X^r \colon J^r\pi \to TJ^r\pi$ by prolonging its flow; its coordinate expression is well-known (see, *e.g.*, [18, 91, 98]).

The fundamental geometric structure on jets is the *contact distribution*, or *Cartan distribution*, $C^r \subset TJ^r\pi$. It is the distribution on $J^r\pi$ generated by all vectors which are tangent to the image $j_r s(M) \subset J^r\pi$ of a prolonged section $j_r s$. It is locally generated by the vector fields

$$D_\lambda = \frac{\partial}{\partial x^\lambda} + u^i_{\boldsymbol{\sigma}\lambda} \frac{\partial}{\partial u^i_{\boldsymbol{\sigma}}}, \qquad \frac{\partial}{\partial u^i_{\boldsymbol{\tau}}}, \tag{1.1}$$

with $0 \leq |\boldsymbol{\sigma}| \leq r-1$, $|\boldsymbol{\tau}| = r$.

Remark 1.1 The contact distribution on finite order jets is not involutive. Indeed, despite the fact that $[D_\lambda, D_\mu] = 0$, if $\boldsymbol{\tau} = \boldsymbol{\sigma} + \lambda$ then $[D_\lambda, \partial/\partial u^i_{\boldsymbol{\tau}}] = -\partial/\partial u^i_{\boldsymbol{\sigma}}$. Moreover, the contact distribution on finite order jets does not admit a natural direct summand that complement it to $TJ^r\pi$. The above two facts are the main motivation to the passage to infinite order jets in order to formulate the variational sequence.

While the contact distribution has an essential importance in the symmetry analysis of PDE [18, 91], in this context the dual concept of contact differential form will play a central role.

2 Contact forms

2.1 Contact forms

Let us denote by $\mathcal{F}_r = C^\infty(J^r\pi)$ the ring of smooth functions on $J^r\pi$.

We denote by Ω^k_r the $\mathcal{F}_r$-module of k-forms on $J^r\pi$.

We denote by Ω^*_r the exterior algebra of forms on $J^r\pi$.

Definition 2.1 We say that a form $\alpha \in \Omega^k_r$ is a *contact k-form* if

$$(j_r s)^*\alpha = 0$$

for all sections s of π.

We denote by $\mathcal{C}^1\Omega^k_r$ the $\mathcal{F}_r$-module of contact k-forms on $J^r\pi$.

We denote by $\mathcal{C}^1\Omega^*_r$ the exterior algebra of contact forms on $J^r\pi$.

Note that if $k > n$ then every form is contact, *i.e.*, $\mathcal{C}^1\Omega^k_r = \Omega^k_r$.

It is obvious from the commutation of d and pull-back that $d\mathcal{C}^1\Omega^k_r \subset \mathcal{C}^1\Omega^{k+1}_r$. Moreover, it is obvious that $\mathcal{C}^1\Omega^*_r$ is an ideal of Ω^*_r. Hence, the following lemma holds.

Lemma 2.2 *The space $\mathcal{C}^1\Omega^*_r$ is a differential ideal of Ω^*_r.*

Unfortunately, the above ideal does not coincide with the ideal generated by 1-forms which annihilate the contact distribution (for this would contradict the non-integrability). More precisely, the following lemma can be easily proved (see, *e.g.*, [70]).

Lemma 2.3 *The space $\mathcal{C}^1\Omega^1_r$ is locally generated (on $\mathcal{F}_r$) by the 1-forms*

$$\omega^i_{\boldsymbol{\sigma}} \overset{def}{=} du^i_{\boldsymbol{\sigma}} - u^i_{\boldsymbol{\sigma}+\lambda}dx^\lambda, \quad 0 \le |\boldsymbol{\sigma}| \le r-1.$$

The above differential forms generate the annihilator of the contact distribution, which is an ideal of Ω^*_r. However, such an ideal is not differential, hence it does not coincide with $\mathcal{C}^1\Omega^*_r$. To realize it, the following formula can be easily proved

$$d\omega^i_{\boldsymbol{\sigma}} = -\omega^i_{\boldsymbol{\sigma}+\lambda} \wedge dx^\lambda, \tag{2.1}$$

from which it follows that, when $|\boldsymbol{\sigma}| = r-1$, then $d\omega^i_{\boldsymbol{\sigma}}$, which is a contact 2-form, cannot be expressed through the 1-forms of lemma 2.3 because $\omega^i_{\boldsymbol{\sigma}+\lambda}$ contains derivatives of order $r+1$.

The following theorem has been first conjectured in [35] ($\mathcal{C}^1\Omega$-hypothesis), then proved in [70, 71].

Theorem 2.4 *Let $k \ge 2$. The space $\mathcal{C}^1\Omega^k_r$ is locally generated (on $\mathcal{F}_r$) by the forms*

$$\omega^i_{\boldsymbol{\sigma}}, \quad d\omega^i_{\boldsymbol{\tau}}, \qquad 0 \le |\boldsymbol{\sigma}| \le r-1, \quad |\boldsymbol{\tau}| = r-1.$$

We can consider forms which are generated by p-th exterior powers of contact forms. More precisely, we have the following definition.

Definition 2.5 Let $p \geq 1$. We say that a form $\alpha \in \Omega^k_r$ is a *p-contact k-form* if it is generated by p-th exterior powers of contact forms.

We denote by $\mathcal{C}^p\Omega^k_r$ the $\mathcal{F}_r$-module of p-contact k-forms on $J^r\pi$.

We denote by $\mathcal{C}^p\Omega^*_r$ the exterior algebra of p-contact forms on $J^r\pi$.

Finally, we set $\mathcal{C}^0\Omega^*_r \stackrel{\text{def}}{=} \Omega^*_r$.

In other words, $\mathcal{C}^p\Omega^*_r$ is the p-th power of the ideal $\mathcal{C}^1\Omega^*_r$ in Ω^*_r. Of course, a 1-contact form is just a contact form. The following lemma is trivial.

Lemma 2.6 *Let $p \geq 0$. We have the inclusion*

$$\mathcal{C}^{p+1}\Omega^*_r \subset \mathcal{C}^p\Omega^*_r.$$

It follows that the space $\mathcal{C}^{p+1}\Omega^*_r$ is a differential ideal of $\mathcal{C}^p\Omega^*_r$, hence of Ω^*_r.

2.2 Horizontalization

Following the discussion in the Introduction, we would like to introduce a tool to extract from a form $\alpha \in \Omega^k_r$ the non-trivial part (to the purposes of calculus of variations). In other words, we would like to introduce a map whose kernel is precisely the set of contact forms. First of all, we observe that eq. (2.1) and Theorem 2.4 suggest that such a map can be constructed if we allow it to increase the jet order by 1. More precisely, it can be easily proved that the contact 1-forms $\omega^i_{\boldsymbol{\sigma}}$, with $0 \leq |\boldsymbol{\sigma}| \leq r-1$ generate a natural subbundle $C^*_r \subset T^*J^r\pi$ [122]. We have the following lemma (see [82, 98]).

Lemma 2.7 *We have the splitting*

$$J^{r+1}\pi \underset{J^r\pi}{\times} T^*J^r\pi = \left(J^{r+1}\pi \underset{M}{\times} T^*M\right) \underset{J^{r+1}\pi}{\oplus} C^*_{r+1}, \tag{2.2}$$

with projections

$$D^{r+1}\colon J^{r+1}\pi \to T^*M \underset{M}{\otimes} TJ^r\pi, \quad \omega^{r+1}\colon J^{r+1}\pi \to T^*J^r\pi \underset{J^r\pi}{\otimes} VJ^r\pi,$$

with coordinate expression

$$D^{r+1} = dx^\lambda \otimes D_\lambda = dx^\lambda \otimes \left(\frac{\partial}{\partial x^\lambda} + u^i_{\boldsymbol{\sigma}+\lambda}\frac{\partial}{\partial u^i_{\boldsymbol{\sigma}}}\right),$$

$$\omega^{r+1} = \omega^i_{\boldsymbol{\sigma}} \otimes \frac{\partial}{\partial u^i_{\boldsymbol{\sigma}}} = (du^i_{\boldsymbol{\sigma}} - u^i_{\boldsymbol{\sigma}+\lambda}dx^\lambda) \otimes \frac{\partial}{\partial u^i_{\boldsymbol{\sigma}}}.$$

Note that the above construction makes sense through the natural inclusions $VJ^r\pi \subset TJ^r\pi$ and $J^{r+1}\pi \times_M T^*M \subset T^*J^{r+1}\pi$, the latter being provided by $T^*\pi_r$.

From elementary multilinear algebra (see the Appendix) it turns out that we have the splitting

$$J^{r+1}\pi \times_{J^r\pi} \wedge^k T^*J^r\pi = \bigoplus_{p+q=k} \left(J^{r+1}\pi \underset{M}{\times} \wedge^q T^*M\right) \underset{J^{r+1}\pi}{\oplus} \wedge^p C^*_{r+1}.$$

Now, we observe that a form $\alpha \in \Omega^k_r$ fulfills

$$\pi^*_{r+1,r}(\alpha)\colon J^{r+1}\pi \to \wedge^k T^* J^r\pi \subset \wedge^k T^* J^{r+1}\pi,$$

where the inclusion is realized through the map $T^*\pi_{r+1,r}$. Hence, $\pi^*_{r+1,r}(\alpha)$ can be split into $k+1$ factors which, respectively, have 0 contact factors, 1 contact factor, ..., k contact factors. More precisely, let us denote by $\mathcal{H}^q_r$ the set of q-forms of the type

$$\alpha\colon J^r\pi \to \wedge^q T^* M.$$

We have the following proposition (for a proof, see [70, 122, 124]).

Proposition 2.8 *We have the natural decomposition*

$$\Omega^k_r \subset \bigoplus_{p+q=k} \mathcal{C}^p\Omega^p_{r+1} \wedge \mathcal{H}^q_{r+1},$$

with splitting projections

$$\mathrm{pr}^{p,q}\colon \Omega^k_r \to \mathcal{C}^p\Omega^p_{r+1} \wedge \mathcal{H}^q_{r+1}, \quad \mathrm{pr}^{p,q}(\alpha) = \left(\binom{p+q}{q} \odot^p i_{D^{r+1}} \odot \odot^q i_{\omega^{r+1}}\right) \circ \pi^*_{r+1,r},$$

where $i_{D^{r+1}}$, $i_{\omega^{r+1}}$ stand for contractions followed by a wedge product (see [98] and the Appendix).

Note that the above maps $\mathrm{pr}^{p,q}$ are not surjective. See [122] for more details.

Definition 2.9 We say the *horizontalization* to be the map

$$h^{p,q}\colon \mathcal{C}^p\Omega^{p+q}_r \to \mathcal{C}^p\Omega^p_{r+1} \wedge \mathcal{H}^q_{r+1}, \quad \alpha \mapsto \mathrm{pr}^{p,q}(\alpha).$$

Horizontalization is not surjective, unless $n = 1$ [72]. We denote by

$$\overline{\Omega}^{p,q}_r \overset{\text{def}}{=} h^{p,q}(\mathcal{C}^p\Omega^{p+q}_r) \tag{2.3}$$

the image of the horizontalization; we say an element $\bar{\alpha} \in \overline{\Omega}^{0,q}_r$ to be a *horizontal form*.

Probably the first occurrence of horizontalization is in [69]. Of course, horizontalization is just the above projection on forms which have 0 contact factors. Note that, if $q > n$, then horizontalization is the zero map. In coordinates, if $0 < q \leq n$, then

$$\alpha = \alpha^{\boldsymbol{\sigma}_1\cdots\boldsymbol{\sigma}_h}_{i_1\ \cdots i_h\ \lambda_{h+1}\cdots\lambda_q}\, du^{i_1}_{\boldsymbol{\sigma}_1} \wedge \cdots \wedge du^{i_h}_{\boldsymbol{\sigma}_h} \wedge dx^{\lambda_{h+1}} \wedge \cdots \wedge dx^{\lambda_q}$$

and

$$h^{0,q}(\alpha) = u^{i_1}_{\boldsymbol{\sigma}_1+\lambda_1} \cdots u^{i_h}_{\boldsymbol{\sigma}_h+\lambda_h} \alpha^{\boldsymbol{\sigma}_1\cdots\boldsymbol{\sigma}_h}_{i_1\ \cdots i_h\ \lambda_{h+1}\cdots\lambda_q}\, dx^{\lambda_1} \wedge \cdots \wedge dx^{\lambda_q}, \tag{2.4}$$

where $0 \leq h \leq q$ (see [5, 70, 71, 122, 125]). Note that the above form is not the most general polynomial in $(r+1)$-st derivatives, even if $q = 1$. For $q > 1$ the skew-symmetrization in the indexes $\lambda_1, \ldots, \lambda_h$ yields a peculiar structure in the polynomial, in which the sums of all terms of the same degree are said to be *hyperjacobians* [90]. Note that the coordinate expressions of $h^{p,q}$ can be obtained in a similar way.

The technical importance of horizontalization is in the next two results.

Lemma 2.10 *Let $\alpha \in \Omega^q_r$, with $0 \le q \le n$, and $s\colon M \to E$ be a section. Then*

$$(j_r s)^*(\alpha) = (j_{r+1}s)^*(h^{0,q}(\alpha))$$

Proposition 2.11 *Let $p \ge 0$. The kernel of $h^{p,q}$ coincides with $p+1$-contact q-forms,* i.e.,

$$\mathcal{C}^{p+1}\Omega^{p+q}_r = \ker h^{p,q}.$$

For a proof of both results, see, for example, [124, 125].

2.3 Horizontal and vertical differential

The above decomposition also affects the exterior differential. Namely, the pull-back of the differential can be split in two operators, one of which raises the contact degree by one, and the other raises the horizontal degree by one. More precisely, in view of proposition 2.8 and following [98], we introduce the maps

$$i_H\colon \Omega^k_r \to \Omega^k_{r+1}, \quad i_H = i_{D^{r+1}} \circ \pi^*_{r+1,r}, \tag{2.5a}$$

$$i_V\colon \Omega^k_r \to \Omega^k_{r+1}, \quad i_V = i_{\omega^{r+1}} \circ \pi^*_{r+1,r}. \tag{2.5b}$$

The maps i_H and i_V are two derivations along $\pi_{r+1,r}$ of degree 0. Together with the exterior differential d they yield two derivations along $\pi_{r+1,r}$ of degree 1, the *horizontal* and *vertical differential*

$$d_H \overset{\text{def}}{=} i_H \circ d - d \circ i_H : \Omega^k_r \to \Omega^k_{r+1},$$

$$d_V \overset{\text{def}}{=} i_V \circ d - d \circ i_V : \Omega^k_r \to \Omega^k_{r+1},$$

It can be proved (see [98]) that d_H and d_V fulfill the properties

$$d_H^2 = d_V^2 = 0, \qquad d_H \circ d_V + d_V \circ d_H = 0, \tag{2.6a}$$

$$d_H + d_V = (\pi^{r+1}_r)^* \circ d, \tag{2.6b}$$

$$(j_{r+1}s)^* \circ d_V = 0, \qquad d \circ (j_r s)^* = (j_{r+1}s)^* \circ d_H. \tag{2.6c}$$

The action of d_H and d_V on functions $f : J^rY \to \mathbb{R}$ and one–forms on J^rY uniquely characterizes d_H and d_V. We have the coordinate expressions

$$d_H f = D_\lambda f\, dx^\lambda = \left(\frac{\partial f}{\partial x^\lambda} + u^i_{\boldsymbol{\sigma}+\lambda}\frac{\partial f}{\partial u^i_{\boldsymbol{\sigma}}}\right) dx^\lambda, \tag{2.7a}$$

$$d_H dx^\lambda = 0, \qquad d_H du^i_{\boldsymbol{\sigma}} = -du^i_{\boldsymbol{\sigma}+\lambda} \wedge dx^\lambda, \qquad d_H \omega^i_{\boldsymbol{\sigma}} = -\omega^i_{\boldsymbol{\sigma}+\lambda} \wedge dx^\lambda, \tag{2.7b}$$

$$d_V f = \frac{\partial f}{\partial u^i_{\boldsymbol{\sigma}}}\,\omega^i_{\boldsymbol{\sigma}}, \tag{2.7c}$$

$$d_V dx^\lambda = 0\,, \qquad d_V du^i_{\boldsymbol{\sigma}} = d^i_{\boldsymbol{\sigma}+\lambda} \wedge dx^\lambda, \qquad d_V \omega^i_{\boldsymbol{\sigma}} = 0. \tag{2.7d}$$

We note that $d_H du^i_{\boldsymbol{\sigma}} = d_H \omega^i_{\boldsymbol{\sigma}}$.

2.4 Infinite order jets

From subsections 2.1, 2.2, 2.3, it is clear that there are fundamental operations in the geometry of jets which do not preserve the order. For this reason the first formulations of variational sequences were derived in infinite order frameworks (with a partial exception in [5]). At the level of forms, this amounts at defining spaces containing all forms on any arbitrary (but finite) order jet. At the level of vector fields, this is done by considering infinite sequences of tangent vectors which are related by the maps $T\pi_{r,s}$.

In what follows we will use the notions of projective (or inverse) system, projective (or inverse) limit, injective (or direct) system, injective (or direct) limit. Such notions can be found in any book of homological algebra (see, *e.g.*, [96]).

We start with the following definition. Consider the projective system

$$\cdots \xrightarrow{\pi_{r+2,r+1}} J^{r+1}\pi \xrightarrow{\pi_{r+1,r}} J^r\pi \xrightarrow{\pi_{r,r-1}} \cdots \xrightarrow{\pi_{1,0}} E \xrightarrow{\pi} M.$$

Definition 2.12 We define the *infinite order jet space* to be the projective limit

$$J^\infty\pi \overset{\text{def}}{=} \varprojlim J^r\pi.$$

Any element $\theta \in J^\infty\pi$ is a sequence of points $\{\theta_r\}_{r\geq 0}$, $\theta_r \in J^r\pi$, which are related by the projections of the system, $\pi_{r,s}(\theta_r) = \theta_s$, $r \geq s$. Hence, we have obvious projections

$$\pi_{\infty,r}\colon J^\infty\pi \to J^r\pi, \qquad \pi_\infty\colon J^\infty\pi \to M.$$

Any section $s\colon M \to E$ induces an element $j^\infty s(x) \in J^\infty\pi$, for $x \in M$, in an obvious way, and conversely, any element $\theta \in J^\infty\pi$ is of the form $\theta = j^\infty s(x)$, with $x = \pi_\infty(\theta)$, for a well-known result of analysis.

Several results can be proved on the infinite order jet: it has the structure of a bundle on E whose fibres are $\mathbb{R}^\infty$, the space of sequences of real numbers; local coordinates on $J^\infty\pi$ are $(x^\lambda, u^i_{\boldsymbol{\sigma}})$, where $0 \leq |\boldsymbol{\sigma}| < +\infty$; it is connected if E is connected, it is Hausdorff and second countable [98]; it is paracompact [102]. Unfortunately, $\mathbb{R}^\infty$ is a Fréchet space which cannot be made into a Banach space [98], hence several important parts of the theory of infinite dimensional Banach manifolds fail to be true. But, to the purposes of building a variational sequence, we need just the ability to deal with functions, tangent vectors and forms which are defined on any finite order jet space. This does not amount at defining all possible functions, tangent vectors, forms on $J^\infty\pi$, but only at considering their inductive or projective counterparts. This is a more or less implicit assumption in the literature; see [18] for an exposition which is close to the present one.

We begin with the projective structure of the tangent space. Namely, we have the following projective system

$$\cdots \xrightarrow{T\pi_{r+2,r+1}} TJ^{r+1}\pi \xrightarrow{T\pi_{r+1,r}} TJ^r\pi \xrightarrow{T\pi_{r,r-1}} \cdots \xrightarrow{T\pi_{1,0}} TE \xrightarrow{T\pi} TM.$$

We define the *tangent space* $TJ^\infty\pi$ to be the projective limit of the above projective system. Hence a tangent vector at $\theta \in J^\infty\pi$ is a sequence of vectors $\{\bar{X}, X_r\}_{k\geq 0}$ tangent to M and to $J^r\pi$ respectively such that $T\pi_r(X_r) = \bar{X}$, $T\pi_{r,s}(X_r) = X_s$ for all $r \geq s \geq 0$. Any tangent vector can be presented in coordinates as the formal sum

$$X = X^\lambda \frac{\partial}{\partial x^\lambda} + X^i_{\boldsymbol{\sigma}} \frac{\partial}{\partial u^i_{\boldsymbol{\sigma}}}, \qquad 0 \leq |\boldsymbol{\sigma}| < +\infty, \tag{2.8}$$

where X^λ, $X^i_{\boldsymbol{\sigma}} \in \mathbb{R}$. Of course we have obvious projections

$$T\pi_{\infty,r} \colon TJ^\infty\pi \to TJ^r\pi, \qquad T\pi_\infty \colon TJ^\infty\pi \to TM,$$

by which it is possible to define pull-back of forms on the infinite order jet. Moreover, we define the *vertical subbundle* $VJ^\infty\pi \subset TJ^\infty\pi$ as the subspace $VJ^\infty\pi \stackrel{\text{def}}{=} \ker T\pi_\infty$. It could also be introduced as a projective limit of finite-order vertical bundles. In coordinates, a vertical tangent vector can be expressed as in (2.8), with the condition $X^\lambda = 0$.

Analogously, we define the *cotangent space* $T^*J^\infty\pi$ to be the injective limit of the injective system $\cdots T^*J^r\pi \to T^*J^{r+1}\pi \cdots$. Hence a cotangent vector at $\theta \in J^\infty\pi$ is an equivalence class of the direct sum $\oplus_{r\in\mathbb{N}} T^*_\theta J^r\pi$ under the following equivalence relation: for all $\alpha, \beta \in \oplus_{r\in\mathbb{N}} T^*_\theta J^r\pi$ we set $\alpha \sim \beta$ if and only if there exist $r, s \in \mathbb{N}$, $r < s$, such that $T^*\pi_{s,r}(\alpha) = \beta$. Moreover, we define the *horizontal subbundle* $\pi^*_\infty(T^*M) \subset T^*J^\infty\pi$.

Any tangent vector can be presented in coordinates as the formal sum

$$X = X^\lambda \frac{\partial}{\partial x^\lambda} + X^i_{\boldsymbol{\sigma}} \frac{\partial}{\partial u^i_{\boldsymbol{\sigma}}}, \qquad 0 \le |\boldsymbol{\sigma}| < +\infty, \tag{2.9}$$

where X^λ, $X^i_{\boldsymbol{\sigma}} \in \mathbb{R}$. Any cotangent vector can be presented in coordinates as the finite sum

$$\alpha = \alpha_\lambda dx^\lambda + \alpha^{\boldsymbol{\sigma}}_i du^i_{\boldsymbol{\sigma}}, \qquad 0 \le |\boldsymbol{\sigma}| \le r, \tag{2.10}$$

for an $r \in \mathbb{N}$.

According with lemma 2.7, we have the following lemma (see [98]).

Lemma 2.13 *We have the splittings*

$$TJ^\infty\pi = C^\infty \underset{J^\infty\pi}{\oplus} VJ^\infty\pi, \tag{2.11}$$

$$T^*J^\infty\pi = \pi^*_\infty(T^*M) \underset{J^\infty\pi}{\oplus} C^*_\infty, \tag{2.12}$$

where

- C^∞ *is the projective limit of the projective system* $\cdots C^{r+1} \to C^r \cdots$, *where the projection is the restriction of* $T\pi_{r+1,r}$ *to* C^{r+1}*;*
- C^*_∞ *is the injective limit of the injective system* $\cdots C^*_r \to C^*_{r+1} \cdots$, *where the injection is the restriction of* $T^*\pi_{r+1,r}$ *to* C^*_r.

The splitting projections are just the direct limit of the maps i_H *and* i_V *of* (2.5)*, that we indicate with the same symbol.*

Now, we could introduce functions and differential forms on $J^\infty\pi$ as functions on $J^\infty\pi$ or sections of exterior powers of $T^*J^\infty\pi$. But we prefer to insist with our 'injective limit' approach because it makes more clear the ideas exposed in the beginning of this section.

The composition with $\pi_{r+1,r}$ provides the injective system of rings $\cdots \subset \mathcal{F}_r \subset \mathcal{F}_{r+1} \subset \cdots$. We can regard the above system also as a *filtered algebra* [18]. Accordingly, pull-back via $\pi_{r+1,r}$ provides several injective (or direct) systems of modules over the above injective system of rings, namely $\cdots \subset \Omega^k_r \subset \Omega^k_{r+1} \subset \cdots$, $\cdots \subset \overline{\Omega}^{0,q}_r \subset \overline{\Omega}^{0,q}_{r+1} \subset \cdots$, $\cdots \subset \mathcal{C}^p\Omega^{p+q}_r \subset \mathcal{C}^p\Omega^{p+q}_{r+1} \subset \cdots$.

Let us introduce the injective (or direct) limits of the above injective systems

$$\mathcal{F} \stackrel{\text{def}}{=} \varinjlim \mathcal{F}_r, \qquad \Omega^k \stackrel{\text{def}}{=} \varinjlim \Omega^k_r, \qquad \overline{\Omega}^{0,q} \stackrel{\text{def}}{=} \varinjlim \overline{\Omega}^{0,q}_r, \qquad \mathcal{C}^p\Omega^{p+q} \stackrel{\text{def}}{=} \varinjlim \mathcal{C}^p\Omega^{p+q}_r.$$

Definition 2.14 We say:

- $f \in \mathcal{F}$ to be a *(smooth) function on* $J^\infty\pi$;
- $\alpha \in \Omega^k$ to be a *(differential) k-form on* $J^\infty\pi$;
- $\bar{\alpha} \in \overline{\Omega}^{0,q}$ to be a *horizontal q-form on* $J^\infty\pi$;
- $\gamma \in \mathcal{C}^p\Omega^k$ to be a *p-contact k-form on* $J^\infty\pi$.

From the definition of direct limit it follows that elements $f \in \mathcal{F}$ are equivalence classes of the direct sum $\oplus_{r\in\mathbb{N}}\mathcal{F}_r$ under the following equivalence relation: for all $g, h \in \oplus_{r\in\mathbb{N}}\mathcal{F}_r$ we set $g \sim h$ if and only if there exist $r, s \in \mathbb{N}$, $r < s$, such that $\pi^*_{s,r}h = g$. Of course, the same holds for the other spaces in the above definition, so that:

- $\mathcal{F}$ is made by all functions on a jet space $J^r\pi$ of arbitrary, but finite, order;
- Ω^k is made by all k-forms on a jet space $J^r\pi$ of arbitrary, but finite, order;
- $\overline{\Omega}^{0,q}$ is made by all horizontal q-forms on a jet space $J^r\pi$ of arbitrary, but finite, order; if $\alpha \in \overline{\Omega}^{0,q}$, then, locally,

  $$\alpha = \alpha_{\lambda_1\cdots\lambda_k}\, dx^{\lambda_1} \wedge \cdots \wedge dx^{\lambda_k}, \qquad \alpha_{\lambda_1\cdots\lambda_k} \in \mathcal{F};$$

  hence, if $\alpha \in \overline{\Omega}^{0,q}$ then $\alpha\colon J^r\pi \to \wedge^k T^*M$, for some $r \in \mathbb{N}$. For this reason, if we consider the inductive system $\cdots \subset \mathcal{H}^q_r \subset \mathcal{H}^q_{r+1} \subset \cdots$ and its injective limit $\mathcal{H}^q$, we have the equality $\mathcal{H}^q = \overline{\Omega}^{0,q}$, which *does not* holt at any finite order level;
- $\mathcal{C}^p\Omega^k$ is made by all p-contact k-forms on a jet space $J^r\pi$ of arbitrary, but finite, order; if $\alpha \in \mathcal{C}^p\Omega^k$, then, locally,

  $$\alpha = \omega^{i_1}_{\boldsymbol{\sigma}_1} \wedge \cdots \wedge \omega^{i_p}_{\boldsymbol{\sigma}_p} \wedge \alpha^{\boldsymbol{\sigma}_1\cdots\boldsymbol{\sigma}_p}_{i_1\cdots i_p}, \qquad \alpha^{\boldsymbol{\sigma}_1\cdots\boldsymbol{\sigma}_p}_{i_1\cdots i_p} \in \Omega^{k-p},$$

  where the multiindexes $\boldsymbol{\sigma}_1,\ldots,\boldsymbol{\sigma}_p$ have arbitrary, but finite, length.

The differential d, the projections $\mathrm{pr}^{p,q}$ (hence also the horizontalization $h^{p,q}$) and the differentials d_H, d_V on finite order jets induce the maps

$$d\colon \Omega^k \to \Omega^{k+1},\ \alpha \mapsto d\alpha, \qquad \mathrm{pr}^{p,q}\colon \Omega^{p+q} \to \mathcal{C}^p\Omega^p \wedge \overline{\Omega}^{0,q},\ \alpha \mapsto \mathrm{pr}^{p,q}(\alpha),$$
$$d_H\colon \Omega^k \to \Omega^{k+1},\ \alpha \mapsto d_H\alpha, \qquad d_V\colon \Omega^k \to \Omega^{k+1},\ \alpha \mapsto d_V\alpha,$$

for each $k \geq 0$, where, being $\alpha \in \Omega^k_r$ for some r, $d\alpha$ coincides with the differential of α on Ω^k_r, and analogously for $\mathrm{pr}^{p,q}$, d_H and d_V.

The proof of the following proposition follows easily from the definitions, the coordinate expressions (2.7), and proposition 2.8.

Proposition 2.15 *We have the natural splitting*

$$\Omega^k = \bigoplus_{p+q=k} \mathcal{C}^p\Omega^p \wedge \overline{\Omega}^{0,q};$$

with splitting projections

$$\mathrm{pr}^{p,q}\colon \Omega^{p+q} \to \mathcal{C}^p\Omega^p \wedge \overline{\Omega}^{0,q}, \quad \mathrm{pr}^{p,q}(\alpha) = \binom{p+q}{q} \odot^p i_V \odot \odot^q i_H.$$

Moreover, we have the following inclusions

$$d_H(\mathcal{C}^p\Omega^p \wedge \overline{\Omega}^{0,q}) \subset \mathcal{C}^p\Omega^p \wedge \overline{\Omega}^{0,q+1}, \qquad d_V(\mathcal{C}^p\Omega^p \wedge \overline{\Omega}^{0,q}) \subset \mathcal{C}^{p+1}\Omega^{p+1} \wedge \overline{\Omega}^{0,q}.$$

Remark 2.16 The above splitting represents one of the major differences between the finite order and the infinite order case. The simple structure of the splitting and the behaviour of d_H and d_V will allow us to give an easy definition of the variational sequence in the infinite order case.

Remark 2.17 The differentials d_H and d_V can also be defined through the above splitting. More precisely, it can be easily proved that

$$d(\mathcal{C}^p\Omega^p \wedge \overline{\Omega}^{0,q}) \subset \mathcal{C}^p\Omega^p \wedge \overline{\Omega}^{0,q+1} \oplus \mathcal{C}^{p+1}\Omega^{p+1} \wedge \overline{\Omega}^{0,q};$$

then d_H is the projection onto the first factor and d_V is the projection onto the second factor of the restriction of d to $\mathcal{C}^p\Omega^p \wedge \overline{\Omega}^{0,q}$ (see [4]).

Finally, a *vector field on* $J^\infty\pi$ is a filtered derivation of $\mathcal{F}$, *i.e.*, an $\mathbb{R}$-linear derivation $X\colon \mathcal{F} \to \mathcal{F}$ such that $X(\mathcal{F}_r) \subset \mathcal{F}_{r+s}$ for all r, and for $l \geq 0$ which depends on X. The number l is said to be the *filtration degree* of the field X. The set of all vector fields is a filtered Lie algebra over $\mathbb{R}$ with respect to commutator $[X, Y]$. Of course, any vector field X on $J^\infty\pi$ can be regarded as a section of $TJ^\infty\pi$ with coordinate expression

$$X = X^\lambda \frac{\partial}{\partial x^\lambda} + X^i_{\boldsymbol{\sigma}} \frac{\partial}{\partial u^i_{\boldsymbol{\sigma}}}, \qquad 0 \leq |\boldsymbol{\sigma}| < +\infty.$$

where $X^\lambda \in \mathcal{F}_s$ and $X^i_{\boldsymbol{\sigma}} \in \mathcal{F}_{r+s}$ [18].

Let X be a vector field on $J^\infty\pi$. Then X can be split according with (2.11) as

$$X = X_H + X_V, \tag{2.13}$$

$$X_H = X^\lambda D_\lambda, \qquad X_V = (X^i_{\boldsymbol{\sigma}} - u^i_{\boldsymbol{\sigma}+\lambda} X^\lambda) \frac{\partial}{\partial u^i_{\boldsymbol{\sigma}}} \quad 0 \leq |\boldsymbol{\sigma}| < +\infty.$$

We observe that any vector field $X\colon E \to TE$ which projects onto a vector field $X\colon M \to TM$ can be prolonged to a vector field X^∞ with filtration degree 0. We have

$$X^\infty = X^\lambda D_\lambda + D_{\boldsymbol{\sigma}}(X^i - u^i_\lambda X^\lambda) \frac{\partial}{\partial u^i_{\boldsymbol{\sigma}}}, \quad 0 \leq |\boldsymbol{\sigma}| < +\infty, \tag{2.14}$$

where $X^\lambda \in C^\infty(M)$. The vector field X^∞_V is said to be the *evolutionary vector field* with *generating function* X (see, *e.g.*, [18, 91]).

We can consider more general evolutionary vector fields. Namely, it can be proved (see, *e.g.*, [18, 91]) that a vector field X on $J^\infty\pi$ is a symmetry of C^∞ if and only if it its vertical part is of the form $X_V = E_\varphi$, where $\varphi\colon J^r\pi \to V\pi$ and

$$E_\varphi\colon J^\infty\pi \to VJ^\infty\pi, \quad E_\varphi = D_{\boldsymbol{\sigma}}\varphi^i\frac{\partial}{\partial u^i_{\boldsymbol{\sigma}}}. \tag{2.15}$$

We say E_φ to be an evolutionary vector field with generating function φ; of course, the filtration degree of E_φ is the order r of the jet space on which φ is defined. It can be proved [18, 91] that evolutionary vector fields are uniquely determined by their generating functions. We denote the $\mathcal{F}_r$-module of generating functions on $J^r\pi$ with $\varkappa_r$. Composing with projections $\pi_{r+1,r}$ yields the chain of inclusions $\cdots \subset \varkappa_r \subset \varkappa_{r+1} \subset \cdots$, hence the direct limit $\varkappa$. This module plays an important role in section 4.

3 Variational bicomplex and variational sequence

Variational sequences has been introduced basically in two ways.

The first way is through the properties of d_H, d_V on infinite order jets [108, 109, 102, 110]. Another way to describe this approach is to consider the splitting (2.11) as a connection on $J^\infty\pi$ which has zero curvature.

The second way is through a spectral sequence [29, 30, 115, 116, 118]; this approach will be described in section 4.

Partial exceptions to this classification are [5], where the approach is (at least partially) on finite order jets, and [17], where the approach is based on the properties of the interior Euler operator (see subsection 3.1). In this section we describe the approach of [108, 109, 102, 110] in a modern language which is close to that of [4, 98].

For all $p \geq 0$ we introduce the following notation:

$$E_0^{0,q} \stackrel{\text{def}}{=} \overline{\Omega}^{0,q}, \qquad E_0^{p,q} \stackrel{\text{def}}{=} \mathcal{C}^p\Omega^p \wedge \overline{\Omega}^{0,q}, \tag{3.1}$$

$$E_1^{p,n} \stackrel{\text{def}}{=} E_0^{p,n}/d_H(E_0^{p,n-1}) = \mathcal{C}^p\Omega^p \wedge \overline{\Omega}^{0,n}/d_h(\mathcal{C}^p\Omega^p \wedge \overline{\Omega}^{0,n-1}). \tag{3.2}$$

The integers p, q are called, respectively, the *contact* and the *horizontal degree*.

We also denote by $\Omega^k(M)$ the space of k-forms on M.

We define the map

$$e_1\colon E_1^{p,n} \to E_1^{p+1,n}, \quad e_1([\alpha]) = [d_V\alpha]. \tag{3.3}$$

The above map is well-defined because $d_V \circ d_H = -d_H \circ d_V$.

In view of the properties (2.6a) of d_H and d_V the following diagram commutes

$$
\begin{array}{ccccccccc}
 & & 0 & & 0 & & 0 & & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
0 & \longrightarrow & E_0^{0,0} & \xrightarrow{d_V} & E_0^{1,0} & \cdots & E_0^{p,0} & \xrightarrow{d_V} & E_0^{p+1,0} \quad \cdots \\
 & & \downarrow{\scriptstyle d_H} & & \downarrow{\scriptstyle -d_H} & & \downarrow{\scriptstyle (-1)^p d_H} & & \downarrow{\scriptstyle (-1)^{p+1} d_H} \\
0 & \longrightarrow & E_0^{0,1} & \xrightarrow{d_V} & E_0^{1,1} & \cdots & E_0^{p,1} & \xrightarrow{d_V} & E_0^{p+1,1} \quad \cdots \\
 & & \vdots & & \vdots & & \vdots & & \vdots \\
0 & \longrightarrow & E_0^{0,n-1} & \xrightarrow{d_V} & E_0^{1,n-1} & \cdots & E_0^{p,n-1} & \xrightarrow{d_V} & E_0^{p+1,n-1} \quad \cdots \\
 & & \downarrow{\scriptstyle d_H} & & \downarrow{\scriptstyle -d_H} & & \downarrow{\scriptstyle (-1)^p d_H} & & \downarrow{\scriptstyle (-1)^{p+1} d_H} \\
0 & \longrightarrow & E_0^{0,n} & \xrightarrow{d_V} & E_0^{1,n} & \cdots & E_0^{p,n} & \xrightarrow{d_V} & E_0^{p+1,n} \quad \cdots \\
 & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
 & & 0 & & 0 & & 0 & & 0
\end{array}
\tag{3.4}
$$

and rows and columns are complexes (in the sense that the kernel of a map contains the image of the previous). According with standard terminology from homological algebra, the above diagram (3.4) is a *double complex*, or a *bicomplex* [20]. The diagram (3.4) can be *augmented* (again, a standard procedure from homological algebra) by the natural inclusion of de Rham complex of M on the left edge and the natural quotient projection on the complex

$$
0 \longrightarrow E_1^{0,n} \xrightarrow{e_1} E_1^{1,n} \cdots\cdots E_1^{p,n} \xrightarrow{e_1} E_1^{p+1,n} \cdots\cdots
$$

on the bottom edge. The resulting bicomplex is

$$
\begin{array}{cccccccccccc}
 & & 0 & & 0 & & & & & & & \\
 & & \downarrow & & \downarrow & & & & & & & \\
 & & \mathbb{R} & & \mathbb{R} & & 0 & & 0 & & 0 & \\
 & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & \\
0 & \longrightarrow & \Omega^0(M) & \xrightarrow{\pi_\infty^*} & E_0^{0,0} & \xrightarrow{d_V} & E_0^{1,0} & \cdots & E_0^{p,0} & \xrightarrow{d_V} & E_0^{p+1,0} & \cdots \\
 & & \downarrow d & & \downarrow d_H & & \downarrow -d_H & & \downarrow (-1)^p d_H & & \downarrow (-1)^{p+1} d_H & \\
0 & \longrightarrow & \Omega^1(M) & \xrightarrow{\pi_\infty^*} & E_0^{0,1} & \xrightarrow{d_V} & E_0^{1,1} & \cdots & E_0^{p,1} & \xrightarrow{d_V} & E_0^{p+1,1} & \cdots \\
 & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & \\
0 & \longrightarrow & \Omega^{n-1}(M) & \xrightarrow{\pi_\infty^*} & E_0^{0,n-1} & \xrightarrow{d_V} & E_0^{1,n-1} & \cdots & E_0^{p,n-1} & \xrightarrow{d_V} & E_0^{p+1,n-1} & \cdots \\
 & & \downarrow d & & \downarrow d_H & & \downarrow -d_H & & \downarrow (-1)^p d_H & & \downarrow (-1)^{p+1} d_H & \\
0 & \longrightarrow & \Omega^n(M) & \xrightarrow{\pi_\infty^*} & E_0^{0,n} & \xrightarrow{d_V} & E_0^{1,n} & \cdots & E_0^{p,n} & \xrightarrow{d_V} & E_0^{p+1,n} & \cdots \\
 & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & \\
 & & 0 & \longrightarrow & E_1^{0,n} & \xrightarrow{e_1} & E_1^{1,n} & \cdots & E_1^{p,n} & \xrightarrow{e_1} & E_1^{p+1,n} & \cdots \\
 & & & & \downarrow & & \downarrow & & \downarrow & & \downarrow & \\
 & & & & 0 & & 0 & & 0 & & 0 &
\end{array}
\tag{3.5}
$$

Definition 3.1 We say the *variational bicomplex* associated with the fibred manifold $\pi\colon E \to M$ to be the bicomplex (3.5).

The variational sequence can be extracted from the variational bicomplex.

Definition 3.2 The following complex

$$
0 \longrightarrow \mathbb{R} \longrightarrow E_0^{0,0} \xrightarrow{d_H} E_0^{0,1} \cdots\cdots E_0^{0,n-1} \xrightarrow{d_H} E_0^{0,n} \xrightarrow{\mathcal{E}} \tag{3.6}
$$

$$
\xrightarrow{\mathcal{E}} E_n^{1,n} \xrightarrow{e_1} E_1^{2,n} \cdots\cdots E_1^{p,n} \xrightarrow{e_1} E_1^{p+1,n} \cdots\cdots
$$

where the map $\mathcal{E}$ is just the composition of the quotient projection $E_0^{0,n} \to E_1^{0,n}$ with the differential $e_1\colon E_1^{0,n} \to E_1^{1,n}$, is said to be the *variational sequence*[3].

The motivation for the above definition is immediate after the analysis of the quotient spaces $E_1^{p,q}$ that we are going to perform in next subsection.

The second column of the variational bicomplex has a special importance and will be studied later on.

---

[3]Some authors use the term *Euler–Lagrange complex* instead, see [4]

Definition 3.3 We say the following sequence

$$0 \longrightarrow \mathbb{R} \longrightarrow E_0^{0,0} \xrightarrow{d_H} E_0^{0,1} \cdots\cdots E_0^{0,n-1} \xrightarrow{d_H} E_0^{0,n} \longrightarrow 0 \tag{3.7}$$

to be the *horizontal de Rham sequence*.

3.1 Representation of the variational sequence by forms

The problem of representing the elements of the quotients $E_1^{p,n}$ for $p > 1$ has been independently solved by many authors [109, 110, 115, 116, 79, 17]. We recognize two different approaches to the problem: with differential forms [109, 110, 79, 17] and with differential operators [115, 116]. In this section we follow the first approach. The interpretation of the variational sequence (3.6) in terms of objects of the calculus of variations will follow at once.

Following [109, 110, 4], let us introduce the map

$$I\colon E_0^{p,n} \to E_0^{p,n}, \quad I(\alpha) = \frac{1}{p}\omega^i \wedge (-1)^{|\boldsymbol{\sigma}|} D_{\boldsymbol{\sigma}} \left(i_{\partial/\partial u^i_{\boldsymbol{\sigma}}} \alpha\right) \tag{3.8}$$

where $D_{\boldsymbol{\sigma}}$ stands for the iterated Lie derivative $(L_{D_1})^{\sigma_1} \cdots (L_{D_n})^{\sigma_n}$.

Definition 3.4 We say the map I to be the *interior Euler operator*[4].

Note that I is denoted by τ in [109, 110] and by D^* in [17]. For a proof of the following theorem, see [4, 67, 110].

Theorem 3.5 *The following properties of I holds*

- *I is a natural map,* i.e., $L_{X^\infty}(I(\alpha)) = I(L_{X^\infty}(\alpha))$*, hence I is a global map;*
- *if* $\alpha \in E_0^{p,n}$ *then there exists a unique form* $\beta \in E_0^{n,p}$*, which is of the type* $\beta = d_H\gamma$ *with* $\gamma \in E_0^{p,n-1}$*, such that*

$$\alpha = I(\alpha) + \beta. \tag{3.9}$$

Remark 3.6 The above form γ is not uniquely defined, in general. For $p = 1$, if the order of α is 1 it is easily proved that γ is uniquely defined; if the order of α is 2 then there exists a unique γ fulfilling a certain intrinsic property; if the order is 3 it is proved in [61] that no natural γ of the above type exists. In [37, 41, 61] it is proved that suitable linear connections on M and on the fibres of $\pi\colon E \to M$ can be used to determine a unique γ. See [2, 4] for the case $p > 1$.

It follows from the above theorem that I is a global map, and if $\gamma \in E_1^{p,n-1}$ then $I(d_H(\gamma)) = 0$, so that $I^2 = I$. For this reason I induces an isomorphism (denoted by the same letter)

$$I\colon E_1^{p,n} \to \mathcal{V}^p, \quad [\alpha] \mapsto I(\alpha),$$

where $\mathcal{V}^p \subset E_0^{p,n}$ is a suitable subspace. The map I also allows us to represent the differentials $\mathcal{E}$, e_1 through forms: $I(\mathcal{E}(\lambda)) = I([d_V\lambda])$, and $I(e_1([\alpha])) = I([d_V\alpha])$.

[4]This name is due to I. Anderson.

Definition 3.7 We say the elements of $\mathcal{V}^p$ to be the p-th degree *variational (or functional, as in [4]) forms*.

Let us see the coordinate expression of I in the most meaningful cases. We set $\nu \overset{\text{def}}{=} dx^1 \wedge \cdots \wedge dx^n$.

Case $p = 1$**:** let $[\alpha] \in E_1^{1,n}$. Then $\alpha = \alpha_i^{\boldsymbol{\sigma}} \omega_{\boldsymbol{\sigma}}^i \wedge \nu$ and

$$I([\alpha]) = (-1)^{|\boldsymbol{\sigma}|} D_{\boldsymbol{\sigma}} \alpha_i^{\boldsymbol{\sigma}}\, \omega^i \wedge \nu.$$

Hence, if $\lambda \in E_0^{0,n}$, then $\lambda = L\nu$, $\mathcal{E}(\lambda) = [\partial L/\partial u_{\boldsymbol{\sigma}}^i \omega_{\boldsymbol{\sigma}}^i \wedge \nu]$ and

$$I(\mathcal{E}(\lambda)) = (-1)^{|\boldsymbol{\sigma}|} D_{\boldsymbol{\sigma}} \frac{\partial L}{\partial u_{\boldsymbol{\sigma}}^i}\, \omega^i \wedge \nu, \tag{3.10}$$

which is just the expression of the *Euler–Lagrange form* corresponding to the *Lagrangian form* λ. It can be proved that the Euler–Lagrange form is the only natural operator in a broad class of differential operators [62, 63]. It is not difficult to prove the following result (see [98, 102]).

Proposition 3.8 *The space $\mathcal{V}^1$ is equal to the injective limit of the system $\cdots \mathcal{V}_r^1 \subset \mathcal{V}_{r+1}^1 \cdots$, where $\mathcal{V}_r^1$ is the space of sections of the bundle*

$$(\pi_{r,1}^* C_1^*) \wedge (\pi_r^* \wedge^n T^*M)\,.$$

Following [102], the elements of $\mathcal{V}^1$ are called *source forms*. A source form $\eta \in \mathcal{V}^1$ has the coordinate expression

$$\eta = \eta_i\, \omega^i \wedge \nu, \qquad \eta_i \in \mathcal{F},\ i = 1, \dots, m.$$

Case $p = 2$**:** let $[\alpha] \in E_1^{1,n}$. Then $\alpha = \alpha_{i\,j}^{\boldsymbol{\sigma}\boldsymbol{\tau}} \omega_{\boldsymbol{\sigma}}^i \wedge \omega_{\boldsymbol{\tau}}^j \wedge \nu$ (with $\alpha_{i\,j}^{\boldsymbol{\sigma}\boldsymbol{\tau}} = -\alpha_{j\,i}^{\boldsymbol{\tau}\boldsymbol{\sigma}}$) and, if α is a form on the r-th order jet, then

$$\begin{aligned} I([\alpha]) &= \frac{1}{2}\omega^i \wedge (-1)^{|\boldsymbol{\sigma}|} D_{\boldsymbol{\sigma}}(\alpha_{i\,j}^{\boldsymbol{\sigma}\boldsymbol{\tau}}\, \omega_{\boldsymbol{\tau}}^j) \wedge \nu \\ &= \frac{1}{2} \sum_{\substack{\boldsymbol{\mu}+\boldsymbol{\tau}=\boldsymbol{\rho} \\ 0 \le |\boldsymbol{\rho}| \le 2r}} (-1)^{|\boldsymbol{\xi}+\boldsymbol{\mu}|} \frac{(\boldsymbol{\xi}+\boldsymbol{\mu})!}{\boldsymbol{\xi}!\boldsymbol{\mu}!} D_{\boldsymbol{\xi}} \alpha_{i\,\,\,\,j}^{\boldsymbol{\xi}+\boldsymbol{\mu}\boldsymbol{\tau}} \omega^i \wedge \omega_{\boldsymbol{\rho}}^j \wedge \nu. \end{aligned} \tag{3.11}$$

If $\eta \in \mathcal{V}^1$, then η represents the element $[\eta] \in E_1^{p,q}$. Hence, if η is a form on the r-th order jet, then

$$\begin{aligned} I(e_1[\eta]) =& I(d_V \eta) \\ =& I\left(\frac{\partial \eta_k}{\partial u_{\boldsymbol{\sigma}}^h} \omega_{\boldsymbol{\sigma}}^h \wedge \omega^k \wedge \nu\right) \\ =& \frac{1}{2}\omega^i \wedge \left(\frac{\partial \eta_j}{\partial u^i} - \frac{\partial \eta_i}{\partial u^j}\right) \omega^j \wedge \nu \\ &+ \frac{1}{2}\omega^i \wedge \sum_{\substack{0 \le \boldsymbol{\rho} \le 2r \\ |\boldsymbol{\mu}+\boldsymbol{\rho}| \ge 1}} (-1)^{|\boldsymbol{\mu}+\boldsymbol{\rho}|} \frac{(\boldsymbol{\mu}+\boldsymbol{\rho})!}{\boldsymbol{\mu}!\boldsymbol{\rho}!} D_{\boldsymbol{\mu}} \frac{\partial \eta_j}{\partial u_{\boldsymbol{\mu}+\boldsymbol{\rho}}^i} \omega_{\boldsymbol{\rho}}^j \wedge \nu \end{aligned} \tag{3.12}$$

> The above form is the well-known *Helmholtz form* corresponding to the source form η. The above coordinate expression dates back to [17], even if the local expression of the Helmholtz conditions $I(d_V\eta) = 0$ of local variationality of η were known much before, even in the general case of arbitrary values of r and n.
>
> The Helmholtz conditions may be also expressed by the Helmholtz tensor [64]. It has the same components of the Helmholtz form without skew-symmetrization with respect to the pair of indexes (i, j). It has been proved that the Helmholtz tensor is the only natural operator in a broad class [64, 87]. Note that the Helmholtz form is also connected with the second variation of functionals [39].

Note that if $p \geq 2$ then the spaces $\mathcal{V}^p$ cannot be characterized as spaces of sections of a vector bundle, like $\mathcal{V}^1$. This can be realized by the fact that $\mathcal{V}^p$ fail to be modules over $\mathcal{F}$. We will see in section 4 how to characterize the elements of $\mathcal{V}^p$.

The forms in $\mathcal{V}^p$ may also be interpreted as functionals. The case $p = 1$ was clear in the papers [40, 45, 69]; the case $p > 1$ was dealt with first in [17] (see also [4]). This provides a relationship between 'standard' calculus of variations and the theory of variational sequences.

Definition 3.9 Let $\alpha \in E_0^{p,n}$. Then we define the family of functionals $\mathcal{A}(\alpha)$

$$\mathcal{A}(\alpha)(X_1, \ldots, X_p)(s)_U \overset{\text{def}}{=} z \int_U (j^\infty s)^* \alpha(X_1^\infty, \ldots, X_p^\infty),$$

depending on an oriented open set with compact closure and oriented regular boundary $U \subset M$, on p vertical vector fields $\{X_i \colon E \to VE\}_{1 \leq i \leq p}$ which vanish on $\pi^{-1}(\partial U)$, and on a section s of π.

The vector fields $X_1, \ldots, X_p$ are called *variation fields*. We denote by

$$\mathcal{F}^p \overset{\text{def}}{=} \{\mathcal{A}(\alpha) \mid \alpha \in E_0^{p,n}\}$$

the space of functionals.

The following proposition is proved in [4, 17].

Proposition 3.10 *Let α, $\alpha' \in E_0^{p,n}$. Then $\mathcal{A}(\alpha) = \mathcal{A}(\alpha')$ if and only if $[\alpha] = [\alpha'] \in E_1^{p,n}$, or, equivalently, if and only if $\alpha' = \alpha + d_H\beta$, with $\beta \in E_0^{p,n-1}$. Hence $\mathcal{V}^p \simeq \mathcal{F}^p$.*

Of course, in the case $p = 1$ we recover the standard integral of a source form evaluated on a variation field.

3.2 Local properties of the variational bicomplex

In this section we show that the variational bicomplex is locally exact. More precisely, recall that an exact sequence is a complex where the kernel of each map is equal to the image of the previous one. Then, we prove that for all $p \in E$ there exists an open neighbourhood $U \subset E$ of p such that the variational sequence on the fibred manifold $\pi|_U \colon U \to \pi(U)$ is an exact sequence.

We begin by proving an exactness result for the rows of the variational bicomplex.

Theorem 3.11 *Let $q \geq 0$. Then for all $p \in E$ there exists an open neighbourhood $U \subset E$ of p such that the rows*

$$0 \longrightarrow \Omega^q(M) \xrightarrow{\pi_\infty^*} E_0^{0,q} \xrightarrow{d_V} E_0^{1,q} \cdots\cdots E_0^{p,q} \xrightarrow{d_V} E_0^{p+1,q} \cdots\cdots$$

of the variational bicomplex associated with the fibred manifold $\pi|_U : U \to \pi(U)$ *are exact.*

The above theorem was proved in [109] for the case of jets of n-velocities (see, *e.g.*, [63] for a definition) and [110] for the case of jets of fibred manifolds (see also [4, 98] for a detailed exposition). The proof is just a 'vertical' version of the Poincaré lemma. In [115, 116] an alternative proof was proposed, see section 4.

A more complex homotopy operator is constructed in order to prove next theorem. Several proofs of the following result has been provided: [109, 110] (see also [4, 98] for a detailed exposition) and [102] with two different homotopy operators, [115, 116] with Spencer sequences, [106] with Koszul complexes (indeed, in [115, 116] the authors proved the *global exactness*, see also [118, 18] and section 4 for a detailed exposition).

Theorem 3.12 *Let* $p \geq 1$. *Then for all* $\mathrm{p} \in E$ *there exists an open neighbourhood* $U \subset E$ *of* p *such that the columns*

$$0 \longrightarrow E_0^{p,0} \xrightarrow{(-1)^p d_H} E_0^{p,1} \cdots\cdots E_0^{p,n-1} \xrightarrow{(-1)^p d_H} E_0^{p,n} \longrightarrow 0$$

of the variational bicomplex associated with the fibred manifold $\pi|_U : U \to \pi(U)$ *are exact for horizontal degrees* $0 \leq q \leq n-1$.

Note that in the case $p = 1$ the global exactness was also established in [61] (see also references therein) by direct computation of the potential of d_H-closed forms using an auxiliary symmetric linear connection on M.

As relatively trivial consequences of the above theorems we have the following corollary, obtained via 'diagram chasing' techniques [109, 110, 102] (see also [4, 98] for a detailed exposition), or via spectral sequences [115, 116] (see also [118, 18] and section 4 for a detailed exposition).

Corollary 3.13 *For all* $\mathrm{p} \in E$ *there exists an open neighbourhood* $U \subset E$ *of* p *such that in the variational bicomplex associated with the fibred manifold* $\pi|_U : U \to \pi(U)$ *the following complexes are exact:*

- *the horizontal de Rham sequence* (3.7)*;*
- *the bottom row of the variational sequence*

$$0 \longrightarrow E_1^{0,n} \xrightarrow{e_1} E_1^{1,n} \cdots\cdots E_1^{p,n} \xrightarrow{e_1} E_1^{p+1,n} \cdots\cdots \quad . \tag{3.13}$$

It follows that the variational sequence associated with the fibred manifold $\pi|_U : U \to \pi(U)$ *is exact.*

Note that also in [79] a variational sequence is constructed and the local exactness at the vertices of degree n and $n+1$ is proved.

3.3 Global properties of the variational bicomplex

In this section we collect results about the cohomology of rows and columns of the variational bicomplex on the given (but arbitrary) fibred manifold $\pi : E \to M$. We recall that the cohomology of a complex is the sequence of the quotients of the kernel of a map with the image of the previous one. The cohomology of the columns is the most studied because it allows to compute the cohomology of the variational sequence.

Theorem 3.14 *Let $p \geq 1$. Then the columns*

$$0 \longrightarrow E_0^{p,0} \xrightarrow{(-1)^p d_H} E_0^{p,1} \cdots\cdots E_0^{p,n-1} \xrightarrow{(-1)^p d_H} E_0^{p,1} \longrightarrow E_1^{p,n} \longrightarrow 0$$

*are exact (*i.e.*, the above sequence have zero cohomology) for horizontal degrees* $0 \leq q \leq n-1$.

The above theorem has been proved in several ways. The first proofs appeared in [115, 116] (using Spencer sequences; see also the longer paper [118]), in [102] (using a sheaf-theoretical approach) in [106] (using an isomorphism with the polynomial Koszul complex; see also the more modern texts [18, 66, 114, 121]) and in [4] (using local exactness and a Mayer–Vietoris argument). Note that the approach of [102] implies passing from modules of sections $E_0^{p,q}$ to the corresponding sheaves of germs of local sections. Those sheaves consists of sections which are defined on finite order jet spaces only locally (see [43, 123]). The following corollary holds.

Corollary 3.15 *The cohomology of the variational sequence is (not naturally) isomorphic to the de Rham cohomology of E.*

Note that the above corollary implies that the cohomology of the horizontal de Rham sequence is isomorphic to the de Rham cohomology of E for horizontal degrees $0 \leq q \leq n-1$. Such a cohomology is also called *characteristic cohomology* in the framework of exterior differential systems [22, 23].

The above corollary can be proved using spectral sequences [115, 116] (but see the more modern texts [18, 66, 114, 121]), sheaf-theoretical arguments [102] or just basic diagram chasing [4]. Note that all proofs show first that the cohomology of the variational sequence is isomorphic to the cohomology of the complex (Ω^*, d), which is, by definition, the de Rham cohomology $H^*(J^\infty\pi)$ of $J^\infty\pi$. Then it is quickly seen that $H^*(J^\infty\pi)$ is isomorphic to $H^*(E)$, just by the fact that, in this case, the cohomology functor commutes with direct limits.

The cohomology of the rows of the variational bicomplex is much less studied. We have the following results [4].

Theorem 3.16 *The cohomology of the rows*

$$0 \longrightarrow \Omega^q(M) \xrightarrow{\pi_\infty^*} E_0^{0,q} \xrightarrow{d_V} E_0^{1,q} \cdots\cdots E_0^{p,q} \xrightarrow{d_V} E_0^{p+1,q} \cdots\cdots$$

vanish for vertical degrees $p > m$.

Some restrictions on the topology of E have to be asked in order to compute the cohomology of vertical rows.

Theorem 3.17 *Let π be a bundle with typical fibre F. Suppose that F admit a finite covering $\{U_i\}_{0\leq i\leq k}$ such that each U_i and each non-empty intersection $U_{i_1} \cap \cdots \cap U_{i_l}$ is diffeomorphic to $\mathbb{R}^n$ for any l (finite* good cover, *see [20]). Suppose that for each p there are a finite number $\{\beta_i\}_{1\leq i\leq d}$ of p-forms on E whose restriction to the fibres of E is a basis for the cohomology of the fibres. Then*

$$H(E_0^{p,q}, d_V) \simeq H^p(F) \otimes H^q(M).$$

More precisely, the forms $\alpha_i \overset{def}{=} \pi^*_{\infty,0}(\beta_i) \in E_0^{p,0}$ *are* d_V*-closed, and if* $\alpha \in E_0^{p,q}$ *is* d_V*-closed, then there are forms* $\{\xi_i\}_{1\leq i\leq d}$ *in* $\Omega^q(M)$ *and a form* $\eta \in E_0^{p-1,q}$ *such that*

$$\alpha = \sum_{i=1}^{d} \xi_i \wedge \alpha_i + d_V\eta.$$

The forms $\{\xi_i\}_{1\leq i\leq d}$ *are unique in the sense that* α *is* d_V*-exact if and only if* $\{\xi_i\}_{1\leq i\leq d}$ *vanish.*

The above theorem is clearly inspired by the Leray–Hirsch theorem [20], but its hypotheses are weaker because the forms β_i are not assumed to be closed on E.

4 $\mathcal{C}$-spectral sequence and variational sequence

In this section we derive the variational sequence as a by-product of a spectral sequence, the $\mathcal{C}$-spectral sequence. To the author's knowledge the first formulations of the $\mathcal{C}$-spectral sequence (on infinite order jets) were done in [29] and [116], independently. But the computation of all terms of the $\mathcal{C}$-spectral sequence was done in [116] (using results from [115]), in the more general setting of differential equations (see also the longer paper [118]). Note that the variational sequence was already formulated in [115], without using the $\mathcal{C}$-spectral sequence. See the notes in section 7 for more details.

The $\mathcal{C}$-spectral sequence allows us not only to recover the variational bicomplex as was formulated in the previous section, but also to formulate a variational sequence on infinite order jets of submanifolds and on infinite prolongations of (ordinary or partial) differential equations (which, we recall, are submanifolds of a jet space of a certain finite order). In this section we will recall the main results on the $\mathcal{C}$-spectral sequence on the fibred manifold π. We will follow the most recent presentation of the subject [18, 66].

We will use the language of differential operators, as in [115, 116]. There are a number of reasons for doing that. First of all this language is used by a part of the scientists that work in this field. Then, it yields the same construction as the interior Euler operator using the adjoint of a differential operator. Moreover, differential operators and the operations on them constitute a calculus which is complementary to that of differential forms and is of independent interest with respect to variational sequences. An important domain of application of this calculus is, for example, the Hamiltonian formalism for evolution equations [56, 66].

4.1 The $\mathcal{C}$-spectral sequence and its 0-th term

Here we introduce the $\mathcal{C}$-spectral sequence and compute its first term.

We begin by recalling the basic facts on spectral sequences, but we suggest the interested reader to consult a book on algebraic topology (like [20, 81]; see also [66]).

We recall that a *filtered module* is a module P endowed with a chain[5] of submodules

$$P \overset{\text{def}}{=} F^0P \supset F^1P \supset F^2P \supset \cdots \supset F^pP \supset \cdots$$

[5] We will only use decreasing filtrations.

A filtered module yields the *associated graded module* $S_0^*(P)$, where

$$S_0^p(P) \stackrel{\text{def}}{=} F^pP/F^{p+1}P.$$

A *(graded) filtered complex* is a graded filtered module P endowed with a differential d of degree 1 which preserves the filtration, *i.e.*, $d(F^pP) \subset F^pP$. With every filtered complex it is associated a filtration of its cohomology $H^*(P)$ as follows:

$$H^*(P) = F^0H^*(P) \supset F^1H^*(P) \supset F^2H^*(P) \supset \cdots \supset F^pH^*(P) \supset \cdots \qquad (4.1)$$

where $F^pH^*(P)$ is the image of the cohomological map $H^*(F^pP) \to H^*(P)$ induced by the inclusion $F^pP \subset P$. In general (4.1) is a filtration without a natural differential.

Any filtered complex yields a *spectral sequence*. A spectral sequence is a sequence of differential Abelian groups (S_n, s_n) where the cohomology of each term is equal to the next term: $H(S_n, s_n) = S_{n+1}$.

A spectral sequence is said to *converge* if there exists $k \in \mathbb{N}$ such that for every $k' \in \mathbb{N}$, $k' > k$, we have $S_k = S_{k'}$. In this case we set $S_\infty \stackrel{\text{def}}{=} S_k$. For spectral sequences associated with filtered complexes the notion of convergence is more specific. Namely, a spectral sequence associated with a filtered complex is said to *converge* if it exists a graded filtered module Q such that $S_\infty = S_0^*(Q)$. It can be proved [20, 81, 66] that if a spectral sequence associated with a filtered complex P lays in the *first quadrant* (*i.e.*, $S_r^{p,q} = 0$ whenever $p < 0$ or $q < 0$), then it converges to the graded module $S_0(H^*(P))$ associated with the filtration (4.1) of $H^*(P)$.

In view of lemma 2.6, the following infinite chain of module inclusions

$$\Omega^* = \mathcal{C}^0\Omega^* \supset \mathcal{C}^1\Omega^* \supset \mathcal{C}^2\Omega^* \supset \cdots \supset \mathcal{C}^p\Omega^* \supset \cdots \qquad (4.2)$$

is a filtered complex.

Definition 4.1 The above filtered complex (4.2) is said to be the $\mathcal{C}$-*filtration*.

The induced spectral sequence is said to be the $\mathcal{C}$-*spectral sequence*.

We recall that, from the definition of spectral sequence associated with a filtered complex, the first term $(S_0^{p,q}, s_0)$ of the $\mathcal{C}$-spectral sequence is just the graded module associated with the $\mathcal{C}$-filtration, *i.e.*,

$$S_0^{p,q} = \mathcal{C}^p\Omega^{p+q}/\mathcal{C}^{p+1}\Omega^{p+q},$$

with differential $s_0 \colon S_0^{p,q} \to S_0^{p,q+1}$, $s_0([\alpha]) = [d\alpha]$. The $\mathcal{C}$-spectral sequence is a first quadrant spectral sequence, hence it converges to the graded group associated with the de Rham cohomology $H^*(J^\infty\pi)$ of the initial complex Ω^* of the $\mathcal{C}$-filtration. We stress that $H^*(J^\infty\pi)$ is filtered according with (4.1).

The proof of the following proposition is quite simple and derives from proposition 2.11 [18, 115, 116, 118, 121].

Proposition 4.2 *The horizontalization* $h^{p,q}$ *yields an isomorphism, denoted by the same symbol,*

$$h^{p,q} \colon S_0^{p,q} \to E_0^{p,q}, \quad [\alpha] \mapsto h^{p,q}(\alpha).$$

Moreover, the above isomorphism yields $s_0 = d_H$. It turns out that the first term of the $\mathcal{C}$-spectral sequence is just the family of complexes $(E_0^{p,*}, d_H)_{0\leq p<+\infty}$, or, equivalently, the family of columns of the diagram (3.4).

Remark 4.3 The reader may wonder about how to recover rows of the variational bicomplex within the $\mathcal{C}$-spectral sequence approach. There is another natural filtration of Ω^*: it is provided by horizontal forms. Namely, one could consider the ideal of forms generated by the codistribution $T^*\pi\colon T^*M \to T^*J^\infty\pi$ and its powers. This filtration is preserved by d and yields another spectral sequence, whose 0-term consists of the rows of the diagram (3.4).

The computations of the remaining terms of the $\mathcal{C}$-spectral sequence will be done in subsection 4.3.

4.2 Forms and differential operators

The computation of the $\mathcal{C}$-spectral sequence has been performed in [115, 116, 118] in the language of differential operators. More precisely, there is an isomorphism between the spaces $E_0^{p,q}$ and suitable spaces of differential operators. As a by-product, we will obtain a description of the spaces $E_1^{p,q}$. The purpose of this subsection is to recall the basic facts about differential operators and to state the above mentioned isomorphism.

We now recall the basic algebraic and geometric setting for differential operators. The interested reader could consult [18, 66, 121] for more details.

Let P, Q be modules over an algebra A over $\mathbb{R}$. We recall ([1]) that a *linear differential operator* of order k is defined to be an $\mathbb{R}$-linear map $\Delta : P \to Q$ such that

$$[\delta_{a_0}, [\ldots, [\delta_{a_k}, \Delta]\ldots]] = 0 \tag{4.3}$$

for all $a_0, \ldots, a_k \in A$. Here, square brackets stand for commutators and δ_{a_i} is the multiplication morphism. Of course, linear differential operators of order zero are morphisms of modules. The A-module of differential operators of order k from P to Q is denoted by $\mathrm{Diff}_k(P, Q)$. The A-module of differential operators of any order from P to Q is denoted by $\mathrm{Diff}(P, Q)$. This definition can be generalized to maps between the product of the A-modules $P_1, \ldots P_l$ and Q which are differential operators of order k in each argument, *i.e.*, *multidifferential operators*. The corresponding space is denoted by $\mathrm{Diff}_k(P_1, \ldots, P_l; Q)$, or, if $P_1 = \cdots = P_l = P$, by $\mathrm{Diff}_{(l)\,k}(P, Q)$. Accordingly, we define $\mathrm{Diff}_{(l)}(P, Q)$.

Let P, Q be modules of sections of a vector bundle over the same basis M, and suppose that $(e_i)_{0\leq i\leq p}$, $(f_j)_{1\leq j\leq q}$ are local bases for their respective sections. Then it can be proved that a differential operator $\Delta \in \mathrm{Diff}_k(P, Q)$ acts in coordinates as expected:

$$\Delta(s) = a_i^{j\boldsymbol{\sigma}} \frac{\partial^{|\boldsymbol{\sigma}|} s^i}{\partial x^{\sigma_1}\cdots\partial x^{\sigma_n}} f_j, \qquad 0 \leq |\boldsymbol{\sigma}| \leq k, \quad \text{for all } s \in P, \tag{4.4}$$

where we used the coordinate expression $s = s^i e_i$. The proof makes use of Taylor expansions of the coefficients s^i and of the property (4.3).

Consider the chain of algebrae $\cdots \subset \mathcal{F}_k \subset \mathcal{F}_{k+1} \subset \cdots$, and two chains of modules of sections of vector bundles $\cdots \subset P_k \subset P_{k+1} \subset \cdots$ and $\cdots \subset Q_k \subset Q_{k+1} \subset \cdots$ over the previous algebrae, with direct limits P and Q. Then a differential operator $\Delta\colon P \to Q$ is an $\mathbb{R}$-linear map such that for all k the restriction $\Delta|_{P_k}$ is a differential operator $\Delta|_{P_k}\colon P_k \to Q_{k+l}$, where l can depend on k.

We will mainly use differential operators whose expressions contain total derivatives instead of standard ones. To do that, we say a $\mathcal{F}$-module P to be *horizontal* if it is the module of sections of $\pi_\infty^* V \to J^\infty\pi$, where $V \to M$ is a vector bundle. Of course, P can be seen as the direct limit of the chain of modules of sections of $\pi_r^* V \to J^r\pi$. Then, we say a differential operator $\Delta\colon P \to Q$ (of order k) between two horizontal modules P and Q to be *$\mathcal{C}$-differential* if it can be restricted to the manifolds of the form $j^\infty s(M)$, where s is a section of π. In other words, Δ is a $\mathcal{C}$-differential operator if the equality $j^\infty s(M)^*(\varphi) = 0$, $\varphi \in P$, implies $j^\infty s(M)^*(\Delta(\varphi)) = 0$ for any section $s\colon M \to E$. In local coordinates, we have $\Delta = a_i^{j\boldsymbol{\sigma}} D_{\boldsymbol{\sigma}}$, where $a_i^{j\boldsymbol{\sigma}} \in \mathcal{F}$.

We denote by $\mathcal{C}\mathrm{Diff}_k(P, Q)$ the $\mathcal{F}$-module of $\mathcal{C}$-differential operators of order k from P to Q. We also introduce the $\mathcal{F}$-module $\mathcal{C}\mathrm{Diff}(P, Q)$ of differential operators from P to Q of any order. We can generalize the definition to multi-$\mathcal{C}$-differential operators. In particular, we will be interested to spaces of antisymmetric multi-$\mathcal{C}$-differential operators, which we denote by $\mathcal{C}\mathrm{Diff}^{\mathrm{alt}}_{(l)\,k}(P, Q)$. Analogously, we introduce $\mathcal{C}\mathrm{Diff}^{\mathrm{alt}}_{(l)}(P, Q)$.

Now, we consider the two horizontal modules $\varkappa$ (2.15) and $E_0^{0,q} = \overline{\Omega}^{0,q}$. For a proof of the following proposition, see [18].

Proposition 4.4 *We have the natural isomorphism*

$$E_0^{p,q} \to \mathcal{C}\mathrm{Diff}^{\mathrm{alt}}_{(p)}(\varkappa, E_0^{0,q}), \qquad \alpha \mapsto \nabla_\alpha \tag{4.5}$$

where $\nabla_\alpha(\varphi_1, \dots, \varphi_p) = E_{\varphi_p} \lrcorner (\dots \lrcorner (E_{\varphi_1} \lrcorner \alpha) \dots)$.

Note that the isomorphism holds because for any vertical tangent vector to $J^r\pi$ there exists an evolutionary field passing through it. In coordinates, if α

$$\alpha = \alpha^{\boldsymbol{\sigma}_1 \cdots \boldsymbol{\sigma}_p}_{i_1 \cdots i_p\ \lambda_1 \cdots \lambda_q}\, \omega^{i_1}_{\boldsymbol{\sigma}_1} \wedge \cdots \wedge \omega^{i_p}_{\boldsymbol{\sigma}_p} \wedge dx^{\lambda_1} \wedge \cdots \wedge dx^{\lambda_q}$$

then

$$\nabla_\alpha(\varphi_1, \dots, \varphi_p) = p!\alpha^{\boldsymbol{\sigma}_1 \cdots \boldsymbol{\sigma}_p}_{i_1 \cdots i_p\ \lambda_1 \cdots \lambda_q}\, D_{\boldsymbol{\sigma}_1}\varphi^{i_1} \cdots D_{\boldsymbol{\sigma}_p}\varphi^{i_p}\, dx^{\lambda_1} \wedge \cdots \wedge dx^{\lambda_q}$$

4.3 The $\mathcal{C}$-spectral sequence and its 1-st and 2-nd terms

The term $S_1^{p,q}$ of the $\mathcal{C}$-spectral sequence is computed in two steps. The following lemma yields the cohomology of the terms with $0 \le q \le n-1$. Let P be a horizontal module. Recall that $E_0^{0,q} = \overline{\Omega}^{0,q}$, the space of horizontal forms, and $E_0^{0,0} = \mathcal{F}$. We introduce the *adjoint module* $P^* \overset{\mathrm{def}}{=} \mathrm{Hom}(P, E_0^{0,n})$. Consider the complex

$$0 \longrightarrow \mathcal{C}\mathrm{Diff}_{(p)}(P, E_0^{0,0}) \xrightarrow{\ w\ } \mathcal{C}\mathrm{Diff}_{(p)}(P, E_0^{0,1}) \cdots\cdots \mathcal{C}\mathrm{Diff}_{(p)}(P, E_0^{0,n}) \longrightarrow 0 \tag{4.6}$$

where the maps w are defined by $w(\nabla) \overset{\mathrm{def}}{=} d_H \circ \nabla$.

Theorem 4.5 *The cohomology of the complex* (4.6) *is zero at* $\mathcal{C}\mathrm{Diff}_{(p)}(P, E_0^{0,q})$ *for* $0 \le q \le n-1$ *and is* $\mathcal{C}\mathrm{Diff}_{(p-1)}(P, P^*)$ *for* $q = n$.

The first version of the above theorem appeared in [116] (corollary 2; see also the longer paper [117]. The proof was published later [118] and used Spencer cohomology. Another proof based on the Koszul complex appeared in [106], but the formulation involved only differential forms in $E_0^{p,q}$. In the language of differential operators (see proposition 4.4), this amounts at considering the subspace $\mathcal{C}\mathrm{Diff}^{\mathrm{alt}}_{(p)}(P, E_0^{0,q}) \subset$

$\mathcal{C}\mathrm{Diff}_{(p)}(P, E_0^{0,q})$, with $P = \varkappa$. The statement of the above theorem 4.5 is taken from [18, p. 190], but the proof is essentially the same as in [106]. Note that there is an obvious inclusion

$$\mathcal{C}\mathrm{Diff}_{(p-1)}(P, P^*) \subset \mathcal{C}\mathrm{Diff}_{(p)}(P, E_0^{0,n}).$$

There is an action of the permutation group S_p of p elements on $\mathcal{C}\mathrm{Diff}_{(p)}(P, E_0^{0,q})$. Namely, if $\tau \in S_p$ and $\nabla \in \mathcal{C}\mathrm{Diff}_{(p)}(P, E_0^{0,q})$ then for all $s_1, \ldots, s_p \in P$ we set $\tau(\nabla)(s_1, \ldots, s_p) \stackrel{\text{def}}{=} \nabla(s_{\tau(1)}, \ldots, s_{\tau(p)})$. This action commutes with w, so that we have the following corollary.

Corollary 4.6 *The skew-symmetric part of the complex* (4.6) *has zero cohomology at* $\mathcal{C}\mathrm{Diff}^{\mathrm{alt}}_{(p)}(P, E_0^{0,q})$ *for* $0 \le q \le n-1$.

It is easy to realize through the isomorphism of proposition 4.4 that, if $P = \varkappa$, then $w = d_H$ up to the isomorphism 4.5. Another set of terms of the $\mathcal{C}$-spectral sequence follows.

Corollary 4.7 *We have:*

- $S_1^{p,q} = 0$ *for* $p > 0$ *and* $0 \le q \le n-1$*;*
- $S_1^{0,q} = H^q(E)$ *for* $0 \le q \le n-1$.

The content of the above theorem is the same of theorem 3.14 and corollary 3.15. The proof of the second statement of above corollary follows from the convergence of the $\mathcal{C}$-spectral sequence to the de Rham cohomology. Indeed it can be quickly realized that the differential s_1 is the zero map on $S_1^{0,q}$ for $0 \le q \le n-1$ and that s_1 is never $S_1^{0,q}$-valued for $0 \le q \le n-1$. It follows that $S_2^{0,q} = S_\infty^{0,q}$ for $0 \le q \le n-1$. Moreover, we note that there is at most one nonzero term among $S_\infty^{p,q}$ with $p + q = k$. Then there exists a $\tilde{p}$ such that $F^{\tilde{p}} H^{\tilde{p}+q}(\Omega_r^*) \ne F^{\tilde{p}+1} H^{\tilde{p}+q}(\Omega_r^*)$. This implies that the filtration of the de Rham cohomology of the initial complex (Ω^*, d) is trivial:

$$H^{\tilde{p}+q}(\Omega_r^*) = F^0 H^{\tilde{p}+q}(\Omega_r^*) = \cdots = F^{\tilde{p}} H^{\tilde{p}+q}(\Omega_r^*) \supset 0 \cdots \supset 0, \tag{4.7}$$

whence $S_\infty^{\tilde{p},q} = H^{\tilde{p}+q}(\Omega_r^*)$.

We now calculate the last set of terms of E_1. The following elementary lemma comes directly from the definition of spectral sequence and the isomorphism (4.5).

Lemma 4.8 *We have*

$$S_1^{p,n} = E_0^{p,n}/d_H(E_0^{p,n-1}) = \mathcal{C}^p\Omega^p \wedge \overline{\Omega}^{0,n}/d_H(\mathcal{C}^p\Omega^p \wedge \overline{\Omega}^{0,n-1}) = E_1^{p,n}.$$

It turns out that

$$E_1^{p,n} \simeq \mathcal{C}\mathrm{Diff}^{\mathrm{alt}}_{(p)}(\varkappa, E_0^{0,n})/d_H(\mathcal{C}\mathrm{Diff}^{\mathrm{alt}}_{(p)}(\varkappa, E_0^{0,n-1})).$$

In view of the discussion preceding corollary 4.6, the term $E_1^{p,n}$ is isomorphic to the subspace $K_p(\varkappa) \subset \mathcal{C}\mathrm{Diff}_{(p-1)}(\varkappa, \varkappa^*)$ of elements which are invariant under the action of the permutation group S_p. To do that we need the notion of adjoint operator. Let P, Q be horizontal modules and $\Delta\colon P \to Q$ a $\mathcal{C}$-differential operator. Then Δ induces a map

$$\Delta'\colon \mathcal{C}\mathrm{Diff}(Q, E_0^{0,q}) \to \mathcal{C}\mathrm{Diff}(P, E_0^{0,q}), \quad \Delta'(\nabla) = \nabla \circ \Delta.$$

The map Δ' is a cochain map for the complex (4.6) (with $p = 1$), in the sense that $\Delta' \circ w = w \circ \Delta'$. Hence Δ' yields a cohomology map which, according to theorem 4.5, is trivial if $0 \leq q \leq n-1$ and is denoted by $\Delta^* \colon Q^* \to P^*$ if $q = n$.

Definition 4.9 The operator Δ^* is said to be the *adjoint operator* to Δ.

In coordinates, following the same notation of (4.4), we have $\Delta = a_i^{j\boldsymbol{\sigma}} D_{\boldsymbol{\sigma}}$. If $\{e^i \otimes \nu\}$, $\{f^j \otimes \nu\}$ are two local bases respectively of P^* and Q^*, and $s^* \in P^*$, $t^* \in Q^*$, we have $s^* = s_i e^i \otimes \nu$, $t^* = t_j f^j \otimes \nu$, and

$$\Delta^*(t^*) = (-1)^{|\boldsymbol{\sigma}|} D_{\boldsymbol{\sigma}}(a_i^{j\boldsymbol{\sigma}} t_j) e^i \otimes \nu. \tag{4.8}$$

In fact, it can be easily proved that the composition $\nabla \circ \Delta$ is locally equal to the above expression up to an operator in $\operatorname{Im} w$. Of course, locally this is just integration by parts. The global meaning of the expression (4.8) appears in the two following statements (for a proof, see [18, 66]). If $\Delta \in \mathcal{C}\mathrm{Diff}_{(p)}(P, Q)$, then for any $p_1, \ldots, p_{p-1}$ we define $\Delta(p_1, \ldots, p_{p-1}) \in \mathcal{C}\mathrm{Diff}(P, Q)$ in the following obvious way:

$$\Delta(p_1, \ldots, p_{p-1})(p_p) = \Delta(p_1, \ldots, p_{p-1}, p_p). \tag{4.9}$$

Next lemma shows how to determine the representative of each n-th cohomology class of the complex (4.6).

Lemma 4.10 *Let P be a horizontal module and $\Delta \in \mathcal{C}\mathrm{Diff}_{(p)}(P, E_0^{0,n})$. Then*

$$\Delta(p_1, \ldots, p_{p-1}) = \Delta(p_1, \ldots, p_{p-1})^*(1) + w(\nabla(p_1, \ldots, p_{p-1})), \tag{4.10}$$

where $\nabla(p_1, \ldots, p_{p-1}) \in \mathcal{C}\mathrm{Diff}(P, E_0^{0,n-1})$. It turns out that $w(\nabla(p_1, \ldots, p_{p-1})) = w(\tilde{\nabla})$, with $\tilde{\nabla} \in \mathcal{C}\mathrm{Diff}_{(p)}(P, E_0^{0,n})$.

The proof is achieved first locally, with a relatively easy computation, then globally by observing that $\Delta(p_1, \ldots, p_{p-1})^*(1)$ is a natural operator and the representative of a cohomology class, hence the difference $\Delta(p_1, \ldots, p_{p-1}) - \Delta(p_1, \ldots, p_{p-1})^*(1)$ must lie in the image of w. See also [18, 66].

The above operator $w(\tilde{\nabla})$ is uniquely determined, but $\tilde{\nabla}$ is not. The problem of determining under which additional requirements $\tilde{\nabla}$ is uniquely determined has been thoroughly analysed in [2, 4, 61] (see remark 3.6). Eq. (4.10) is a consequence of the fact that every object in the n-th cohomology class of the complex (4.6) is globally represented by a single homomorphism in P^*.

Proposition 4.11 (Green's formula) *Let P, Q be horizontal modules and $\Delta \colon P \to Q$ a $\mathcal{C}$-differential operator. Then*

$$q^*(\Delta(p)) - (\Delta^*(q^*))(p) = d_H(\omega_{p,q^*}(\Delta)) \tag{4.11}$$

for all $q^* \in Q^*$, $p \in P$, where $\omega_{p,q^*}(\Delta) \in E_0^{0,n-1}$ and $\omega_{p,q^*}(\Delta)$ is a $\mathcal{C}$-differential operator with respect to p and q^*.

The above formula has been introduced in [115] (but see also [18, 66, 118]); its proof is a simple consequence of lemma 4.10.

Now, it is easy to see that the action of a permutation of the first $p-1$ arguments of $\square \in \mathcal{C}\mathrm{Diff}_{(p-1)}(\varkappa, \varkappa^*)$ commutes with the splitting of lemma 4.10, hence $K_p(\varkappa) \subset$

$\mathcal{C}\mathrm{Diff}^{\mathrm{alt}}_{(p-1)}(\varkappa,\varkappa^*)$. Then, for $\Delta\in\mathcal{C}\mathrm{Diff}_{(p-1)}(\varkappa,\varkappa^*)$ and for any $p_1,\ldots,p_p$ we define $\Delta_j(p_1,\ldots,\widehat{p_j},\ldots,p_{p-1})\in\mathcal{C}\mathrm{Diff}(\varkappa,\varkappa^*)$ in the following obvious way:

$$\Delta_j(p_1,\ldots,\widehat{p_j},\ldots,p_{p-1})(p_j)(p_p)=\Delta(p_1,\ldots,p_{p-1})(p_p). \tag{4.12}$$

Due to Green's formula we have

$$\Delta_j(p_j)(p_p)=\Delta_j^*(p_p)(p_j)+d_H(\omega_{p,q^*}(\Delta)).$$

This implies that $K_p(\varkappa)\subset\mathcal{C}\mathrm{Diff}^{\mathrm{alt}}_{(p-1)}(\varkappa,\varkappa^*)$ is the subset of skew-adjoint operators with respect to the exchange of one of the first $p-1$ arguments with the last one. Hence, we proved the following theorem.

Theorem 4.12 *There is an isomorphism*

$$I\colon E_1^{p,n}\to K_p(\varkappa),\quad [\Delta]\mapsto\Delta^*(1), \tag{4.13}$$

where the adjoint is taken with respect to one of the arguments of Δ.

Let us see the coordinate expression of I in the most meaningful cases. We will represent elements of $E_0^{p,n}$ through the isomorphism 4.5. We set $\nu\stackrel{\mathrm{def}}{=}dx^1\wedge\cdots\wedge dx^n$.

Case $p=1$**:** let $[\alpha]\in E_1^{1,n}$. Then $\nabla_\alpha(\varphi)=\alpha_i^{\boldsymbol{\sigma}}\,D_{\boldsymbol{\sigma}}\varphi^i\,\nu$ and

$$I([\alpha])(\varphi)=(-1)^{|\boldsymbol{\sigma}|}D_{\boldsymbol{\sigma}}\alpha_i^{\boldsymbol{\sigma}}\,\varphi^i\,\nu.$$

Considerations similar to what exposed in section 3.1 apply also here.

Case $p=2$**:** let $[\alpha]\in E_1^{1,n}$. Then $\nabla_\alpha(\varphi_1,\varphi_2)=2\alpha_{i\ j}^{\boldsymbol{\sigma}\boldsymbol{\tau}}D_{\boldsymbol{\sigma}}\varphi_1^iD_{\boldsymbol{\tau}}\varphi_2^j\,\nu$ (with $\alpha_{i\ j}^{\boldsymbol{\sigma}\boldsymbol{\tau}}=-\alpha_{j\ i}^{\boldsymbol{\tau}\boldsymbol{\sigma}}$) and, if α is a form on the r-th order jet, then

$$\begin{aligned}I([\alpha])(\varphi_1)(\varphi_2)&=(-1)^{|\boldsymbol{\tau}|}D_{\boldsymbol{\tau}}(2\alpha_{i\ j}^{\boldsymbol{\sigma}\boldsymbol{\tau}}D_{\boldsymbol{\sigma}}\varphi_1^i)\varphi_2^j\,\nu\\&=\sum_{\substack{\boldsymbol{\mu}+\boldsymbol{\sigma}=\boldsymbol{\rho}\\0\le|\boldsymbol{\rho}|\le 2r}}(-1)^{|\boldsymbol{\xi}+\boldsymbol{\mu}|}\frac{(\boldsymbol{\xi}+\boldsymbol{\mu})!}{\boldsymbol{\xi}!\boldsymbol{\mu}!}2D_{\boldsymbol{\xi}}\alpha_{ij}^{\boldsymbol{\sigma}\boldsymbol{\xi}+\boldsymbol{\mu}}D_{\boldsymbol{\rho}}\varphi_1^i\varphi_2^j\,\nu.\end{aligned} \tag{4.14}$$

Note that the above expressions coincide with the expressions of subsection 3.1 up to the isomorphism (4.5) and a constant factor (which depends on different conventions about numerical factors and the ordering in contractions and wedge products).

Remark 4.13 We observe that in [116] an intrinsic expression of e_1 which makes use of the above isomorphism was provided (see also [18, p. 195–197]).

A last step is needed in order to complete the computation of the $\mathcal{C}$-spectral sequence.

Theorem 4.14 *We have:*

- $S_2^{0,q}=H^q(E)$ *for* $0\le q\le n-1$*;*
- $S_2^{p,q}=0$ *for* $p>0$ *and* $0\le q\le n-1$*;*
- $S_2^{p,n}=H^{p+n}(E)$ *for* $0\le p$.

It turns out that $S_2=S_\infty$.

The only non-trivial statement of the above theorem is the last one. This follows from the convergence of the $\mathcal{C}$-spectral sequence to the de Rham cohomology and the fact that the differential e_2 always point either from 0 to $S_2^{p,n}$ or from $S_2^{p,n}$, so that it is the trivial map in both cases. For more details, see [18, 66, 118]. We just recall that the computation of the $\mathcal{C}$-spectral sequence for $J^\infty\pi$ is called *one-line theorem* [118, 121].

Corollary 4.15 *The variational sequence is obtained from the* $\mathcal{C}$*-spectral sequence by joining the two complexes* $(E_0^{0,q}, d_H)$ *and* $(E_1^{p,n}, e_1)$.

The above corollary is proved after proving that $s_1 = e_1$. This is quite easy, see [124].

We stress once again that the above construction yields the same results as in section 3 with the only exception of the cohomology of the rows. For this another spectral sequence would produce the results, namely the one arising from a filtration through horizontal forms.

5 Finite order variational sequence

The variational bicomplex and its derivation through the spectral sequence have been derived so far on infinite order jets. The reasons for doing that have been explained in section 2. But both the variational bicomplex and its derivation through the spectral sequence admit a finite-order counterpart, which has been studied in recent years.

The first statement of a partial version of finite order variational sequence was in [5]. This finite order variational sequence stopped with a trivial projection to 0 just after the space of finite order source forms (see section 3.1). The local exactness of this sequence was proved, together with an original solution of the global inverse problem (despite the fact that in order to do that the authors used infinite order jets). For more detailed comments about that variational sequence see remark 5.7.

The first formulation of a (long) variational sequence on finite order jet spaces is in [70] (see [72] for the case $n = 1$). Below we will describe the main points of the approach of [70], and compare it with other approaches. We also observe that more details can be found in [74]. The $\mathcal{C}$-spectral sequence on finite order jets of fibrings has been recently computed; the interested reader can find it in [125].

For the sake of completeness we also mention the paper [46]. In that paper the exactness of the horizontal de Rham sequence on finite order jets of submanifolds is proved. Nonetheless, we stress that this result could also be easily derived from the exactness results in [5, 70]. Another contribution has been given in [95], where the author stresses the relationship between a part of the finite order variational sequence and the Spencer sequence. This relationship was already explored in [115, 116] in the case of infinite order jet spaces.

The scheme of the finite order approach of [70] is the following. First of all we stress that the approach is developed in the language of sheaves. In [70] a natural exact subsequence of the de Rham sequence on $J^r\pi$ is defined. This subsequence is made by contact forms and their differentials. Then we define the r–th order variational sequence to be the quotient of the de Rham sequence on $J^r\pi$ by means of the above exact subsequence. Local and global results about the variational sequence are proved using the fact that the above subsequence is globally exact and using the abstract de Rham theorem.

Let us consider the sheaf of 1-contact forms $\mathcal{C}^1\Omega^*$, and denote by $(d\mathcal{C}^1\Omega^k)\tilde{}$ the sheaf generated by the presheaf $d\mathcal{C}^1\Omega^k$. We set

$$
\begin{aligned}
&\Theta^q_r \overset{\text{def}}{=} \mathcal{C}^1\Omega^q_r + (d\mathcal{C}^1\Omega^{q-1}_r)^{\smile} \quad 0 \leq q \leq n, \\
&\Theta^{p+n}_r \overset{\text{def}}{=} \mathcal{C}^p\Omega^{p+n}_r + (d\mathcal{C}^p\Omega^{p+n-1}_r)^{\smile} \quad 1 \leq p \leq \dim J^r\pi.
\end{aligned}
\tag{5.1}
$$

We observe that $d\mathcal{C}^1\Omega^{q-1}_r \subset \mathcal{C}^1\Omega^q_r$, so that the second summand of the above first equation yields no contribution to $\mathcal{C}^1\Omega^q_r$. Moreover, if we denote by c_r the dimension of the contact distribution on $J^r\pi$, the we observe that $\Theta^{p+n}_r = 0$ if $p+n > P \overset{\text{def}}{=} \dim J^r\pi - c_r$. Moreover, we have the following property (proved in [70]).

Lemma 5.1 *Let $0 \leq k \leq \dim J^r\pi$. Then the sheaves Θ^k_r are soft sheaves.*

We have the following natural soft subsequence of the de Rham sequence on $J^r\pi$

$$
0 \longrightarrow \Theta^1_r \xrightarrow{\ d\ } \Theta^2_r \xrightarrow{\ d\ } \cdots \xrightarrow{\ d\ } \Theta^P_r \xrightarrow{\ d\ } 0 \tag{5.2}
$$

Definition 5.2 The sheaf sequence (5.2) is said to be the *contact sequence*.

Theorem 5.3 *The contact sequence is an exact soft resolution of $\mathcal{C}^1\Omega^1_r$, hence the cohomology of the associated cochain complex of sections on any open subset of $J^r\pi$ vanishes.*

The above theorem is proved in [70] by first proving the local exactness of the contact sequence and then using standard results from sheaf theory (for which an adequate source is [126]).

Standard arguments of homological algebra prove that the following diagram is commutative, and its rows and columns are exact.

$$
\begin{array}{ccccccccccccccc}
 & & & & & & 0 & & 0 & & & & 0 & & \\
 & & & & & & \downarrow & & \downarrow & & & & \downarrow & & \\
 & & & & 0 & \longrightarrow & \Theta^1_r & \xrightarrow{d} & \Theta^2_r & \xrightarrow{d} & \cdots & \xrightarrow{d} & \Theta^I_r & \xrightarrow{d} & 0 \\
 & & & & & & \downarrow & & \downarrow & & & & \downarrow & & \\
0 \longrightarrow & \mathbb{R} & \longrightarrow & \Omega^0_r & & \xrightarrow{d} & \Omega^1_r & \xrightarrow{d} & \Omega^2_r & \xrightarrow{d} & \cdots & \xrightarrow{d} & \Omega^I_r & \xrightarrow{d} & \Omega^{I+1}_r \xrightarrow{d} \cdots 0 \\
 & & & & \searrow & & \downarrow & & \downarrow & & & & \downarrow & \nearrow_{\mathcal{E}_I} & \\
 & & & & & & \Omega^1_r/\Theta^1_r & \xrightarrow{\mathcal{E}_1} & \Omega^2_r/\Theta^2_r & \xrightarrow{\mathcal{E}_2} & \cdots & \xrightarrow{\mathcal{E}_{I-1}} & \Omega^I_r/\Theta^I_r & & \\
 & & & & & & \downarrow & & \downarrow & & & & \downarrow & & \\
 & & & & & & 0 & & 0 & & & & 0 & &
\end{array}
$$

Definition 5.4 The above diagram is said to be the *r-th order variational bicomplex* associated with the fibred manifold $\pi\colon E \to M$. We say the bottom row of the above diagram to be the *r-th order variational sequence* associated with the fibred manifold $\pi\colon E \to M$.

Due to theorem 5.3 the finite order variational sequence is an exact sheaf sequence (this means that the sequence is locally exact, [126]). Hence both the de Rham sequence and the variational sequence are acyclic resolutions of the constant sheaf $\mathbb{R}$ ('acyclic' means that the sequences are locally exact with the exception of the first sheaf $\mathbb{R}$). Next corollary follows at once.

Corollary 5.5 *The cohomology of the variational sequence is naturally isomorphic to the de Rham cohomology of* $J^r\pi$.

Having already dealt with local and global properties of the r-th order variational sequence, we are left with the problem of representing the quotient sheaves. Now it is obvious that, for $0 \leq q \leq n$, horizontalization provides such a representation (see [70, 122]).

Proposition 5.6 *Let* $0 \leq q \leq n$. *Then we have the isomorphism*

$$J_q \colon \Omega^q_r/\Theta^q_r \to \overline{\Omega}^{0,q}_r, \quad [\alpha] \mapsto h^{0,q}(\alpha).$$

The quotient differential $\mathcal{E}_q$ *reads through the above isomorphism as*

$$J_{q+1}(\mathcal{E}_q([\alpha])) = J_{q+1}([d\alpha]) = h^{0,q+1}(d\alpha) = d_H h^{0,q+1}(\alpha).$$

The last equality of the above equation is the least obvious, and was first proved in [5]. The proof depends on the fact that $D_\lambda u^i_{\boldsymbol{\sigma}\mu} = u^i_{\boldsymbol{\sigma}\mu\lambda}$, and that the indexes λ, μ are skew-symmetrized in the coefficients of α (see the coordinate expression of $h^{0,q}$).

Remark 5.7 In [5] the finite order variational sequence is developed starting from the idea of finding a subsequence of forms whose order do not change under d_H. The authors prove that the above property characterizes the forms which are the image of $h^{0,q}$ (see also [4]). Conversely, in [70] the idea is to start with forms on finite order jets, but the result is the same up to the degree $q = n$.

When the degree of forms is greater than n we are able to provide isomorphisms of the quotient sheaves with other quotient sheaves made with proper subsheaves. This helps both to the purpose of representing quotient sheaves and to the purpose of comparing the current approach with others, as we will see.

Proposition 5.8 *The horizontalization* $h^{p,n}$ *induces the natural sheaf isomorphism*

$$J_{p+n} \colon \Omega^{p+n}_r/\Theta^{p+n}_r \to \overline{\Omega}^{p,n}_r/h^{p,n}((d\mathcal{C}^p\Omega^{p+n-1}_r)), \quad [\alpha] \mapsto [h^{p,n}(\alpha)].$$

The quotient differential $\mathcal{E}_{p+n}$ *reads through the above isomorphism as*

$$J_{p+1+n}(\mathcal{E}_{p+n}([\alpha])) = J_{p+1+n}([d\alpha]) = [h^{p+1,n}(d\alpha)] = [d_H h^{p,n}(\alpha)].$$

For a proof, see [124]. A similar approach is in [67, 68].

Now, it is clear from proposition 5.8 that we are able to represent the quotient sheaves $\overline{\Omega}^{p,n}_r/h^{p,n}((d\mathcal{C}^p\Omega^{p+n-1}_r))$ using the interior Euler operator restricted to $\overline{\Omega}^{p,n}_r$; this is precisely the approach of [67, 68]. See [47] for a different approach to this problem. A further approach to the problem of representation appeared in [76]. Here the concept of *Lepagean equivalent* is introduced in full generality (older version of this concept can be found *e.g.*, in [69], with references to older foundational works). Namely, let $\alpha \in \Omega^{p+n}_r$. Then a Lepage equivalent of $[\alpha] \in E^{p,n}_1$ is a differential form $\beta \in \Omega^{p+n}_r$ such that

$$h^{p,n}(\beta) = h^{p,n}(\alpha), \qquad h^{p,q}(d\beta) = I(h^{p+1,n}(d\alpha)) = e_1([\alpha]).$$

The most important example of a Lepagean equivalent is the Poincaré–Cartan form of a Lagrangian (see, *e.g.*, [74]). A representation of forms in the variational sequence through Lepagian equivalents is currently being studied also in exterior differential systems theories [23].

Remark 5.9 It is interesting to observe that, either in view of theorem 4.5 or in view of the results by several authors referred to in remark 3.6, every form $\alpha \in \overline{\Omega}_r^{p,n}$ can be written as the sum $\alpha = \sigma + d_H\gamma$, where σ can be seen either as a skew-adjoint differential operator (from the isomorphism of proposition 4.4 and theorem 4.12) or as a form in the image of the interior Euler operator (which admits an equivalent characterization as skew-adjoint form, see [4]).

This means that, despite the fact that the denominator in proposition 5.8 is made by forms which are *locally* total divergences, only global divergences really matter.

The finite order formulation of [70] yields a variational sequence which can be proved to be equal to the finite order variational sequence obtained from a finite order analogue of the $\mathcal{C}$-spectral sequence [125]. Moreover, as one could expect, for $0 \leq s < r$ pull-back via $\pi_{r,s}$ yields a natural inclusion of the s-th order variational bicomplex into the s-th order variational bicomplex. More precisely, we have the following lemma (see [70]).

Lemma 5.10 *Let* $0 \leq s < r$*. Then we have the injective sheaf morphism*

$$\chi_s^r : \left(\Omega_s^k/\Theta_s^k\right) \to \left(\Omega_r^k/\Theta_r^k\right), \quad [\alpha] \mapsto [\pi_{r,s}^*\alpha].$$

Hence, there is an inclusion of the s*–th order variational bicomplex into the* r*–th order variational bicomplex.*

It can be proved that there exists an infinite order analogue of the above r-th order variational bicomplex [123]. This is defined in view of the above lemma via a direct limit of the injective family of r-th order variational bicomplexes. Nonetheless the direct limit infinite order bicomplex will be a bicomplex of presheaves, because gluing forms defined on jets of increasing order provides ‘forms’ which are only locally of finite order (see [43, 123] and the comments after theorem 3.14).

Remark 5.11 The main motivation for the finite order variational sequence has been a refinement in inverse problems of the calculus of variations. For example, a source form $[\alpha] \in \Omega_r^{n+1}/\Theta_r^{n+1}$ which is locally variational, *i.e.* $e_1([\alpha]) = 0$, admits a (local) Lagrangian $[\beta] \in \Omega_r^n/\Theta_r^n$. A representative of $[\beta]$ is $h^{0,n}(\beta)$, which is defined on the $r+1$-st order jet and depends on highest order derivatives through hyperjacobians (proposition 5.6). See [73, 123, 124, 125] for a comparison between the finite order and infinite order approaches.

6 Special topics

Due to space and time constraints it is not possible to go further in describing in details the current achievements in variational sequence theory. It is also impossible to reserve to applications and examples more than just a mention. The above tasks would require writing a whole book. But in this section at least the most important research directions of the last 15 years will be exposed, with reference to the literature for the readers who are interested in knowing more.

6.1 Inverse problem of the calculus of variations

The variational sequence is intimately related with the inverse problem of the calculus of variations (see the Introduction). This problem has a long history for which possible sources are the notes [91, p. 377] and references quoted therein, and [14, 88, 78] for the case of mechanics ($n = 1$). Here we briefly describe some inverse problems arising in the variational sequence, including the inverse problem of the calculus of variations. We just recall that the cohomology of the de Rham sequence on E is isomorphic to the cohomology of the variational sequence.

Variationally trivial Lagrangians. A variationally trivial Lagrangian is an element $[\alpha] \in E_1^{0,n}$ such that $e_1([\alpha]) = 0$. If $[\alpha]$ is a variationally trivial Lagrangian, then $[\alpha]$ is locally a total divergence, *i.e.*, $[\alpha] = d_H[\beta]$ with $[\beta] \in E_0^{0,n-1}$. A global horizontal $n-1$-form $[\beta] \in E_0^{0,n-1}$ such that $[\alpha] = d_H[\beta]$ exists if and only if $[[\alpha]] = 0 \in H^n(E)$. A refinement of this result is the following theorem.

Theorem 6.1 *Let* $\lambda\colon J^r\pi \to \wedge^n T^*M$ *induce a variationally trivial Lagrangian* $[\lambda]$*. Then, locally,* $\lambda = d_H\mu$*, where* $\mu = h^{0,n-1}(\alpha)$ *and* $\alpha \in \Omega_{r-1}^{n-1}$.

In other words, according to the above hypotheses, $\lambda = h^{0,n}(d\alpha)$, hence it depends on r-th order derivatives through hyperjacobians. This result has been proved in [5, 15, 75, 48][6] using various techniques. Note that the result is better with respect to the order of jets than what can be obtained by the local exactness of the finite order variational sequence. In fact, from the finite order variational sequence we would obtain $\alpha \in \Omega_r^{n-1}$. Of course, the result is sharp: the order cannot be further lowered.

Locally variational source forms. A locally variational source form is an element $[\alpha] \in E_1^{1,n}$ such that $e_1([\alpha]) = 0$. If $[\alpha]$ is a locally variational source form, then $[\alpha]$ is locally the Euler–Lagrange expression of a (local) Lagrangian, *i.e.*, $[\alpha] = \mathcal{E}[\beta]$ with $[\beta] \in E_1^{0,n}$. A global Lagrangian $[\beta] \in E_0^{0,n}$ such that $[\alpha] = \mathcal{E}[\beta]$ exists if and only if $[[\alpha]] = 0 \in H^{n+1}(E)$. A refinement of this result, like in the previous inverse problem, is much more difficult. We list the results which have been achieved till now.

Theorem 6.2 *Let* $[\alpha] \in \Omega_r^{n+1}/\Theta_r^{n+1}$ *be locally variational. Then there exists a (local) Lagrangian* $[\beta] \in \Omega_r^n/\Theta_r^n$ *such that* $[\alpha] = \mathcal{E}[\beta]$.

The above result is a direct consequence of the local exactness of the finite order variational sequence, and, as before, it is sharp with respect to the order [70, 122]. However, it can be very difficult to check that a source form is in the space $\Omega_r^{n+1}/\Theta_r^{n+1}$. A result proved in [4] is helpful in this sense. Let $u^{(r)}$ denote all derivative coordinates of order r on a jet space. Let $f \in C^\infty(J^{2r}\pi)$, and suppose that

$$f(x^\lambda, u^{(0)}, \dots, u^{(r)}, tu^{(r+1)}, t^2u^{(r+2)}, \dots, t^r u^{(2r)})$$

is a polynomial of degree less than or equal to r in $u^{(s)}$, with $r+1 \leq s \leq 2r$. Then f is said to be a weighted polynomial of degree r in the derivative coordinates of order $r+1 \leq s \leq 2r$.

Theorem 6.3 *Let* $[\Delta]$ *be a locally variational source form induced by* $\Delta\colon J^{2r}\pi \to C_0^* \wedge \wedge^n T^*M$*. Suppose that the coefficients of* Δ *are weighted polynomials of degree less than or equal to* r*. Then* $\Delta = \mathcal{E}(\lambda)$*, where* $\lambda\colon J^r\pi \to \wedge^n T^*M$.

Again, the result is sharp with respect to the order of the jet space where the Lagrangian is defined. The above theorem is complemented in [4] by a rather complex algorithm for

[6] In [15] the proof is for the special case when the Lagrangian does not depend on (x^λ).

building the lowest order Lagrangian. This algorithm is an improvement of the well-known Volterra Lagrangian

$$L = \int_0^1 u^i \Delta_i(x^\lambda, tu^j_{\boldsymbol{\sigma}})dt$$

for a locally variational source form Δ. In fact, the above Lagrangian is defined on the same jet space as Δ. The finite order variational sequence yields another method for computing lower order Lagrangians, provided we know that $\Delta = [\alpha] \in \Omega^{n+1}_r/\Theta^{n+1}_r$. Namely, we apply the contact homotopy operator to the closed form $d\alpha \in \Theta^{n+2}_r$, finding $\beta \in \Theta^{n+1}_r$ such that $d\beta = d\alpha$. By using once again using the (standard) homotopy operator we find $\gamma \in \Omega^n_r$ such that $d\gamma = \beta - \alpha$, and $\lambda \overset{\text{def}}{=} h^{0,n}(\gamma)$ is the required Lagrangian. Of course, the most difficult point is to invert the representation of quotients in the variational sequence, *i.e.*, to find a least order α such that $\Delta = [\alpha]$.

The above theorem does not exhaust the finite order inverse problem. A locally variational source form on $J^{2r}\pi$ seems to have a definite form of the coefficients with respect to its derivatives of order s, with $r + 1 \leq s \leq 2r$. A conjecture in this sense is formulated in [4] in an admittedly imprecise way. We *conjecture* that locally variational source forms defined on $J^{2r}\pi$ could be elements of $\Omega^{n+1}_r/\Theta^{n+1}_r$. Note that the representation through I of elements in $\Omega^{n+1}_r/\Theta^{n+1}_r$ yields source forms which are of order $2r+1$ and are obtained through the adjoint of the horizontalization of a form in Ω^{n+1}_r (which is a hyperjacobian polynomial of degree at most n in derivatives of order r); see [122] for more details about the structure of such forms.

Finally, we recall that recently some geometric results on variational first-order partial differential equations have been obtained in [54]. Such equations arise in multisymplectic field theories.

Symplectic structures. In [33] the symplectic structures for evolution equations are introduced. They are dual to the Hamiltonian structures mentioned in the introduction. A symplectic structure is an element $[\alpha] \in E^{2,n}_1$ such that $e_1(\alpha) = 0$ (see also[18]). It is clear that another inverse problem arises here. But there are no results as on the above section. It seems natural to formulate a conjecture on the structure of symplectic structures by analogy with the above conjecture.

Variational problems defined by local data. There are some examples of global source forms which do not admit a global Lagrangian. For instance, Galilean relativistic mechanics [97] and Chern–Simons field theories (where a global Lagrangian indeed exists but it is not gauge-invariant). Some authors proposed a general formalism for dealing with such situations. Namely, they introduce a sheaf of local n-forms all of which produce the same Euler–Lagrange source form under the action of $\mathcal{E}$. See [21, 19, 103, 104] for more details.

6.2 Variational sequence on jets of submanifolds

Let E be an $n+m$-dimensional manifold, and $x \in E$. We say that two n-dimensional submanifolds L_1, L_2 such that $x \in L_1 \cap L_2$ are *r-equivalent* if they have a contact of order r at x. It is possible to choose a chart of E at x of the form (x^λ, u^i), $1 \leq \lambda \leq n$, $1 \leq i \leq m$, where both L_1 and L_2 can be expressed as graphs $u^i = f^i_1(x^\lambda)$, $u^i = f^i_2(x^\lambda)$.

Then the contact condition is the equality of the derivatives of the above functions at x up to the order r. This is an equivalence relation whose quotient set is $J^r(E,n)$, the *r-th order jet space of n-dimensional submanifolds of E*[7]. If E is endowed with a fibring π, then $J^r\pi$ is the open and dense subspace of $J^r(E,n)$ which is made by submanifolds which are transverse to the fibring at a point (which, of course, can be locally identified with the images of sections, hence with local sections themselves).

Of course, jets of submanifolds have a contact distribution, hence a $\mathcal{C}$-spectral sequence can be formulated [29, 115, 116]. As a by-product a variational sequence is obtained. Jets of submanifolds can also be seen as jets of parametrizations of submanifolds (*i.e.*, jets of local n-dimensional immersions) up to the action of the reparametrization group [63]. In this setting another approach to the variational sequence is [99]. In [84] the finite-order $\mathcal{C}$-spectral sequence on jets of submanifolds is computed. See also the more comprehensive treaties [1, 119, 121] on the geometry of jets of submanifolds, partial differential equations and the calculus of variations. Another approach to the calculus of variations on jets of submanifolds can be found in [49].

6.3 Variational sequence on differential equations

There are several books on the geometric theory of differential equations (see the Introduction). We invite the interested reader to consult them. Here we just recall the main result related to the variational sequence on differential equations.

A differential equation (ordinary or partial, scalar or system) is a submanifold $S \subset J^r(E,n)$. Such a submanifold inherits the contact distribution from $J^r(E,n)$, hence the $\mathcal{C}$-spectral sequence can be defined on it. Let us describe what are the main differences with the 'trivial equation case', *i.e.*, the case of $S = J^r(E,n)$ or $S = J^r\pi$.

First of all, we observe that the term $E_1^{0,n-1}$ of the $\mathcal{C}$-spectral sequence of an equation is made by equivalence classes of conservation laws of the given equation up to trivial conservation laws. To realize it, it is sufficient to recall that conservation laws take the form of a total divergence which vanishes on the given equation (like, *e.g.*, continuity equations).

If S is closed then it can be represented as $F = 0$, where F is a section of a vector bundle over $J^r(E,n)$. Any differential equation $S = S^{(0)} \subset J^r(E,n)$ can be prolonged to a differential equation $S^{(1)} \subset J^{r+1}(E,n)$ which is locally described as $D_\lambda F^i = 0$. By iterating this procedure we obtain a sequence $\{S^{(i)}\}_{0\leq i\leq +\infty}$. We require that the equation S be *formally integrable*: this amounts at requiring that for every $i \in \mathbb{N}$ the restriction of $\pi_{i+1,i}$ to $S^{(i+1)}$ be a bundle over $S^{(i)}$. Hence the inverse limit $S^{(\infty)}$ can be constructed. We also require that the equation be *regular*: this means that the ideal of functions on $S^{(\infty)}$ is functionally generated by the differential consequences $D_{\boldsymbol{\sigma}} F^i$ of F. Finally, we say that S is ℓ-normal if the linearization of F has maximal rank (see [18, p. 198] for more details).

In [116, 118] the following theorem is proved ('two-lines theorem'): if S is formally integrable, regular and ℓ-normal, then the terms $E_1^{p,q}$ of the $\mathcal{C}$-spectral sequence on $S^{(\infty)}$ with $p > 0$, $1 \leq q \leq n-2$ are trivial. In other words, non-trivial terms of the $\mathcal{C}$-spectral sequence are distributed on the column $E_i^{0,q}$ for $1 \leq q \leq n-2$ and on the rows $E_j^{p,n-1}$, $E_j^{p,n}$ for $p \geq 1$; this explains the name of the theorem. Note that $E_\infty^{0,q} = E_1^{0,q}$ for

[7]The synonyms 'manifold of contact elements' [28] and 'extended jet bundles' [91] are also used.

$1 \leq q \leq n-2$ and $E_\infty^{p,n-1} = E_3^{p,n-1}$, $E_\infty^{p,n} = E_3^{p,n}$ for $p \geq 1$. An explicit description of the non-vanishing terms is also provided by the two-line theorem.

Most 'classical' differential equations of mathematical physics (KdV equation, heat equation, etc.) are ℓ-normal, but gauge equations (like Yang-Mills equation and Einstein equation) are not; the structure of their conservation laws is more complex than that of ℓ-normal equations [51]. This fact was not considered in [116, 118]. In [107] the method of compatibility complex was proposed to compute the number of non-trivial lines. That approach has been generalized in [114] (k-lines theorem) and compared with the Koszul–Tate resolution method in [113]. In [22] the same problem was considered in the framework of exterior differential systems (the author used the term 'characteristic cohomology' to indicate what we called the horizontal de Rham cohomology); see also [23].

Since then, several papers dealt with the $\mathcal{C}$-spectral sequence on differential equations. We recall the works [50, 59] on evolution equations and the works [7, 8, 9] on second-order parabolic and hyperbolic equations in the case $n = 2$.

6.4 Variational sequence and symmetries

Invariant variational problems. There are a number of variational problems which admit a group of symmetries G. The way to find invariant solutions for these problems is to find solutions of a reduced system on the space of invariants of G; this is related to Palais'principle of symmetric criticality. In the paper [6] the solution of this problem is related to the existence of a cochain map between the G-invariant variational bicomplex (see below) and the variational bicomplex on the space of invariants of G. The local existence of the cochain map is related to a relative Lie algebra cohomology group.

Lie derivatives of variational forms. The Lie derivative of variational forms, *i.e.*, elements of $E_1^{p,n}$ or equivalently $\mathcal{V}^p$, is interesting for the determination of symmetries of Lagrangians and source forms. However, the result of a Lie derivative with respect to a prolonged vector field is a form which, in general, contains d_H-exact terms. For this reason it is natural to derive a new operator, the variational Lie derivative, which is defined up to d_H-exact terms. Such a formula first appeared in [118] ('infinitesimal Stokes'formula').

Theorem 6.4 *Let $X\colon E \to TE$ be a vector field, and $[\alpha] \in E_1^{p,n}$. Then*

$$[\mathcal{L}_{X^\infty}\alpha] = e_1([i_{X^\infty}\alpha]) + i_{X^\infty}(e_1([\alpha])),$$

where the contraction $i_{X^\infty}(e_1([\alpha]))$ is defined by virtue of the identity $i_{X_V^\infty} \circ d_H = d_H \circ i_{X_V^\infty}$ and the fact that the action of X_H^∞ is trivial.

The above theorem can also be found in [4], and in [38, 77] in the finite order case. It has clear connections with Noether's theorem, for which we invite the reader to consult the above literature.

Evolutionary vector fields are one example of first-order differential operators with no constant term that preserve the contact distribution. For this reason, they yield operators on all the spaces $E_k^{*,*}$ of the $\mathcal{C}$-spectral sequence. More generally, the problem of finding 'secondary' differential operators, *i.e.*, higher order differential operators which preserve the contact distribution, has been faced [53]. A complete classification has not been achieved yet.

Takens'problem. It is well-known that, by virtue of Noether's theorem, any infinitesimal symmetry of a Lagrangian yields a conservation law of the corresponding Euler–Lagrange equations. Takens'problem [101] can be formulated as follows: *when a source form, endowed with a space of infinitesimal symmetries each of which generates a conservation law, is locally variational*.

The problem has been solved in several cases, besides the simplest ones in [101].

(1) Among the main results of [13], we have the following one. Consider the bundle $A \to M$, where A is the space of electromagnetic vector potentials, and let Δ be a source form. Suppose that Δ has translational and gauge symmetries and corresponding conservation laws. Then, if $n = 2$ and Δ is of third order, or $n \geq 3$ and Δ is of second order, Δ is locally variational.

(2) In [10] the case of second-order scalar differential equations is considered. A number of conditions on symmetries and conservation laws about which Takens'problem for the above equations admits an affirmative answer is derived.

(3) In [12] the case of polynomial differential equations which admit the algebra of Euclidean isometries and corresponding conservation laws is considered. The authors make use of the formal differential calculus by [42].

(4) Finally, in [92] the problem is considered for the case of systems of first order differential equations which admit the group of translations and corresponding conservation laws.

Invariant inverse problem. This problem can be described as follows: given a locally variational source forms which is invariant under the action of a group G, find (if it exists) a Lagrangian which is invariant under the action of G.

The problem admits a formulation in cohomological terms: consider a Lie group G (or a Lie pseudogroup G') acting on a manifold E[8]. Lift the action to $J^\infty(E, n)$. Then consider the G-invariant subcomplex of the variational sequence. Its cohomology is the G-invariant cohomology; it determines the solvability of the invariant inverse problem. The main difference with the non-invariant case is that the G-invariant cohomology could be different from zero even locally. The same consideration holds for infinitesimal actions.

The invariant variational bicomplex appeared in [106] together with several examples of applications, but without any specific mention to the invariant inverse problem. In a subsequent paper [3] (where the reader can also find a short story of the invariant inverse problem) the following invariant inverse problem was considered: to find natural Lagrangians for natural source forms on the bundle of Riemannian metrics on a given manifold M. Among the results it is interesting to note that, while the invariant $n+1$-st cohomology vanishes for $\dim M = 0,\ 1,\ 2 \bmod 4$, it is nonvanishing for $\dim M = 3 \bmod 4$, thus leading to an obstruction of Chern–Simons type to the existence of natural Lagrangians for natural source forms.

Further results in mechanics ($n = 1$) are exposed in [85, 86], where the obstruction to the existence of Lagrangians is found in the cohomology of the Lie algebra of G. It is proved that such an obstruction can be removed by a central extension of the group G.

[8] if the manifold is fibred, then the action is required to be projectable

In [11] the local inverse problem invariant with respect to a finite-dimensional Lie group action is completely solved. Namely, conditions under which the local invariant cohomology of the variational sequence is isomorphic to the local invariant de Rham cohomology of the total space E are given. Moreover, considering the action of a finite-dimensional Lie algebra, conditions under which the local invariant cohomology of the variational sequence is isomorphic to the cohomology of the Lie algebra are given. The paper is completed by several examples. In [93] the case of an infinite-dimensional Lie pseudogroup has been considered, and the local invariant cohomology is computed in terms of the Lie algebra cohomology of the formal infinitesimal generators of the pseudogroup. An application of the above methods and results is presented in [94].

In [60] the authors make use of a method of invariantization from the moving frames theory and compute the invariant counterparts of operators like the horizontal differential and the Euler–Lagrange operator.

The invariant variational bicomplex seems to be an important part of the BRST theory of quantized gauge fields [16], despite the fact that the mathematical side of that theory still needs deep investigation.

Differential invariants. The works [3, 106] (see also references therein) showed that the $\mathcal{C}$-spectral sequence invariant with respect to the pseudogroup of local diffeomorphisms provides a new approach to characteristic classes. In [52, 120] characteristic classes are interpreted as cohomologies of the regular spectra of the algebra of differential invariants.

6.5 Further topics

Variational multivectors. Variational forms, *i.e.*, elements of $E_1^{p,n}$, admit a dual counterpart. More precisely, 'standard' differential forms on a manifold M admit as a counterpart multivector fields, *i.e.*, sections of the bundle $\wedge^k TM$. The counterpart of the 'standard' exterior differential is the Schouten bracket. The counterpart for variational forms is constituted by variational multivectors. In [91] (where the word 'functional' is used instead of 'variational', see also references therein) the approach to variational multivectors is an 'integral' one, and multivectors are described in coordinates up to total divergences. A variational Poisson bracket is introduced. In [56] multivectors are explicitly described through the calculus of differential operators, and their bracket is analyzed in the graded case, which leads both to a variational Poisson bracket and to a variational Schouten bracket. We stress that such a bracket allows to define operators which are Hamiltonian, in the sense that their 'squared' bracket vanishes, without respect to a given Hamiltonian [58].

Variational sequences on supermanifolds. The problem of computing the analogue of the $\mathcal{C}$-spectral sequence for supermanifolds is almost completely open. We quote the paper [112] with a comprehensive list of references. There, integration, adjoint operators, Green's formula, the Euler operator and Noether's theorem are introduced in a noncommutative setting. As a by-product, an interesting characterization of Berezin volume forms is obtained.

7 Notes on the development of the subject

To the author's knowledge, the first papers where a variational sequence appeared are by Horndeski [55] and by Gel′fand and Dikii [42]. Horndeski constructed an analogue of the sequence (3.13) for a class of tensors (rather than forms) in coordinates, using jets in an implicit way, in order to study the inverse problem of the calculus of variations. Gel′fand and Dikii introduced the differentials d_H and $\mathcal{E}$ of the variational sequence (3.6) in the case $n = m = 1$, only for polynomial functions of $u^i_{\boldsymbol{\sigma}}$. The calculus that they developed was called by them the *formal calculus of variations*. This calculus was used to study the Hamiltonian formalism for evolution equations, of which they are among the main contributors. Their variational sequence was studied by Olver and Shakiban, who computed its cohomology [89]. An alternative approach to this problem is in [31].

At the same time Tulczyjew, studying the Euler–Lagrange differential [108], and speaking with Horndeski[9], matured the ideas that led to the variational bicomplex, first for higher n-dimensional tangent bundles $T^r_n M$ [109], then for jets of fibrings [110]. His results included the local exactness of the variational bicomplex, achieved through local homotopy operators. However, his results did not include the solution of the global inverse problem, *i.e.*, cohomological results about the variational sequence, until [111].

The $\mathcal{C}$-spectral sequence approach was developed independently by Dedecker [29] and Vinogradov [115, 116]. However, the most complete achievements about the $\mathcal{C}$-spectral sequence are due to Vinogradov. In fact, in [29] there is only the definition of the $\mathcal{C}$-spectral sequence, together with the definition of the variational sequence on jets of fibrings and submanifolds (see also the later paper [30]). Previous works by Dedecker made use of spectral sequences for the calculus of variations [25, 26, 27, 28], but none dealt with variational sequences. In [115, 116] all terms of the $\mathcal{C}$-spectral sequence are computed for jets of fibrings and jets of submanifolds ('one-line theorem'). The computation included a complete description of all terms of $E_1^{p,n}$ through the theory of adjoint operators and Green's formula. Moreover, the $\mathcal{C}$-spectral sequence was computed for the first time also on differential equations ('two-line theorem'). This last achievement led to the interpretation of conservation laws in terms of cohomology classes of the horizontal de Rham complex on the given equation and their computation. Vinogradov did not publish the detailed proofs of his results in [115, 116]; however he published a longer exposition of his results in [117] followed by a detailed exposition with proofs in [118]. Manin's review [83] of the geometry of partial differential equations devotes a section to the variational sequence. The material is based on results by Vinogradov and Kuperschmidt.

Independently, Takens [102] provided a formulation of the variational bicomplex together with local exactness and global cohomological results on jets of fibrings. His proofs of the local exactness relied are different with respect to those of Tulczyjew. After [102], Takens left this field of research and become an outstanding scientist in dynamical systems.

All the above approaches to variational sequences were developed on infinite order jets. Independently from the previous authors, Anderson and Duchamp [5] developed a new approach to variational sequences. The main novelty in their approach was the use of finite order jets. Their approach was formulated trying to find spaces of forms for which d_H was stationary with respect to the order of jets. Their approach did not provide a 'long' variational sequence, stopping with zero just after the space of source forms. Moreover,

[9] W. M. Tulczyjew, private communication

in the paper there is a cohomological computation about the global inverse problem, but this is performed on the infinite order jet. Another important result in the paper is the local classification of trivial Lagrangians of order r (but see also [15] for Lagrangians which do not depend on (x^λ)). Such a result has never been derived in an infinite order jets framework. Anderson is the author of the book [4], which, unfortunately, has never been finished. However, it is still a source of interesting proofs, examples, and facts, especially about the finite order inverse problem.

After that the foundations were established, a number of important contributions and improvements appeared in the literature.

In [106] Tsujishita reviewed the $\mathcal{C}$-spectral sequence and presented some new proofs of old facts together with new ideas and theorems (remarkably, the invariant $\mathcal{C}$-spectral sequence with interesting examples). A deeper analysis by several authors (Gessler [51], Krasil′shchik [65], Marvan [80], Tsujishita [107], Verbovetsky [114]) led to the generalization of Vinogradov's 'two-lines theorem' to the so-called 'k-lines theorem'. The fundamental tool for the computation of non-trivial lines in the $\mathcal{C}$-spectral sequence was the compatibility complex (see [114] and references therein).

The k-lines theorem was also proved in [16] in the framework of the BRST theory of quantized gauge fields [16]. A comparison between the approach of [16] (Koszul-Tate resolution) and the compatibility complex method was recently performed [113].

Bryant and Griffiths [22] proved similar results on the horizontal de Rham cohomology in the framework of exterior differential systems. They call such a cohomology the *characteristic cohomology* of an exterior differential system.

Duzhin began to study the finite order $\mathcal{C}$-spectral sequence, but he only completed the computations for first order jets of the trivial bundle $\pi = \mathrm{pr}_1 \colon M \times \mathbb{R} \to M$ (here pr_1 is the projection on the first factor) [34].

Krupka was the first one to formulate a 'long' variational sequence on finite order jets in [70]. This approach was formulated in the language of sheaves entirely in terms of finite order jet spaces. The idea is described in section 5. The results included local exactness and global cohomology of the finite order variational sequence, which turned out to be the same as the infinite order case. More precisely, it was proved that the direct limit of Krupka's variational bicomplex was the same as the variational bicomplex [123, 124], and that the $\mathcal{C}$-spectral sequence on finite order jets provides a finite order variational sequence which is the same as Krupka's one [124, 125]. The representation of Krupka's variational sequence was obtained by Krbek and Musilová in [67, 68] using the interior Euler operator adapted to the finite order case. The classification of variationally trivial Lagrangians was proved using local exactness of the finite order variational sequence [48, 75]. The Lepagean equivalent theory provided yet another representation of the variational sequence [76].

As a final remark, we observe that there are many research topics which are connected with variational sequences (such as the inverse problem of the calculus of variations). It is impossible to provide historical notes for all of them, for space and time constraints. The interested reader can consult the references indicated in section 6.

Appendix: splitting the exterior algebra

In propositions 2.8 and 2.15 we deal with two splittings of exterior algebrae which are induced by the splittings (2.2) and (2.12) of the underlying space. In order to make this paper self-contained we briefly describe how to obtain the exterior algebra projections from the underlying splitting projections [122, 124].

Let V be a vector space such that $\dim V = n$. Suppose that $V = W_1 \oplus W_2$, with $p_1 : V \to W_1$ and $p_2 : V \to W_2$ the related projections. Then, we have the splitting

$$\wedge^m V = \bigoplus_{k+h=m} \wedge^k W_1 \wedge \wedge^h W_2, \tag{7.1}$$

where $\wedge^k W_1 \wedge \wedge^h W_2$ is the subspace of $\wedge^m V$ generated by the wedge products of elements of $\wedge^k W_1$ and $\wedge^h W_2$.

There exists a natural inclusion $\odot^k L(V, V) \subset L(\wedge^k V, \wedge^k V)$. Then, the following identity can be easily proved:

$$\odot^n (p_1 + p_2) = \sum_{i=0}^{n} \binom{n}{i} \odot^i p_1 \odot \odot^{n-i} p_2.$$

It follows that the projections $p_{k,h}$ related to the splitting (7.1) turn out to be the maps

$$p_{k,h} = \binom{k}{p} \odot^k p_1 \odot \odot^h p_2 : \wedge^m V \to \wedge^k W_1 \wedge \wedge^h W_2.$$

Bibliography

[1] D. V. Alekseevsky, V. V. Lychagin, A. M. Vinogradov: *Basic ideas and concepts of differential geometry, Geometry I* (Encycl. Math. Sci., Vol. 28, Springer-Verlag, Berlin, 1991)

[2] R. Alonso Blanco: On the Green-Vinogradov formula *Acta Appl. Math.* **72(1-2)** (2002) 19–32

[3] I. M. Anderson: Natural variational principles on Riemannian manifolds *Annals of Math.* **120(2)** (1984) 329–370; `http://www.jstor.org/`

[4] I. M. Anderson: *The variational bicomplex* (book preprint, technical report of the Utah State University, 1989); available at `http://www.math.usu.edu/~fg_mp/`

[5] I. M. Anderson, T. Duchamp: On the existence of global variational principles *American J. of Math.* **102** (1980) 781–868; `http://www.jstor.org/`

[6] I. M. Anderson, M. E. Fels: Symmetry reduction of variational bicomplexes and the principle of symmetric criticality *American J. of Math.* **119** (1997) 609–670; `http://www.math.usu.edu/fg_mp/`

[7] I. M. Anderson, N. Kamran: La cohomologie du complexe bi-gradué variationnel pour les équations paraboliques du deuxième ordre dans le plan *Comptes Rendus Acad. Sci. (Paris), Série I* **321** (1995) 1213–1217

[8] I. M. Anderson, N. Kamran: Conservation laws and the variational bi-complex for second-order scalar hyperbolic equations in the plane *Acta Applic. Math.* **41** (1995) 135–144

[9] I. M. Anderson, N. Kamran: The Variational Bicomplex for Hyperbolic Second-Order Scalar Partial Differential Equations in the Plane *Duke J. Math.* **87** (1997) 265–319; `http://www.math.usu.edu/fg_mp/`

[10] I. M. Anderson, J. Pohjanpelto: Variational principles for differential equations with symmetries and conservation laws I: Second order scalar equations *Math. Ann.* **299** (1994) 191–222

[11] I. M. Anderson, J. Pohjanpelto: The cohomology of invariant variational bicomplexes *Acta Appl. Math.* **41** (1995) 3–19; `http://www.math.usu.edu/fg_mp/`

[12] I. M. Anderson, J. Pohjanpelto: Variational principles for differential equations with symmetries and conservation laws II: Polynomial differential equations *Math. Ann.* **301** (1995) 627–653

[13] I. M. Anderson, J. Pohjanpelto: Symmetries, Conservation Laws and Variational Principles for Vector Field Theories *Math. Proc. Royal Phil. Soc.* **120** (1996) 479–539; `http://www.math.usu.edu/fg_mp/`

[14] I. M. Anderson, G. Thompson: *The Inverse Problem of the Calculus of Variations for Ordinary Differential Equations* (Memoirs of the Am. Math. Soc., 98, n. 473, 1992)

[15] J. M. Ball, J. C. Currie, P. J. Olver: Null Lagrangians, weak continuity and variational problems of arbitrary order *J. Funct. Anal.* **41** (1981) 135–174

[16] G. Barnich, F. Brandt, M. Henneaux: Local BRST cohomology in the antifield formalism: I. General theorems *Comm. Math. Phys.* **174** (1995) 57–92; `arXiv:hep-th/9405109`

[17] M. Bauderon: Le problème inverse du calcul des variations *Ann. de l'I.H.P.* **36(2)** (1982) 159–179; `http://www.numdam.org`

[18] A. V. Bocharov, V. N. Chetverikov, S. V. Duzhin, N. G. Khor′kova, I. S. Krasil′shchik, A. V. Samokhin, Yu. N. Torkhov, A. M. Verbovetsky, A. M. Vinogradov: *Symmetries and Conservation Laws for Differential Equations of Mathematical Physics* (I. S. Krasil′shchik and A. M. Vinogradov eds., Amer. Math. Soc., 1999)

[19] A. Borowiec, M. Ferraris, M. Francaviglia, M. Palese: Conservation laws for non-global Lagrangians *In: Proc. Conf. 'Geometry and Topology of Manifolds'* (Krynica 2001; Univ. Iagel. Acta Math. 41, 2003) 319–331; `arXiv:math-ph/0301043`

[20] R. Bott, L. W. Tu: *Differential forms in algebraic topology* (G.T.M., Springer-Verlag, New York, 1982)

[21] J. Brajerčík, D. Krupka: Variational principles for locally variational forms *J. Math. Phys.* **46** (2005) 1–15

[22] R. L. Bryant, P. A. Griffiths: Characteristic cohomology of differential systems I: general theory *J. Amer. Math. Soc.* **8** (1995) 507–596;
R. L. Bryant, P. A. Griffiths: Characteristic cohomology of differential systems II: general theory *Duke Math. J.* **78** (1995) 531–676; `http://fds.duke.edu/db/aas/math/faculty/bryant/publications.html`

[23] R. L. Bryant, P. A. Griffiths, D. Grossman: *Exterior differential systems and the calculus of variations* (Chicago Lectures of Mathematics, Univ. of Chicago Press, 2003); `arXiv:math.DG/0207039`

[24] P. Dedecker: Sur un problème inverse du calcul des variations *Acad. Roy. Bélgique Bull. Cl. Sci.* **36(5)** (1950) 63–70

[25] P. Dedecker: Systemes differentiels exterieurs, invariants integraux et suites spectrales *In: Convegno Internazionale di Geometria Differenziale* (1953, Ed. Cremonese, Roma, 1954) 247–262

[26] P. Dedecker: Quelques applications de la suite spectrale aux intégrales multiples du calcul des variations et aux invariants intégraux, chapitre I: suites spectrales et généralisations *Bull. Soc. Roy. Sci. Liége* **24** (1955) 276–295;
P. Dedecker: Quelques applications de la suite spectrale aux intégrales multiples du calcul des variations et aux invariants intégraux, chapitre II: espaces de jets locaux et faisceaux *Bull. Soc. Roy. Sci. Liége* **25** (1956) 387–399

[27] P. Dedecker: Calcul des variations et topologie algébrique *Mémoires de la Société Royale des Sciences de Liège, Quatrième série, Fascicule I, Université de Liége* **19** (1957) 1-215

[28] P. Dedecker: On the generalization of symplectic geometry to multiple integrals in the calculus of variations *In: Diff. Geom. Methods in Math. Phys.* (Bonn, 1975; Lecture Notes in Mathematics 570, Springer–Verlag, Berlin, 1977) 395–456

[29] P. Dedecker: On applications of homological algebra to calculus of variations and mathematical physics *In: Proceedings of the IV international colloquium on differential geometry* (Santiago de Compostela, Universidad de Santiago de Compostela, Cursos y Congresos de la Universidad de Santiago de Compostela 15, 1978) 285–294

[30] P. Dedecker, W. M. Tulczyjew: Spectral sequences and the inverse problem of the calculus of variations *In: Internat. Coll. on Diff. Geom. Methods in Math. Phys.* (Aix-en-Provence, 1979; Lecture Notes in Mathematics 836, Springer–Verlag, Berlin, 1980) 498–503

[31] L. Dickey: On exactness of the variational bicomplex *In: Mathematical Aspects of Classical Field Theories* (Contemp. Math. 132 (M. Gotay, J. Marsden and V. Moncrief, eds.) Amer. Math. Soc., 1992) 307–314

[32] L. Dickey: *Soliton Equations and Hamiltonian Systems* (Advanced Series in Mathematical Physics, 26, World Scientific, 2nd edition, 2003)

[33] I. Dorfman: *Dirac structures and Integrability of nonlinear Evolution Equations* (Wiley, 1993)

[34] S. V. Duzhin: $\mathcal{C}$-spectral sequence on the manifold J^1M *Uspekhi Math. Nauk* **38** (in Russian), (1983) 165–166; *English translation: Russian Math. Surveys* **38(1)** (1983) 179–181

[35] S. V. Duzhin: $\mathcal{C}$-spectral sequence on manifolds of jets of finite order *Preprint* VINITI UDK 514.763.8, (in Russian) (1982)

[36] J. L. Ericksen: Nilpotent Energies in Liquid Crystal Theories *Arch. Rational Mech. Anal.* **10** (1962) 189–196

[37] M. Ferraris, M. Francaviglia: On the Global Structure of Lagrangian and Hamiltonian Formalism in Higher Order Calculus of Variations *In: Proc. of the Int. Meet. on Geometry and Physics* (Pitagora Editrice, Bologna, 1983) 12–15

[38] M. Francaviglia, M. Palese, R. Vitolo: Symmetries in finite order variational sequences *Czech. Math. J.* **52(1)** (2002) 197–213; `http://poincare.unile.it/vitolo`

[39] M. Francaviglia, M. Palese, R. Vitolo: A geometric formulation of Hessian and Jacobi tensors for higher order Lagrangians *Diff. Geom. Appl.* **22(1)** (2005) 105–120; `http://poincare.unile.it/vitolo`

[40] P. L. Garcia: The Poincaré–Cartan Invariant in the Calculus of Variations *In: Symposia Mathematica* (14, 1974) 219–246

[41] P. L. Garcia, J. Munoz: On the geometrical structure of higher order variational calculus *In: Proc. of the IUTAM–ISIMM Symp. on Modern Developments in Anal. Mech.* (Torino, 1982, Tecnoprint, Bologna, 1983) 127–147

[42] I. M. Gel′fand, L. A. Dikii: Asymptotic behaviour of the resolvent of Sturm–Liouville equations and the algebra of the Kortexeg–De Vries equations *Uspekhi Mat. Nauk* **30(5)** (in Russian), (1975) 67–100; *English translation: Russian Math. Surveys* **30(5)** (1975) 77–113

[43] G. Giachetta, L. Mangiarotti, G. Sardanashvily: Cohomology of the infinite-order jet space and the inverse problem *J. Math. Phys.* **42(9)** (2001) 4272–4282; `arXiv:math.DG/0006074`

[44] M. Giaquinta, M. Modica, J. Souček: *Cartesian currents in the calculus of variations, vol. I and II* (Springer, 1998)

[45] H. Goldschmidt, S. Sternberg: The Hamilton–Cartan formalism in the calculus of variations *Ann. Inst. Fourier, Grenoble* **23(1)** (1973) 203–267; `http://www.numdam.org`

[46] M. Grassi: Local vanishing of characteristic cohomology *Duke Math. J.* **102(2)** (2000) 307–328

[47] D. R. Grigore: The Variational Sequence on Finite Jet Bundle Extensions and the Lagrangian Formalism *Diff. Geom. and Appl.* **10** (1999) 43–77

[48] D. R. Grigore: Variationally trivial Lagrangians and Locally Variational Differential Equations of Arbitrary Order *Diff. Geom. and Appl.* **10** (1999) 79–105

[49] D. R. Grigore: Lagrangian formalism on Grassmann manifolds *In: Handbook of Global Analysis* (Amsterdam: Elsevier, 2007) 327–373

[50] D. M. Gessler: On the Vinogradov $\mathcal{C}$-spectral sequence for determined systems of differential equations *Diff. Geom. Appl.* **7(4)** (1997) 303-324; `http://diffiety.ac.ru`

[51] D. M. Gessler: The "three-line" theorem for the Vinogradov $\mathcal{C}$-spectral sequence of the Yang-Mills equation *Acta Appl. Math.* **56(2-3)** (1999) 139–153; `http://diffiety.ac.ru`

[52] D. M. Gessler, A. M. Verbovetsky, A. M. Vinogradov: Scalar differential invariants and characteristic classes of homogeneous geometrical structures *Mat. Zam.* **51(6)** (1992) 15–26; `http://diffiety.ac.ru`

[53] V. N. Gusyatnikova, A. M. Vinogradov, V. A. Yumaguzhin: Secondary Differential Operators *J. Geom. Phys.* **2** (1985) 23–65; `http://diffiety.ac.ru`

[54] A. Haková, O. Krupková: Variational first-order partial differential equations *J. Diff. Equat.* **191** (2003) 67–89

[55] G. W. Horndeski: Differential operators associated with the Euler–Lagrange operator *Tensor, N. S.* **28** (1974) 303–318

[56] S. Igonin, A. Verbovetsky, R. Vitolo: Variational Multivectors and Brackets in the Geometry of Jet Spaces *In: Proc. V Int. Conf. on Symmetry in Nonlinear Mathematical Physics* (Kyiv 2003 Inst. of Mathematics, Nat. Acad. of Sciences of Ukraine, 2004) 1335–1342; `http://poincare.unile.it/vitolo`.

[57] N. Kamran: *Selected topics in the Geometrical Study of Differential Equations* (CBMS Regional Conference Series in Mathematics 96, American Mathematical Society, 2002)

[58] P. Kersten, I. Krasil′shchik, A. Verbovetsky: On the integrability conditions for some structures related to evolution differential equations *Acta Appl. Math.* **83** (2004) 167–173; `arXiv:math.DG/0310451`

[59] N. G. Khorkova: On the $\mathcal{C}$-spectral sequence of differential equations *Diff. Geom. Appl.* **3** (1993) 219–243

[60] I. A. Kogan, P. J. Olver: Invariant Euler–Lagrange equations and the invariant variational bicomplex *Acta Appl. Math.* **76** (2003) 137–193; `http://www.math.umn.edu/~olver/`

[61] I. Kolář: A geometrical version of the higher order Hamilton formalism in fibred manifolds *J. Geom. Phys.* **1(2)** (1984) 127–137

[62] I. Kolář: Natural operators related with the variational calculus *In: Proc. of the Int. Conf. on Diff. Geom. and its Appl.* (1992, Silesian Univ., Opava, 1993) 461–472

[63] I. Kolář, P. Michor, J. Slovák: *Natural Operations in Differential Geometry* (Springer-Verlag, 1993); `http://www.emis.de/monographs/index.html`

[64] I. Kolář, R. Vitolo: On the Helmholtz operator for Euler morphism *In: Math. Proc. Camb. Phil. Soc.* (135, 2003) 277–290; `http://poincare.unile.it/vitolo`

[65] I. S. Krasil′shchik: Some new cohomological invariants for nonlinear differential equations *Diff. Geom. and Appl.* **2** (1992) 307–350

[66] I. S. Krasil′shchik, A. M. Verbovetsky: Homological methods in equations of mathematical physics *Preprint* Open Education and Sciences, Opava, Czech Rep., `arXiv:math.DG/9808130` (1998)

[67] M. Krbek, J. Musilová: Representation of the variational sequence by differential forms *Rep. Math. Phys.* **51(2/3)** (2003) 251–258

[68] M. Krbek, J. Musilová: Representation of the variational sequence by differential forms *Acta Appl. Math.* **88(2)** (2005) 177–199

[69] D. Krupka: Some geometric aspects of the calculus of variations in fibred manifolds *Folia Fac. Sci. Nat. UJEP Brunensis, Brno University* **14** (1973) 1–65; `arXiv:math-ph/0110005`

[70] D. Krupka: Variational sequences on finite order jet spaces *In: Proceedings of the Conf. on Diff. Geom. and its Appl.* (1989; World Scientific, New York, 1990) 236–254

[71] D. Krupka: The contact ideal *Diff. Geom. and its Appl.* **5** (1995) 257–276

[72] D. Krupka: Variational sequences in mechanics *Calc. Var.* **5** (1997) 557–583

[73] D. Krupka: Variational sequences and variational bicomplexes *In: Differential Geometry and Applications* (Proc. Conf., Edited by I. Kolář, O. Kowalski, D. Krupka, J. Slová k, Brno, August 1998; Masaryk Univ., Brno, Czech Republic, 1999) 525–531; `http://www.emis.de/proceedings`

[74] D. Krupka: Global variational theory in fibred spaces *In: Handbook of Global Analysis* (Amsterdam: Elsevier, 2007) 773–836

[75] D. Krupka, J. Musilova: Trivial Lagrangians in Field Theory *Diff. Geom. and Appl.* **9** (1998) 293–505

[76] D. Krupka, J. Šedenková: Variational sequences and Lepage forms *In: Differential Geometry and Its Applications* (Proc. Conf., Edited by J. Bures, O. Kowalski, D. Krupka, J. Slovák, Prague, August 2004; Charles University in Prague, Czech Republic, 2005) 617–627

[77] D. Krupka, O. Krupkova, G. Prince, W. Sarlet: Contact symmetries and variational sequences *In: Differential Geometry and Its Applications* (Proc. Conf., Edited by J. Bures, O. Kowalski, D. Krupka, J. Slovák, Prague, August 2004; Charles University in Prague, Czech Republic, 2005) 605–615

[78] O. Krupková: *The geometry of ordinary variational equations* (LNM 1678, Springer, 1997)

[79] B. A. Kuperschmidt: Geometry of jet bundles and the structure of Lagrangian and Hamiltonian formalism *In: Lecture Notes in Math.* (n. 775: Geometric Methods in Mathematical Physics, Springer, Berlin, 1980) 162–218

[80] M. Marvan: On the $\mathcal{C}$-spectral sequence with "general" coefficients *In: Proc. of the Conf. on Diff. Geom. and its Appl.* (Brno 1989; World Scientific, 1990) 361–371

[81] J. McCleary: *A user's guide to spectral sequences* (2nd ed., Cambridge Univ. Press, 2001)

[82] L. Mangiarotti, M. Modugno: Fibered Spaces, Jet Spaces and Connections for Field Theories *In: Proc. of the Int. Meet. on Geometry and Physics* (Pitagora Editrice, Bologna, 1983) 135–165

[83] Yu. I. Manin: Algebraic aspects of nonlinear differential equations *Journal of Mathematical Sciences* **11(1)** (1979) 1–122

[84] G. Manno, R. Vitolo: Variational sequences on finite order jets of submanifolds *In: Proc. of the VIII Conf. on Diff. Geom. and Appl.* (Opava, Czech Republic, 2001) 435–446; `http://www.emis.de/proceedings`; see also the longer preprints with expanded proofs and more results `arXiv:math.DG/0311078` and `arXiv:math.DG/0602127`.

[85] G. Marmo, G. Morandi: Symmetries in the Lagrangian and Hamiltonian formalism: the equivariant inverse problem *In: Symmetries in Science III* (Vorarlberg 1989; Plenum Press, 1989) 243–309

[86] G. Marmo, G. Morandi: On the inverse problem with symmetries and the appearance of cohomologies in classical Lagrangian dynamics *Rep. Math. Phys.* **28(3)** (1989) 389–410

[87] W. M. Mikulski: On naturality of the Helmholtz operator *Arch. Math. Brno* **41** (2005) 1–5

[88] G. Morandi, C. Ferrario, G. Lo Vecchio, G. Marmo, C. Rubano: The inverse problem in the calculus of variations and the geometry of the tangent bundle *Physics Reports* **188(3-4)** (1990) 147–284

[89] P. J. Olver, C. Shakiban: A Resolution of the Euler Operator *Proc. Am. Math. Soc.* **69** (1978) 223–229

[90] P. Olver: Hyperjacobians, determinantal ideals and weak solutions to variational problems *Proc. Roy. Soc. Edinburgh* **95A** (1983) 317–340

[91] P. J. Olver: *Applications of Lie Groups to Differential Equations* (GTM 107, 2nd edition, Springer, 1993)

[92] J. Pohjanpelto: Takens' problem for systems of first order differential equations *Ark. Mat.* **33** (1995) 343–456

[93] J. Pohjanpelto, I. M. Anderson: Infinite dimensional Lie algebra cohomology and the cohomology of invariant Euler-Lagrange complexes: a preliminary report *In: Proceedings of the 6th International Conference on Differential Geometry and Applications* (Brno, 1995 (J. Janyska, I. Kolar, J. Slovak, eds.) Masaryk University, Brno, Czech Republic, 1996) 427–448

[94] J. Pohjanpelto: The Cohomology of the Variational Bicomplex Invariant Under the Potential Kadomtsev-Petviashvili Equations *J. Nonlinear Math. Phys.* **4** (1997) 364–376

[95] J. F. Pommaret: Spencer sequence and variational sequence *Acta Appl. Math.* **41** (1995) 285–296

[96] J. J. Rotman: *An introduction to homological algebra* (Academic Press, 1979)

[97] D. Saller, R. Vitolo: Symmetries in covariant classical mechanics *J. Math. Phys.* **41(10)** (October 2000) 6824–6842

[98] D. J. Saunders: *The Geometry of Jet Bundles* (Cambridge Univ. Press, 1989)

[99] D. J. Saunders: Homogeneous variational complexes and bicomplexes *Preprint* `arXiv:math.DG/0512383` (2005)

[100] D. J. Saunders: Jet manifolds and natural bundles *In: Handbook of Global Analysis* (Amsterdam: Elsevier, 2007) 1035–1068

[101] F. Takens: Symmetries, conservation laws and variational principles *In: Lecture Notes in Mathematics No. 597* (Springer, 1977) 581–603

[102] F. Takens: A global version of the inverse problem of the calculus of variations *J. Diff. Geom.* **14** (1979) 543–562

[103] C. Tejero Prieto: Variational problems defined by local data *In: Proceedings of the Eighth International Conference on Differential Geometry and its Applications* (vol. 3, Mathematical Publications, Silesian University at Opava, O. Kowalski, D. Krupka, J. Slovak (Eds.), 2002) 473-483; `http://www.emis.de/proceedings`.

[104] C. Tejero Prieto: Variational formulation of ChernSimons theory for arbitrary Lie groups *J. Geom. Phys.* **50** (2004) 138-161

[105] E. Tonti: Variational formulation of nonlinear differential equations (I) and (II) *Acad. Roy. Belg. C (V)* **55** (1969) 137–165 and 262–278

[106] T. Tsujishita: On variation bicomplexes associated to differential equations *Osaka J. Math.* **19** (1982) 311–363

[107] T. Tsujishita: Homological method of computing invariants of systems of differential equations *Diff. Geom. Appl.* **1(1)** (1991) 3–34

[108] W. M. Tulczyjew: Sur la différentielle de Lagrange *C. R. Acad. Sc. Paris, série A* **280** (1975) 1295–1298

[109] W. M. Tulczyjew: The Lagrange Complex *Bull. Soc. Math. France* **105** (1977) 419–431

[110] W. M. Tulczyjew: The Euler–Lagrange Resolution *In: Internat. Coll. on Diff. Geom. Methods in Math. Phys.* (Aix–en–Provence, 1979; Lecture Notes in Mathematics, n. 836, Springer–Verlag, Berlin, 1980) 22–48

[111] W. M. Tulczyjew: Cohomology of the Lagrange complex *Ann. Scuola Norm. Sup. Pisa* **14** (1988) 217–227

[112] A. M. Verbovetsky: Lagrangian formalism over graded algebras *J. Geom. Phys.* **18** (1996) 195–214; `arXiv:hep-th/9407037`

[113] A. Verbovetsky: Remarks on two approaches to the horizontal cohomology: Compatibility complex and the Koszul-Tate resolution *Acta Appl. Math.* **72** (2002) 123–131; `arXiv:math.DG/0105207`

[114] A. M. Verbovetsky: Notes on the horizontal cohomology *In: Secondary Calculus and Cohomological Physics* (M. Henneaux, I. S. Krasil′shchik, and A. M. Vinogradov, eds., Contemporary Mathematics, vol. 219, Amer. Math. Soc., Providence, R.I., 1998) 211–231; `http://diffiety.ac.ru/`

[115] A. M. Vinogradov: On the algebro–geometric foundations of Lagrangian field theory *Soviet Math. Dokl.* **18** (1977) 1200–1204; `http://diffiety.ac.ru/`

[116] A. M. Vinogradov: A spectral sequence associated with a non–linear differential equation, and algebro–geometric foundations of Lagrangian field theory with constraints *Soviet Math. Dokl.* **19** (1978) 144–148; `http://diffiety.ac.ru/`

[117] A. M. Vinogradov: Geometry of nonlinear differential equations *J. Soviet Math.* **17(1)** (1981) 1624–1649

[118] A. M. Vinogradov: The $\mathcal{C}$–Spectral Sequence, Lagrangian Formalism and Conservation Laws I and II *Journal of Mathematical Analysis and Applications* **100(1)** (1984) 1–40 and 41–129; `http://diffiety.ac.ru/`

[119] A. M. Vinogradov: An informal introduction to the geometry of jet spaces *In: Rend. Seminari Fac. Sci. Univ.* (Cagliari 58, 1988) 301–333; `http://diffiety.ac.ru/`

[120] A. M. Vinogradov: Scalar differential invariants, diffieties and characteristic classes *In: Mechanics, Analysis and Geometry: 200 years after Lagrange* (M. Francaviglia (Editor), Elsevier, 1991) 379-414; `http://diffiety.ac.ru/`

[121] A. M. Vinogradov: *Cohomological analysis of Partial Differential Equations and Secondary Calculus* (Translations of Math. Monographs 204, Amer. Math. Soc., 2001)

[122] R. Vitolo: Finite order variational bicomplexes *Math. Proc. of the Camb. Phil. Soc.* **125** (1998) 321–333; `http://poincare.unile.it/vitolo`. A more detailed version is available at `arXiv:math-ph/0001009`

[123] R. Vitolo: A new infinite order formulation of variational sequences *Arch. Math. Un. Brunensis* **4(34)** (1998) 483–504; `http://poincare.unile.it/vitolo`

[124] R. Vitolo: On different geometric formulations of Lagrangian formalism *Diff. Geom. and its Appl.* **10(3)** (1999) 225–255; updated version at `http://poincare.unile.it/vitolo`

[125] R. Vitolo: Finite order formulation of Vinogradov C–spectral sequence *Acta Appl. Math.* **70(1-2)** (2002) 133–154; `arXiv:math.DG/0111161`; updated version at `http://poincare.unile.it/vitolo`

[126] R. O. Wells: *Differential Analysis on Complex Manifolds* (GTM, n. 65, Springer–Verlag, Berlin, 1980)

[127] V. V. Zharinov: *Lecture notes on geometrical aspects of partial differential equations* (World Scientific, 1992)

RAFFAELE VITOLO

Department of Mathematics "E. De Giorgi", University of Lecce,
Via per Arnesano, 73100 Lecce, Italy

E-mail: `raffaele.vitolo@unile.it`

Web: `http://poincare.unile.it/vitolo`

This page intentionally left blank

Handbook of Global Analysis
Demeter Krupka and David Saunders

The Oka-Grauert-Gromov principle for holomorphic bundles

Pit-Mann Wong

Contents

Introduction

In a very influential paper in 1939 Oka [41] discovered a condition on complex manifolds for which a Cousin II distribution is holomorphically solvable if it is topologically solvable. The reduction of holomorphic problems to topological problems is often refers to as the Oka Principle. Oka's article led to the invention of the concept of plurisubharmonic exhaustion by Grauert, the concept of Stein manifolds by Stein, the cohomological theory on complex manifolds by Dolbeault and H. Cartan, cumulating in the work of Grauert uniting the analytic, algebraic and topological theory into the theory of complex manifolds (especially Stein manifolds) that we know today. One of the spectacular achievements of Grauert is the extension, in 1957/58 [19] and [20], of Oka's Principle to holomorphic fiber bundles over Stein spaces, namely, two holomorphic principal bundles are biholomorphic if and only if they are topologically isomorphic. Grauert also showed that if two global holomorphic sections are homotopic through global continuous sections then they are also homotopic through global holomorphic sections. These results are extended further by Gromov [21], [22], [23] in 1989 to subelliptic bundles, a concept more flexible than fiber bundles. In the last few years, thanks to the effort of Eliashberg [10], Forster-Ramspott [13], Forstneric-Prezelj [16], Henkin-Leiterer [27], Larrusson [36] and many others, the theory (sometimes under the name of the homotopy principle or simply h-principle) is once again attracting a lot of attention. In this article we shall present an account of some of the main results of the theory together with some of the important applications: embedding dimension of Stein spaces (Forster and Ramspott [12, 13], Schürmann [46]), complete intersections (Schneider [45], Forstneric [14]) and complex hyperbolic geometry (Chandler-Wong [8], Wong-Wong [51]). In the last section we introduce briefly the algebraic version of the Oka Principle known as Serre's Problems. The statement that an algebraic on $\mathbf{K}^n$ (where $\mathbf{K}$ is an algebraically closed field) vector bundle is trivial is equivalent to the statement that every projective module over the polynomial ring $\mathrm{K}[t_1, ..., t_n]$ is free. This last statement was resolved in the affirmative by Quillen and Suslin independently in 1976. This can be extended to the ring convergent power series over p-adic numbers. For rigid analytic p-adic number fields, the analogue of complex Stein can be defined and can be used to deal with problems in p-adic hyperbolic geometry (see Anh-Wang-Wong [2]). This is the analogue of the application of the Oka Principle in complex hyperbolic geometry. We only give a very partial list of references due to the limitation of the length of the article, interested readers should look into the bibliography of the short list of articles that we provide.

1 Stein manifolds and Stein spaces

The most natural class of *non-compact* complex manifolds (resp., spaces) are the *closed* complex submanifolds (resp., subvarieties, i.e., common zeros of holomorphic functions) of $\mathbf{C}^N$. These manifolds are Kähler and have many (as many as one can hope for) holomorphic functions. An intrinsic characterization of these manifolds, now known as Stein manifolds, first appeared in a ground breaking paper by K. Stein [47] in 1951:

Definition 1.1. A complex manifold X is said to be a *Stein manifold* if the following three conditions are satisfied:

(i) *Global holomorphic functions separate points*, i.e., for any pair of distinct points $x_1 \neq x_2 \in X$ there exists a holomorphic function on X such that $f(x_i) \neq f(x_2)$.

(ii) *X is holomorphically convex*, i.e., for any compact set K in X, the *holomorphic convex hull*

$$\hat{K} = \{x \in X \mid |f(x)| \leq ||f||_K = \sup_{y \in K} |f(y)|\}$$

is also compact.

(iii) *Global holomorphic functions provide local coordinates*, i.e., for any $x \in X$ there exist holomorphic functions $f_1, ..., f_n, n = \dim_{\mathbf{C}} X$ on X such that $df_1 \wedge ... \wedge df_n(x) \neq 0$. $\square$

Remark 1.2. Condition (ii) above is equivalent to the condition:

for any discrete sequence $\{x_n\}$ in X there exists a global holomorphic function f such that $\limsup_i |f(x_i)| = \infty$. $\square$

We give here a very brief description of complex spaces. A *ringed space* is a Hausdorff topological space X together with a sheaf $\mathcal{O}_X$ (the structure sheaf) of associative and commutative *local* algebras over C. Recall that the *radical* of an ideal I is defined to be:

$$\mathrm{rad}\, I = \{x \in R \mid x^n \in I \text{ for some } n \in \mathbf{N}\}.$$

The radical of the zero ideal is called the *nilradical*. The *nilradical subsheaf* $\mathfrak{n}$ is the subsheaf whose stalk at x is the nilradical of the stalk of the structure sheaf at x. A ringed space $(X, \mathcal{O}_X)$ is said to be *reduced* if $\mathfrak{n} = 0$. In general the ringed space $(X, \mathcal{O}_X/\mathfrak{n})$ is reduced and is called the *reduction* of $(X, \mathcal{O}_X)$.

By a *closed analytic subvariety* (the term *closed analytic subset* is also commonly used in the literature) X of an open set U in C^N we mean the common zeros of global holomorphic functions on U together with a *structure sheaf* defined by $\mathcal{O}_X = \mathcal{O}_U/\mathcal{I}_X$ where $\mathcal{I}_X$ is the ideal sheaf of germs of holomorphic functions vanishing on X. The ringed space $(X, \mathcal{O}_X)$ is reduced.

It is convenient and important to allow *non-reduced* structure. We start from a coherent subsheaf (= subsheaf of finite type) of ideals $\mathcal{I}$ of $\mathcal{O}_U$ where U is an open set U in $\mathbf{C}^N$. Let $X = V(\mathcal{I}) = \mathrm{supp}(\mathcal{O}_U/\mathcal{I}) = \{x \in X \mid \mathcal{O}_{U,x} \neq \mathcal{I}_{U,x}\}$ and $\mathcal{O}_X = \mathcal{O}_U/\mathcal{I}$. The ringed space $(X, \mathcal{O}_X)$ is called a *closed analytic subspace* X of U. The reduction of $(X, \mathcal{O}_U/\mathcal{I})$ is the analytic subvariety $(X, \mathcal{O}_U/\mathcal{I}_X)$ ($\mathcal{I}_X$ is the ideal sheaf of X) introduced in the preceding paragraph.

Example 1.3. Let $X = \{(z, w) \in \mathbf{C}^2 \mid z = 0\}$ be one of the coordinate lines. The ideal sheaf $\mathcal{I}_X =< z >$ is generated by the function z then $(X, \mathcal{O}_{\mathbf{C}^2}/ < z >)$ is reduced. If we take $\mathcal{I} =< z2 >$, the ideal generated by z^2, then $(X, \mathcal{O}_{\mathrm{C}^2}/ < z^2 >)$ is non-reduced. □

Definition 1.4. A *complex space* is a *Hausdorff* space which admits a *countable basis of open sets* with a coordinate covering of open subsets $\{U_i\}$ and homeomorphism $f_i : U_i \to V_i$ where V_i is a subvariety in some open subset in $\mathbf{C}^{n_i}$ such that

$$f_{ij} = f_i \circ f_j^{-1} : f_j(U_i \cap U_j) \to f_i(U_i \cap U_j)$$

is a biholomorphic map. The structure sheaf $\mathcal{O}_X$ is defined via these biholomorphic maps. At each point x the *local embedding dimension* at x is the smallest integer n_x such that there exists an open neighborhood U_x which is biholomorphic to a closed subvariety of an open set in C^{n_x}. A complex space is said to have *bounded embedding dimension* if $\sup_{x \in X} n_x < \infty$. □

Remark 1.5. A local embedding allows us to define the structure sheaf $\mathcal{O}_X$ of a complex space X. The pair $(X, \mathcal{O}_X)$ is a ringed space. The local embedding also shows that X is *locally compact* and *locally pathwise connected*. Some authors omitted the condition that the topology admits a countable basis of open sets. For our purpose this is included in the definition to avoid certain pathological situations. This condition implies that X is *metrizable* hence *paracompact* (every open cover has a locally finite refinement) and that *X is a countable union of of compact sets.* □

In the presence of singular points the third condition in Definition 1.1 obviously requires modification:

Definition 1.6. A complex space X is said to be *Stein* if

(i) (separability) global holomorphic functions separate points;

(ii) (convexity) X is holomorphically convex;

(iii) (uniformizable) global holomorphic functions generate the maximal ideal at each point of X, i.e., for any $x \in X$ there exist finitely many global holomorphic functions such that the germs of these functions at x generate the maximal ideal $\mathfrak{m}_x$(= the sheaf germs of functions *vanishing* at x) of $\mathcal{O}_{X,x}$. □

In the older literature (for example in Grauert's original articles) the terminology *holomorphically complete* is often used instead of Stein space.

The following topological property of a Stein space (due to Andreotti-Frankel [1] in the non-singular case and to Hamm [26] in the general case) plays an important role in many important results in the *cohomological* as well as the *analytical* theory of Stein spaces:

Theorem 1.7. *A Stein space of complex dimension n has the homotopy type of a CW complex of real dimension n.*

In the preceding theorem homotopy can be replaced by homology or by intersection homology (see [17]):

Corollary 1.8. *Let X be a complex n-dimensional Stein space. Then*

$$H_i(X, \mathbf{Z}) = IH_i(X, \mathbf{Z}) = 0$$

for $i > n$, moreover, $H_n(X, \mathbf{Z})$ and $IH_n(X, \mathbf{Z})$ has no torsion.

Example 1.9. (see Patrizio-Wong [42]) Let X be a pure dimension closed subvariety of $\mathbf{C}^N$. For a generic point $z_0 \in \mathbf{C}^N$ the function $\tau = ||z - z_0||^2|_X$ is a Morse function on X. Assume that X is non-singular then τ is a strictly plurisubharmonic non-negative function on X. Let m_0 be the minimum of τ on X. The set $X_0 = \{x \in X \mid \tau(x) = m_0\}$ is called the *center* of X (relative to τ). By a well-known theorem (of Harvey-Wells) the center is a totally real submanifold of real dimension at most n. If τ has no critical point except for those in the center (i.e., $X_0 = \{x \in X \mid d\tau|_x = 0\}$) then τ is said to be a *canonical exhaustion function* of radius $0 < R \leq \infty$. The following assertion is clear.

The center X_0 of a canonical exhaustion has the same homotopy type as X.

If a canonical exhaustion function exists then standard Morse theory implies that X can be homotopically deformed onto the center X_0. Replacing τ by $\tau - m_0$ *we may always assume that the minimum value is* 0. A strictly plurisubharmonic exhaustion function $\tau : X \to [0, R)$, of class C^∞ (C^5 is enough), on X is said to be a *Monge-Ampére exhaustion* if the function $\log \tau$ satisfies the complex homogeneous Monge-Ampére equation:

$$(\partial\bar{\partial} \log \tau)^n = 0, \quad n = \dim_{\mathbf{C}} X$$

on $X \setminus X_0$. The preceding equation is equivalent to the following equation:

$$\sum_{\alpha,\beta} \tau^{\bar{\beta}\alpha} \tau_\alpha \tau_{\bar{\beta}} = \tau$$

where subscripts indicate partial derivatives and superscripts indicates raising of indices. So $\tau_\alpha = \partial\tau/\partial z_\alpha, \tau_\beta = \partial\tau/\partial\bar{z}_\beta; (\tau_{\alpha\bar{\beta}})$ is the Levi-form and $(\tau^{\bar{\beta}\alpha})$ is the inverse matrix. The preceding equation simply means that

$$||d\tau||_h^2 = \tau \tag{*}$$

where h is the Kähler metric defined by the Levi-form $\sqrt{-1}\partial\bar{\partial}||z||^2$. It is clear that:

A Monge-Ampére exhaustion is canonical.

Condition (*) implies that $\tau = 0$ is the only critical value of a Monge-Ampére exhaustion. We give a few examples of Stein manifolds with Monge-Ampére exhaustions. The obvious one is C^n with $\tau = ||z||^2$ (and the unit ball B^n. Indeed it is known that, up to biholomorphisms $(\mathrm{C}^n, ||z||^2)$ and $(B^n, ||z||^2)$ (infinite or finite radius) are the only Stein manifolds with a *smooth* ($\mathcal{C}^5$ is enough; regularity is crucial here) Monge-Ampére exhaustion such that the center consists of exactly one point.

Examples of Stein manifolds which can be deformed into a compact symmetric spaces of rank one are also known. The following are Stein (actually affine algebraic) manifolds:

$$\begin{aligned}
(M_I^n) &= Q^n, \text{ the complex affine quadric}, n \geq 2;\\
(M_{II}^n) &= \mathrm{P}^n(\mathbf{C}) \setminus \overline{Q}^{n-1} \text{ where } \overline{Q}^{n-1} \text{ is compact complex quadric}, n \geq 2;\\
(M_{III}^{2n}) &= (\mathbf{P}^n(\mathbf{C}) \times \mathbf{P}^n(\mathbf{C})) \setminus \mathbf{P}_\infty^N(\mathbf{C}), n \geq 1, N = (n+1)2 - 1;\\
(M_{IV}^{4n}) &= \mathbf{G}\mathrm{r}(2, 2n, \mathbf{C}) \setminus \mathrm{P}_\infty^N(\mathbf{C}), n \geq 1, N = n(2n-1) - 1;\\
(M_V^{16}) & \text{ is a 16 dimensional Stein manifold.}
\end{aligned}$$

Each of these Stein manifolds admit a strictly plurisubharmonic exhaustion function $\tau : M \to [1, \infty)$ with center the respective compact symmetric space of rank one:

$$\begin{aligned}
(I) & \text{ the } n\text{-sphere } S^n, n \geq 2;\\
(II) & \text{ the real projective space } \mathbf{P}^n(\mathbf{R}), n \geq 2;\\
(III) & \text{ the complex projective space } \mathbf{P}^n(\mathbf{C}), n \geq 1;\\
(IV) & \text{ the quaterionic projective space } \mathbf{P}^n(\mathbf{H}), n \geq 1;\\
(V) & \text{ the Cayley projective plane } \mathrm{P}^n(\mathrm{C}ayley).
\end{aligned}$$

Moreover, the function $\cosh^{-1}\tau$ satisfies the complex homogeneous equation:

$$(dd^c \cosh^{-1}\tau)^m = \left(dd^c \log\left(\tau + \sqrt{\tau^2 - 1}\right)\right)^m = 0$$

outside of the center and where $m = \dim M$.

For example the affine hyperquardics

$$Q^n = \{z = (z_1, ..., z_{n+1}) \in \mathbf{C}^{n+1} \mid z_1^2 + \cdots + z_{n+1}^2 = 1\}$$

contains the n-sphere $S^n = \{x_1^2 + \cdots + x_{n+1}^2 = 1\}$ where $x_j = \mathrm{Re}\, z_j$. The function

$$\tau(z) = ||z||^2|_{Q^n} = (|z_1|^2 + \cdots + |z_{n+1}|^2)|_{Q^n}$$

is a strictly plurisubharmonic exhaustion function with center S^n. The other cases, though more complicated, are analogously verified. □

The theorem of Andreotti-Frankel was established by constructing a proper Morse function on the Stein manifold X with the property that all critical points have index bounded from above by the *complex* dimension of X. It turns out that this condition characterize Stein manifolds (Eliashberg [11]):

Theorem 1.10. *Let X be an open smooth almost complex manifold of real dimension $2n$. If there exists a proper Morse function $\rho : X \to [0, \infty)$ such that all critical points have index $\leq n$ then X admits a Stein structure.*

The following theorem is a partial list of various different characterization of Stein manifolds (see Gunning [25] for the proof):

Theorem 1.11. *Let X be a complex manifold. Then the following conditions are equivalent.*

(1) *X is Stein.*

(2) *Global holomorphic functions separate points and X is holomorphically convex.*

(3) *For any discrete sequence, finite or infinite, $\{x_i \mid i \in I\}$ and any sequence of complex numbers $\{c_i \mid i \in I\}$ there exists a holomorphic function on X such that $f(x_i) = c_i$ for all $i \in I$.*

(4) *For any point $x \in X$ there exists a sequence of holomorphic functions $\{f_i\}$ on X such that x is an isolated point of $A = \cap_i \{z \in X \mid f_i(z) = 0\}$ and X is holomorphically convex.*

(5) *X contains no compact complex subvariety of strictly positive dimension and X is holomorphically convex.*

(6) *The sheaf cohomology groups $H^i(X, \mathcal{S}) = 0$ for all $i \geq 1$ and for all coherent sheaf $\mathcal{S}$ on X.*

(7) *There exists a proper real valued function $\rho : X \to \mathrm{R}$ such that the Levi form $dd^c\rho = \sqrt{-1}\partial\bar{\partial}\rho$ is positive definite.*

(8) *There is a holomorphic embedding of X as a closed complex submanifold of C^N for some N.*

Remark 1.12. (i) If X is a domain in a Stein manifold Y then X is Stein if and only if $H^i(X, \mathcal{O}_X) = 0$ for all $1 \leq i \leq n = \dim_{\mathbf{C}} X$ where $\mathcal{O}_X$ is the sheaf of germs of holomorphic functions. If X is a domain in $\mathbf{C}^n$ then X is Stein if and only if $H^i(X, \mathcal{O}_X) = 0$ for all $1 \leq i \leq n-1$.

(ii) For complex spaces the first 6 conditions are equivalent. The equivalence extends also to (7) with the following modification:

$(7')$ *There exists a proper real valued function $\rho : X \to \mathrm{R}$ such that, at each point $x \in X$ there exists an open neighborhood U_x of x, an embedding $h_x : U_x \to V_x$ where V_x is an analytic subset of an open set B in C^{n_x} and a function $\tilde{\rho}_x$ of class $\mathcal{C}^2$ with positive definite Levi form such that $\rho = h_x^* \tilde{\rho}_x = \tilde{\rho}_x \circ h_x$.*

(iii) Stein spaces satisfying condition (8) are precisely those with *bounded local embedding dimension*, more precisely, those with the property that the supremum (over $x \in X$) of the local embedding dimension at x is bounded.

(iv) For the question of the best (i.e., the smallest) possible embedding dimension N see section 7 below.

2 Oka's theorem

Let X be a complex space and $\mathcal{U} = \{U_i\}$ an open cover of X. A *Cousin I distribution* is a collection $\{(U_i, f_i) \mid f_i$ is a meromorphic function on $U_i\}$ satisfying the condition that $f_i - f_j = f_{ij}$ is holomorphic on $U_i \cap U_j$. The condition means that f_i and f_j have the same principal part on $U_i \cap U_j$. Cousin's first (or the additive Cousin) problem is to find condition so that there exists a global meromorphic function on X with the given prescribed principle parts. In one complex variable this is known as the Mittag-Leffler's problem. It is a well-known fact that the problem is solvable if $H^1(X, \mathcal{O}_X) = 0$. In particular this is true if X is Stein (see Theorem 1.11 part (6) or Remark 1.12).

A collection $\{(U_i \cap U_j, f_{ij}) \mid f_{ij}$ is *non-vanishing* and holomorphic on $U_i \cap U_j\}$ is called a *Cousin II distribution* if the multiplicative cocycle condition is satisfied:

$$f_{ij} f_{jk} f_{ki} = 1.$$

The Cousin II (or the multiplicative Cousin) problem is to find conditions so that there exists, for all i, a *non-vanishing* holomorphic function f_i on U_i such that

$$f_{ij} = f_i / f_j$$

on $U_i \cap U_j$.

The theory of Oka's principle originated from the insightful work of Oka in 1939 that, on a domain of holomorphy (= Stein domain) in C^n, the Second Cousin Problem is *holomorphically* solvable if and only if it is *continuously* solvable:

Theorem 2.1. (Oka 1939) *Let X be a complex space satisfying the condition that*

$$H^1(X, \mathcal{O}_X) = 0$$

($\mathit{this\ is\ the\ case\ if}\ X\ \mathit{is\ Stein}$) *and $\{(U_i \cap U_j, f_{ij})\}$ be a Cousin II distribution where $\{U_i\}$ is an open cover of X by Stein open subsets. Assume that there exist continuous non-vanishing functions such that c_i on U_i such that $c_i/c_j = f_{ij}$ is non-vanishing and holomorphic on $U_i \cap U_j$. Then there exists, on each U_i, a non-vanishing holomorphic functions f_i such that $f_i/f_j = c_i/c_j$ on $U_i \cap U_j$ for all i, j.*

Proof. We may write $c_i = \exp \chi_i$ where each χ_i is continuous. The condition that c_i/c_j is holomorphic on $U_i \cap U_j$ is equivalent to the condition that $\chi_i - \chi_j$ is holomorphic on $U_i \cap U_j$. Moreover $\chi_{ij} = \chi_i - \chi_j$ satisfies the additive cocycle condition:

$$\chi_{ij} + \chi_{jk} + \chi_{ki} = 0.$$

If $H^1(X, \mathcal{O}_X) = 0$ then the additive Cousin Problem is solvable, namely, there exist holomorphic functions ϕ_i on U_i such that $\phi_i - \phi_j = \chi_i - \chi_j$ for all i, j. The functions $f_i = \exp \phi_i$ on U_i satisfy the requirement of the theorem. □

Oka's theorem preceded the introduction of the concept of Stein spaces and the original theorem assume that X is a domain of holomorphy but the proof of the more general form requires no new technique. Oka's theorem can also be viewed as follows. The Cousin II distribution $\{U_i \cap U_j, f_{ij}\}$ is a cocycle, i.e., an element of $Z^1(\mathcal{U}, \mathcal{O}_X^*)$ (where $\mathcal{O}_X^*$ is the sheaf of germs of non-vanishing holomorphic functions) hence defines an element in $H^1(X, \mathcal{O}^*)$ = isomorphism classes of holomorphic line bundles on X. Oka's theorem asserts that a holomorphic line bundle over a Stein space ($H^1(X, \mathcal{O}_X)$ is sufficient) is holomorphically trivial if it is topologically trivial. In fact, on a Stein space, the injection $\mathcal{O}_X^* \to \mathcal{C}_X^*$ of the sheaf of germs of non-vanishing holomorphic functions into the sheaf of germs of non-vanishing continuous functions induces an isomorphism:

$$H^1(X, \mathcal{O}^*) \cong H^1(X, \mathcal{C}^*).$$

In other words, *on a Stein space every continuous line bundle admits a holomorphic structure and two holomorphic line bundles are holomorphically isomorphic if and only if they are continuously isomorphic*. This assertion can be established as follows. In the continuous category there is a short exact sequence:

$$0 \to \mathrm{Z} \to \mathcal{C} \xrightarrow{\epsilon} \mathcal{C}^* \to 0$$

(where $\epsilon(f) = \exp(2\pi\sqrt{-1}f)$) inducing a long exact sequence:

$$\cdots \to H^1(X,\mathcal{C}) \to H^1(X,\mathcal{C}^*) \to H^2(X,\mathrm{Z}) \to H^2(X,\mathcal{C}) \to \cdots$$

The sheaf of germs of continuous functions is a fine sheaf hence $H^i(X,\mathcal{C}) = 0$ for all $i \geq 1$. This implies that

$$H1(X,\mathcal{C}^*) \cong H^2(X,\mathbf{Z}).$$

On a complex manifold there is an analogous short exact sequence of sheaves:

$$0 \to \mathrm{Z} \to \mathcal{O} \xrightarrow{\epsilon} \mathcal{O}^* \to 0$$

inducing a long exact sequence:

$$\cdots \to H^1(X,\mathcal{O}) \to H^1(X,\mathcal{O}^*) \to H^2(X,\mathbf{Z}) \to H^2(X,\mathcal{O}) \to \cdots$$

If the groups $H^i(X,\mathcal{O}) = 0$ for $i = 1$ and 2 (for example, if X is Stein) then:

$$H^1(X,\mathcal{O}^*) \cong H^2(X,\mathbf{Z}).$$

Consequently, Oka's Theorem is valid for any such manifolds:

Theorem 2.2. *Let X be a complex manifold satisfying the condition that*

$$H^1(X,\mathcal{O}) = H^2(X,\mathcal{O}) = 0$$

(*this is the case if X is Stein*) *then*

$$H^1(X,\mathcal{O}^*) \cong H^1(X,\mathcal{C}^*) \cong H^2(X,\mathbf{Z});$$

consequently, every continuous complex line bundle admits a holomorphic structure

Corollary 2.3. *Let X be a complex manifold satisfying the condition of theorem 2.2. If in addition, $H^2(X,\mathbf{Z}) = 0$, then all continuous and holomorphic line bundles are trivial.*

Remark 2.4. (a) As remarked in section 1, for domains in $\mathbf{C}^n, n \leq 3$, the condition in Theorem 2.2 is equivalent to Stein but is strictly weaker than Stein if $n \geq 4$.

(b) If X is an open (= non-compact) Riemann surface then the conditions of Corollary 2.3 are satisfied, hence every holomorphic line bundle is trivial. As we shall see later that, in fact, the same is true for vector bundles of arbitrary rank (see Theorem 5.1 below).

(c) The assumptions of Theorem 2.2 are satisfied for some compact complex manifolds (which are of course not Stein), for example, $H^i(\mathbf{P}^n,\mathcal{O}) = 0$ for all $i \geq 0$. However, the condition that $H^2(\mathbf{P}^n,\mathbf{Z}) = 0$ in Corollary 2.3 is never satisfied if X is compact Kähler, for example, $H^2(\mathbf{P}^n,\mathbf{Z}) \cong \mathbf{Z}$.

3 Grauert's Oka principle

Over a Stein manifold Oka's result in the preceding section may be equivalently reformulated as follows.

Theorem 3.1. Let $p : E \to X$ be a holomorphic principle bundle over a Stein manifold. If the fiber is $\mathbf{C}^*$ then every continuous section is homotopic to a holomorphic section.

Oka's result was extended by Grauert in 1957 to any principal G-bundle. Besides Grauert's original papers, the article of H. Cartan [6] is also a good reference for the materials of this section.

Let G be a complex Lie group and F a complex space. We say that G acts on F holomorphically if there is a group homomorphism $\phi : G \to \mathrm{Aut}\, F$ (where $\mathrm{Aut}\, F$ is the group of biholomorphic self maps of F. The action is said to be *effective* if the homomorphism ϕ is injective. We shall always assume that G acts effectively, from the left, on F. Let $p : E \to X$ be a *continuous* (*resp. holomorphic*) *fiber bundle* over X with fiber F and Lie group G. This means that p is a continuous (resp. holomorphic) surjection and the following axioms are satisfied:

(i) p is locally trivial, i.e., there exists an open cover $\{U_i\}$ of X and continuous (resp. holomorphic) bundle isomorphism $\phi_i : p^{-1}(U_i) \to U_i \times G$ for all i,
(ii) $\phi_{ij} = \phi_i \circ \phi_j^{-1} : p^{-1}(U_i \cap U_j) \to p^{-1}(U_i \cap U_j)$ is of the form

$$\phi_{ij}(x, \xi) = (x, \psi_{ij}(x)\xi))$$

where $\psi_{ij} : U_i \cap U_j \to G$ is a continuous (resp. holomorphic) map satisfying the condition:

$$\psi_{ij}\psi_{jk} = \psi_{ik}.$$

A fiber bundle is determined up to isomorphisms by the cocycles $\{\psi_{ij}\}$. In other words, the set of isomorphism classes of continuous (resp. holomorphic) fiber bundle is $H^1(X, \mathcal{G}_c)$ (resp. $H^1(X, \mathcal{G}_h)$) where $\mathcal{G}_c$ (resp. $\mathcal{G}_h$) is the sheaf of germs of continuous (resp. holomorphic) functions on X with values in G. The space of sections over an open set U in X, denoted $\Gamma(U, \mathcal{G}_c)$ and $\Gamma(U, \mathcal{G}_h)$, are equipped with the topology of uniform convergence on compact subsets of U. This induces a topology on the space of cocycles $Z^1(\mathcal{U}, \mathcal{G}_c)$ and hence also a topology on the set $H^1(\mathcal{U}, \mathcal{G}_c)$ where $\mathcal{U}$ is an open cover of X. Taking direct limit over open covers $\mathcal{U}$ yields a topology on the set $H^1(X, \mathcal{G}_c)$. It is well-known that (see for example Hirzebruch [28]) $H^1(X, \mathcal{G}_c)$ (resp. $H^1(X, \mathcal{G}_h)$) is the set of all isomorphic classes of continuous (resp. holomorphic) principal fiber bundles with typical fiber G. If G is not abelian $H^1(X, \mathcal{G}_c)$ and $H^1(X, \mathcal{G}_c)$ do not have the structure of a group but there is a distinguished element corresponding to the trivial fiber bundle.

A fiber bundle E is said to be a *principal G-bundle* if $F = G$ with G acting on itself via left translations. The group G acts naturally on E from the right. Indeed we can directly define a principal G-bundle $p : E \to X$ by postulating a right action by G on E satisfying the axioms:
(i) local trivialization, $\phi_i : U_i \times G \overset{\approx}{\to} p^{-1}(U_i)$,
(ii) $\phi_i(x, g)g' = \phi_i(x, gg')$ for all i and for all $g, g' \in G$.

Let E be a principal G-bundle and F a complex space on which G acts effectively form the left. A fiber bundle with typical fiber F associated to E is constructed as follows.

Define an action of G on $E \times F$ by

$$(\gamma, \xi)g = (\gamma g, g^{-1}\xi)$$

for all $g \in G$ and $\gamma \in E, \xi \in F$. The orbit space $E \times_G F = (E \times F)/G$ is a fiber bundle. The bundles E and $E \times_G F$ are said to be associated to each other. Bundles associated to each other define the same element in the set $H^1(X, \mathcal{G}_c)$ (resp. $H^1(X, \mathcal{G}_h)$).

Theorem 3.2. (Grauert) *Let X be a complex space and and G be a complex Lie group. Denote by $\mathcal{G}_c$ and $\mathcal{G}_h$ the sheaf of germs of continuous and, respectively, holomorphic functions on X with values in G. Assume that X is Stein then the inclusion $\iota : \mathcal{G}_h \to \mathcal{G}_c$ induces a bijection $\iota_* : H^1(X, \mathcal{G}_h) \to H^1(X, \mathcal{G}_c)$. Consequently, (i) every continuous principle G-bundle admits a unique holomorphic structure, (ii) two holomorphic principle G-bundles are holomorphically isomorphic if and only they are continuously isomorphic.*

Corollary 3.3. *On a contractible Stein space X every holomorphic principal G-bundle is holomorphically trivial.*

Let E a holomorphic principal G-bundle over a complex space X. The space of continuous and respectively, holomorphic sections are denoted by $\Gamma(X, \mathcal{C}(E))$ and $\Gamma(X, \mathcal{O}(E))$ where $\mathcal{C}(E)$ and $\mathcal{O}(E)$ are the sheaf of germs of continuous sections and, respectively, holomorphic sections of E. These spaces are equipped with the topology of uniform convergence on compact subsets of U.

Definition 3.4. Let X and Z be complex spaces and $X_0 \subset X$. Let $f_0, f_1 : X \to Z$ be continuous maps with f_0 holomorphic. We say that f_0 can be *deformed to* f_1 (or f_0 *and* f_1 *are homotopic*) relative to Z if there is a continuous map

$$F : X \times [0, 1] \to Z$$

such that $f_0 = F(\cdot, 0), f_1 = F(\cdot, 1)$. If both f_0 and f_1 are holomorphic we say that f_0 can be *deformed through holomorphic maps to*) f_1 (or f_0 *and* f_1 *are homotopic through holomorphic maps to* f_1) relative to Y if, in addition, F is *real analytic* and $f_t = F(\cdot, t) : X \to Y$ is holomorphic for $0 \le t \le 1$.

The following results (useful in applications) are variations of Grauert's Oka Principle formulated in terms of sections.

Theorem 3.5. (Homotopy) *Let E be a holomorphic principal G-bundle. Let Y be a $($possibly empty$)$ closed analytic subset of a Stein space X. Then*

(i) every continuous section $f_0 : X \to E$, such that $f_0|Y$ is holomorphic, is homotopic through continuous sections relative to Y to a holomorphic section $f_1 : X \to E$;

(ii) any two holomorphic sections $f_0, f_1 : X \to E$ which are homotopic through continuous sections relative to Y are also homotopic through holomorphic sections relative to Y.

Theorem 3.6. (Runge Approximation) *Let E be a holomorphic principal G-bundle over a Stein space X. Let $f : U \to E$ be a holomorphic section over a holomorphically convex open subset U of X. If f can be arbitrary approximated, on compact subsets of U, by global continuous sections of E then it can be arbitrary approximated, on compact subsets of U, by global holomorphic sections of E.*

Theorem 3.7. (Extension) *Let E be a holomorphic principle G-bundle. Let U be a holomorphically convex open subset of a Stein space X and Y be a closed analytic subset of X. Let $f : U \to E$ and $g : Y \to E$ be holomorphic sections, defined on U and Y respectively, satisfying the condition that $f = g$ on $Y \cap U$. If f can be approximated, on compact subsets of $U \cap Y$, by global continuous sections of E which are extensions of g then it can be approximated, on compact subsets of $U \cap Y$, by global holomorphic sections of E which are extensions of g.*

Remark 3.8. If we take E to be a trivial G-bundle then the preceding theorems yield the various version of Oka principle for maps into a complex Lie group G. For example we get from Theorem 3.5 that every continuous map from a Stein space into G is homotopic to a holomorphic map and two holomorphic G-valued maps are homotopic through continuous maps then they are also homotopic through holomorphic maps.

Definition 3.9. Let B be a subgroup of a topological group A with unit 1. Denote by $\bar{I}^q$ the q-fold, $q \geq 1$, Cartesian product of the unit interval $\bar{I} = [0, 1]$ and

$$J_{q-1} = \{(t_1, ..., t_q) \in \partial \bar{I}^q \mid t_q \neq 0\}.$$

For $q \geq 1$ the set of all continuous map $f : \bar{I}^q \to A$ such that

$$f(\partial \bar{I}^q) \subset B \text{ and } f(J_{q-1}) = 1$$

is a topological group, denoted $\rho_q(A, B)$, with the obvious structure induced by the group structure of A. We set by convention $\rho_0(A, B) = A$. The connected components of $\rho_q(A, B)$ also form a group and will be denoted by $\pi_q(A, B)$. If $B = \{1\}$ we write $\rho_q(A)$ for $\rho_q(A, 1)$ and $\pi_q(A)$ for $\pi_q(A, 1)$.

Theorem 3.10. *Let X be a Stein space and E be a principal G-bundle over X. Then the injection $\iota : \Gamma(X, \mathcal{O}(E)) \to \Gamma(X, \mathcal{C}(E))$ is a weak homotopy equivalence, i.e., induces a bijection $\iota_* : \pi_q(\Gamma(X, \mathcal{O}(E))) \to \pi_q(\Gamma(X, \mathcal{C}(E)))$ for all integer $q \geq 0$.*

Remark 3.11. In fact Theorems 3.5, 3.6, 3.7 and 3.10 are valid for fiber bundles with complex homogeneous spaces as fibers.

There is also the question about the converse: to what extent does the Oka principle characterize the base space X? For instance, is X Stein? The first result in this direction is due to Kajiwara and Nishihara [31]:

Theorem 3.12. *Let X be a two dimensional Stein manifold and let $D \subsetneqq X$ be an open subset. Let G be a complex Lie group and $\mathcal{G}_h$ and $\mathcal{G}_c$ be the sheaves of germs of holomorphic and resp. continuous functions on D with values in G. If the inclusion $\iota : \mathcal{G}_h \to \mathcal{G}_c$ induces a bijection of $H^1(D, \mathcal{G}_h)$ and $H^1(D, \mathcal{G}_c)$ then D is Stein.*

The theorem is false if the dimension of X is three or higher: simply take $G = \mathbf{C}, D = \mathbf{C}^n \setminus \{0\}$ and $X = \mathbf{C}^n$. For $n \geq 3, H^1(D, \mathcal{O}) = H^1(D, \mathcal{C}) = 0$ but D is not Stein.

Kajiwara [29] also obtained the following result:

Theorem 3.13. *Let X be a Stein manifold of arbitrary dimension and D be a domain in X with continuous boundary. Suppose that there exists a complex Lie group G with the property that for every polydisc P in X and every open covering $\mathcal{U}$ of $D \cap D$ a cocycle $f \in Z^1(\mathcal{U}, G)$ is $\mathcal{G}_c$-trivial then it is also $\mathcal{G}_h$-trivial. Then D is Stein.*

Kajiwara also provided a counter-example of a domain D with discontinuous boundary. Leiterer [38] found a condition which works for all domains and for all dimensions:

Theorem 3.14. *Let X be a Stein manifold of arbitrary dimension and D a domain in X with $H^1(D, \mathcal{O}) = 0$. If every continuously trivial bundle on D is also holomorphically trivial then D is Stein.*

Henkin and Leiterer showed that Grauert's Principle is valid on the following class of manifolds slightly more general than Stein:

Definition 3.15. A complex manifold X is said to be *pseudoconvex* if there exists an exhaustion (i.e., proper) function $\rho : X \to \mathrm{R}$ of class $\mathcal{C}^2$ such that ρ is strictly plurisubharmonic (i.e., the Levi form $dd^c\rho$ is positive definite) on $X - X_0$ where $X_0 = \rho^{-1}((-\infty, 0])$.

Theorem 3.16. *Theorem 3.5 is valid for pseudoconvex manifolds.*

Grauert's original approach is based on induction on the dimension of the base X. Henkin and Leiterer [27] used an approach known as the bumping technique (I believe also due to Grauert for other purpose). However the bumping technique works only if the base X is non-singular.

4 Gromov's Oka principle

In 1989 Gromov further ([21], [22], [23]) extended Oka's principle to much wider classes of spaces that are not fiber bundles. Some of his results were written up in details later by Forstneric and Prezelj [17], [16]. These results extend Grauert's Oka Principle for holomorphic fiber bundles $p : E \to X$ over a Stein manifold to certain submersions $p : Z \to X$. Submersion here means that the map p and its differential $dp : TZ \to TX$ are surjective. The condition on Z is that it should be *subelliptic*. By way of motivation we start with a simpler notion (due to Gromov):

Definition 4.1. Let $p : Z \to X$ be a holomorphic submersion of complex manifolds and $h : E \to Z$ be a holomorphic vector bundle over Z. A holomorphic map $s : E \to Z$ is said to be a *vertical spray* (or simply, a spray) associated to the submersion if its restriction to the zero section (which may be identified with Z) is the identity self-map of Z, i.e., $s(0_z) = z$ and fiber preserving, i.e., $s(E_z) \subset Z_{p(z)}$ (where E_z and $Z_{p(z)}$ are fibers over the respective points) for all $z \in Z$. Denote by $s_z = s|E_z : E_z \to Z_{p(z)}$. A spray s is said to be *dominating at a point* $z \in Z$ if the differential

$$ds_z : T_{0_z}E_z \to T_z Z_{p(z)} = \ker dp_z$$

at the point 0_z is a surjection. A spray s is said to be *dominating* if is dominating at every point $z \in Z$, in other words, $ds : E \to VTZ(=$ vertical tangent bundle = $\ker dp)$ is surjective. □

Essentially, the situation of a vector bundle $E \to X$ is replaced by the situation of a triple $(E \xrightarrow{s} Z \xrightarrow{p} X)$ where s is a vector bundle and p is a submersion. Roughly speaking it is required that the bundle structure is compatible (fiber preserving) with the submersion. With this set up the results of section 3 remain valid:

Theorem 4.2. (Homotopy) *Let* $(E \xrightarrow{s} Z \xrightarrow{p} X)$ *be a dominating vertical spray over a Stein manifold* X. *Let* Y *be a (possibly empty) closed analytic subset of* X. *Then*

(i) *every continuous section* $f_0 : X \to E$ *such that* $f_0|Y$ *is holomorphic is homotopic through continuous sections relative to* Y *to a holomorphic section* $f_1 : X \to E$;

(ii) *any two holomorphic sections* $f_0, f_1 : X \to E$ *which are homotopic through continuous sections relative to* Y *are also homotopic through holomorphic sections relative to* Y.

Theorem 4.3. (Runge Approximation) *Let* $(E \xrightarrow{s} Z \xrightarrow{p} X)$ *be a dominating vertical spray over a Stein manifold* X. *Let* $U \subset X$ *be a holomorphically convex open subset of* X *and* $f : U \to Z$ *be a holomorphic section. If* f *can be arbitrary approximated, on compact subsets of* U, *by global continuous sections of* Z *then it can be arbitrary approximated, on compact subsets of* U, *by global holomorphic sections of* Z.

Theorem 4.4. (Extension) *Let* $(E \xrightarrow{s} Z \xrightarrow{p} X)$ *be a dominating vertical spray over a Stein manifold* X. *Let* U *be a holomorphically convex open subset of a Stein space* X *and* Y *be a closed analytic subset of* X. *Let* $f : U \to E$ *and* $g : Y \to E$ *be holomorphic sections, defined on* U *and* Y *respectively, satisfying the condition that* $f = g$ *on* $Y \cap U$. *If* f *can be arbitrary approximated, on compact subsets of* $U \cap Y$, *by global continuous sections of* E *which are extensions of* g *then it can be arbitrary approximated, on compact subsets of* $U \cap Y$, *by global holomorphic sections of* E *which are extensions of* g.

Corollary 4.5. (Weak Homotopy Equivalence) *Let* $(E \xrightarrow{s} Z \xrightarrow{p} X)$ *be a dominating vertical spray over a Stein manifold* X. *The inclusion* $i : \Gamma_{\text{hol}}(X, Z) \to \Gamma_{\text{cont}}(X, Z)$ *of the space of holomorphic sections in the space of continuous sections is a weak homotopy equivalence, i.e., induces a bijection* $\iota_* : \pi_q(\Gamma_{\text{hol}}(X, Z)) \to \pi_q(\Gamma_{\text{hol}}(X, Z))$.

It is not so easy to come up with interesting examples of dominating sprays. The situation improves dramatically by slightly weaken the notion above (this is again due to Gromov):

Definition 4.6. A holomorphic submersion of complex manifolds $p : Z \to X$ is said to be *subelliptic* if at every point $x \in X$ there exists an open neighborhood U of x such that $p|_U : Z|_U \to U$ admits finitely many sprays $s_i : E_i \to Z, i = 1, ..., k$ such that

$$s_1(E_{1,z}) + \cdots + s_k(E_{k,z}) = VT_z Z.$$

If $k = 1$ at every point then the submersion is said to be *elliptic*.

The proofs of the preceding results work just as well for subellptic submersions:

Theorem 4.7. *The conclusions of Theorems 4.2, 4.3, 4.4 and Corollary 4.5 are valid for a subelliptic holomorphic submersion* $p : Z \to X$ *over a Stein manifold* X.

We give below examples of sprays, elliptic and subelliptic submersions. These examples can be found in the article by Gromov or Forstnerič [16],

Example 4.8. (1) Let G be a complex Lie group with Lie algebra $\mathfrak{g}$. Let $E = G \times \mathfrak{g} \to G$ then

$$s : E \to G, \quad s(g, t) = (\exp t)g$$

is a dominating spray.

(2) Let $\phi^i : \mathrm{C} \times Y \to Y, i = 1, ..., N$ be holomorphic maps with the property that each

$$\phi_t^i : Y \to Y,\ \phi_t^i(y) = \phi^i(t, y),\ t \in \mathbf{C}$$

is an one-parameter subgroup of automorphisms of Y. Then

$$s = (\phi^1, ..., \phi^N) : Y \times \mathrm{C}^N \to Y,\ s(y, t_1, ..., t_N) = (\phi^1(t_1), ..., \phi^N(t_N))$$

is a spray which is dominating if the vector fields associated to $\phi_t^i, i = 1, ..., N$ span the fibers of the tangent bundle TY at each point of Y. This is the case if Y is homogeneous. Using the one parameter subgroups associated to these vector fields we obtain a dominating spray on C^n.

(3) Let Σ be a complex subvariety of codimension at least two in a complex Grassmannian Y. Then $Y \setminus \Sigma$ is subelliptic. More generally, let $p : Z \to X$ be a holomorphic fiber bundle, with typical fiber a complex Grassmannian, over a complex manifold X. Let Σ be a complex subvariety in Z such that $\Sigma \cap Z_x$ is of codimension at least two in Z_x for all $x \in X$. Then $Z \setminus \Sigma \to X$ is a subelliptic bundle over X.

(4) If Y is subelliptic then any unramified cover $\tilde{Y}$ of Y is also subelliptic. The converse, whether $\tilde{Y}$ subelliptic implies Y subelliptic, is open.

(5) If each $Y_i, i = 1, ..., n,$ is subelliptic then the Cartesian product $\Psi_{i=1}^n Y_i$ is also subelliptic.

(6) The open Riemann surface $Y = \mathbf{P}^1 \setminus \{d$ distinct points $\}$ is subelliptic if and only if $d \leq 2$. More generally, $Y = \mathbf{P}^n \setminus \{d$ hyperplanes in general position $\}$ is subelliptic if and only if $d \leq n + 1$. Notice that $\mathbf{P}^1 \setminus \{d$ distinct points $\}$ is hyperbolic if $d > 2$ and more generally, $Y = \mathbf{P}^n \setminus \{d$ hyperplanes in general position $\}$ is measure hyperbolic if $d > n + 2$.

(7) (Rosay) The manifold

$$Y = \mathbf{C}^2 \setminus \big(\{w = 0\} \cup \{w = 1\} \cup \{w = kz\} \cup \{zw = 1\}\big)$$

is not subelliptic.

5 The case of Riemann surfaces

Grauert's Oka principle implies that every holomorphic vector bundle over a contractible Stein space is holomorphically trivial. In the special case of open Riemann surfaces (all of which are Stein) triviality is automatic without any topological assumption:

Theorem 5.1. (Röhrl [43] 1957) *Every holomorphic vector bundle over an open Riemann surface is holomorphically trivial.*

We recall here a classical theorem on vector bundles over the Riemann sphere for comparison:

Theorem 5.2. (Grothendieck [23] 1957) *Every holomorphic vector bundle over the Riemann sphere* $\mathbf{P}1$ *is holomorphically isomorphic to a direct sum of holomorphic line bundles.*

The analogue of Theorem 5.1 in higher dimension was observed by a number of people. The first results are due to Forster-Ramspott in a series of paper beginning in 1966. The next result can be derived from the work of M. Schneider in [45] using Hamm's topological lemma (see Theorem 1.7 in section 1) :

Theorem 5.3. *Let X be a pure n-dimensional Stein space and E a holomorphic vector bundle of rank $r \geq [n/2]$ over X. Then*

(a) *there is a holomorphic vector bundle F of rank $[n/2]$ such that E is holomorphically isomorphic to $F \oplus \mathcal{O}_X^{r-[n/2]}$,*

(b) *if n is even then there is a holomorphic vector bundle F such that E is holomorphically isomorphic to $F \oplus \mathcal{O}_X^{r-(n/2)+1}$ precisely when the Chern class $c_{n/2}(E)$ vanishes.*

Röhrl's Theorem is the case $n = 1$ in part (a) of the preceding theorem. For $n = 2$ part (b) asserts that a vector bundle is trivial if and only if $c_1(E) = 0$. By Hamm's result Theorem 5.3 is reduced to the case of complex vector bundles over CW-complexes of real dimension n (so $H^2(X, \mathbf{Z}) = 0$ if X is an open Riemann surface, Röhrl's Theorem follows immediately because $H^1(X, \mathcal{O}^* = H^2(X, \mathbf{Z}))$. For a proof of the assertions of the theorem in the later case we refer the readers to Husemoller [29]. The next result is another example of combining Oka Principle with Hamm's Lemma:

Theorem 5.4. *Let X be a pure n-dimensional Stein space. Let E and F be holomorphic vector bundles of rank $r \geq [n/2]$ over X. Assume that the torsion in $H^2(X, \mathbf{Z})$ is relatively prime to $(m-1)!$ for all $m \geq 3$. Then E and F are holomorphically isomorphic if the two bundles have the same total Chern class.*

The preceding result (with continuous isomorphism) is known for continuous bundles over a CW complex (see Peterson [43]). The theorem now follows from Oka-Grauert's Principle.

Definition 5.5. Let X, Y be complex spaces. We say that the Oka principle holds for mappings from X to Y if every continuous map $f_0 : X \to Y$ is homotopic to a holomorphic mapping $f : X \to Y$. We shall denote by $[X : Y]$ the set of homotopy classes of continuous maps from X to Y. $\square$

The complete list of pairs of Riemann surfaces (open or otherwise) for which the Oka principle holds can be found in Winkelmann [49]:

Theorem 5.6a. *Let X and Y be Riemann surfaces. Then Oka's principle for mappings from X to Y holds if and only if (X, Y) is in the list below*:

(a) *$X = \mathbf{P}^1, Y \not\cong \mathbf{P}^1$(every continuous map is homotopic to a constant),*

(b) *X non-compact, $Y = \mathbf{P}^1$(every continuous map is homotopic to a constant),*

(c) *either X or Y,*

(d) *X is non-compact, $Y = \mathbf{C}^*$(there is a surjection $\rho : H^0(X, \mathcal{O}^*) \to H^1(X, \mathbf{Z}) = \mathrm{Hom}(\pi_1(X), \mathbf{Z}) = \mathrm{Hom}(\pi_1(M), \pi_1(\mathbf{C}^*)) = [X : \mathbf{C}^*]$),*

(e) *X is non-compact, Y is a torus,*

(f) *$X = \overline{X} \setminus \cup_i \overline{D}_i$ where $\overline{X}$ is a compact Riemann surface and $\{D_i \mid i = 1, .., k\}$ are disjoint discs of positive radii, Y is punctured unit disc Δ^*.*

It is convenient to list all pairs of Riemann surface for which the Oka Principle *fails*.

Theorem 5.6b. *Let X and Y be Riemann surfaces then Oka's principle for mappings from X to Y fails if and only if (X, Y) is in the list below*:

(a) *X is compact and $Y = \mathbf{P}^1$,*

(b) *X is compact and neither X nor Y is simply connected,*

(c) *X is non-compact and not simply connected, Y is not a torus nor any in the list $\mathbf{P}^1, \mathbf{C}, \mathbf{C}^*, \Delta, \Delta^*$ where Δ is the unit disc and $\Delta^* = \Delta \setminus \{0\}$,*

(d) *$X = \overline{X} \setminus \{\text{one point}\}$ where $\overline{X}$ is compact, $Y = \Delta^*$,*

(e) *$H_1(X)$ is not finitely generated and $Y = \Delta^*$.*

6 Complete intersections

Definition 6.1. Let R be a commutative ring. (We consider only rings with 1.) The *dimension* of R is by definition the supremum of the lengths n of all prime ideal chains:

$$\mathfrak{p}_0 \subsetneq \mathfrak{p}_1 \subsetneq \cdots \subsetneq \mathfrak{p}_n.$$

The *height*, $h(\mathfrak{p})$, of a prime ideal is the *supremum* of all the lengths of prime ideal chains terminating at $\mathfrak{p}$ ($\mathfrak{p}_n = \mathfrak{p}$ in the chain above). For an ideal $I \subsetneq R$ the *height* of I, $h(I)$, is the infimum of the heights of prime ideals containing I. For an ideal I, denote by $\mu(I)$ the infimum of the number of generators (over R) of I. $\square$

For a commutative Noetherian ring R it is clear that

$$h(I) \leq \mu(I) < \infty$$

for all ideal I.

Definition 6.2. Let $I \neq R$ be an ideal in a commutative Noetherian ring R. Then I is said to be an *ideal theoretic complete intersection* (or simply a *complete intersection*) in R if $h(I) = \mu(I)$. It is said to be a *set theoretic complete intersection* if there exists $f_1, ..., f_r$ such that $\mathrm{rad}(I) = \mathrm{rad}(f_1, ..., f_r)$ with $r = h(I)$. It is said to be a *local complete intersection* if for all maximal ideal $\mathfrak{m}$ the localization $I_{\mathfrak{m}}$ is an ideal theoretic complete intersection in $R_{\mathfrak{m}}$. $\square$

The corresponding geometric version on varieties of these concepts are defined in terms of the rings of functions:

Definition 6.3. Let X be a complex space and $Y \subset X$ a complex subspace. Denote by $I(Y) := H0(Y, \mathcal{I}_Y)$ (where $\mathcal{I}_Y$ is the ideal sheaf of Y in X) and $R = \mathcal{O}(X) := H^0(X, \mathcal{O}_X)$. Then Y is said to be an *ideal theoretic complete intersection*, or simply a *complete intersection* in X (resp. a *set theoretical complete intersection*; resp. a *local complete intersection*) if $I(Y)$ is an ideal theoretic complete intersection (resp. a *set theoretical complete intersection*; resp. a *local complete intersection*) in R. $\square$

An excellent reference for this section is the article by Schneider [45].

Remark 6.4. (i) A subspace Y is said to be a *local complete intersection at a point $y \in Y$* if the ideal $\mathcal{I}_{Y,y}$ is a local complete intersection in $\mathcal{O}_{X,y}$. The condition that the ideal $I = I(Y)$ is a local complete intersection in $R = \mathcal{O}(X)$ in the sense of Definition 6.3 is equivalent to the condition that $\mathcal{I}_{Y,y}$ is a local complete intersection in $\mathcal{O}_{X,y}$ for all $y \in Y$.

(ii) A family of functions $f_1, ..., f_r$ on X is said to define a complex subspace Y in X if $Y = \{x \in X \mid f_1(x) = \cdots = f_r(x)\}$. Define

$$\mu(Y) = \min\{r \mid \text{ there exists } r \text{ holomorphic functions on } X \text{ defining } Y\}. \qquad (6.1)$$

then

$$\operatorname{codim} Y \le \mu(Y) \le \mu(I(Y)). \tag{6.2}$$

In fact

$$Y \text{ is a complete intersection if and only if } \mu(Y) = \operatorname{codim} Y. \tag{6.3}$$

In other words, if $r = \operatorname{codim} Y$ then Y is a complete intersection if and only if there exists r holomorphic functions $f_1, ..., f_r$ on X such that the germs of these functions at y generate the ideal $\mathcal{I}_{Y,y}$ (over $\mathcal{O}_{X,y}$) for all $y \in Y$. For a local complete intersection we only require that the generators are *locally defined near y and may vary with y.* Analogously, we also have

$$Y \text{ is an ideal complete intersection if and only if } \mu(I(Y)) = \operatorname{codim} Y. \tag{6.4}$$

(iii) If X is affine algebraic and Y an affine algebraic subspace then the preceding concepts can also be formulated in the algebraic category (and one can replace C by an algebraically closed field). In this case the ideal $I(Y)$ is replaced by $I_{\text{alg}}(Y) = \{$algebraic functions on X vanishing on $Y\}$, $\mathcal{O}(X)$ is replaced by $\mathcal{O}_{\text{alg}}(X) = \{$algebraic functions on $X\}$ and the generators are required to be algebraic. If the underlying field is C then an affine algebraic variety over C is Stein. The numbers $\mu_{\text{alg}}(Y)$ and $\mu_{\text{alg}}(I(Y))$ are analogously defined. The assertions (6.2), (6.3) and (6.4) are also valid in the algebraic category. This case will be further discussed in section 8 below concerning the algebraic version of the Oka Principle.

(iv) It is clear that, in either the holomorphic or algebraic category, an ideal complete intersection is a local complete intersection as well as a set theoretic complete intersection.

(v) A smooth subvariety is a local complete intersection.

(vi) Forster showed that, if $Y \subset \mathrm{A}^n$ is an affine algebraic subspace then $\mu_{\text{alg}}(Y) \le n + 1$. Kumar [33] showed later that $\mu_{\text{alg}}(Y) \le n$ if Y is a local complete intersection.

(vii) Kumar [34] showed that affine algebraic locally complete intersection curves in $\mathbf{A}^n$ are set theoretic (algebraic) complete intersections.

$(viii)$ Affine algebraic subspaces of pure dimension $n - 1$ (every irreducible components is of dimension $n - 1$) of $\mathbf{A}^n$ are algebraic complete intersections. It is clear that, in general, affine curves are not set theoretic algebraic complete intersections (just take any curve that is not a local complete intersection). There are also smooth affine curves that are not (ideal theoretic) algebraic complete intersections. An excellent source in the algebraic case is Kunz [35].

(ix) Let $\mathcal{I}_Y$ be the ideal sheaf of Y. The quotient $\mathcal{I}_Y/\mathcal{I}_Y^2$ is called the *conormal sheaf.* It is locally free (in which case it is referred to as the *conormal bundle* and its dual, the normal bundle) if Y is a local complete intersection. Boratynski [4] showed that if $Y \subset \mathrm{A}^n$ is an affine algebraic local complete intersection with trivial normal bundle then Y is a set theoretic (algebraic) complete intersection. □

As far as I know the analogue of Boratynski's result in the Stein (transcendental) case is still open:

Open Problem 6.5. Is every closed complex subspace Y in C^n which is a local complete intersection and with trivial normal bundle a complete intersection?

For a general coherent sheaf $\mathcal{S}$ the number of generators $\mu(\mathcal{S}(X))$ (where $\mathcal{S}(X)$ is the space of global sections of $\mathcal{S}$) may not be finite. It is finite if

$$d = \sup_{x\in X} \dim \mathcal{S}_x/\mathfrak{m}_x\mathcal{S}_x$$

is finite. In fact we have the following result due to Schneider [43]:

Theorem 6.6. *Let X be a pure n-dimensional Stein space and $\mathcal{S}$ a coherent analytic sheaf over X. For $k \geq 1$ let $S_k(\mathcal{S}) = \{x \in X \mid \dim \mathcal{S}_x/\mathfrak{m}_x\mathcal{S}_x \geq k\}$. Then the $\mathcal{O}(X)$-module $\mathcal{S}(X) = H0(X, \mathcal{S})$ is finitely generated if and only if*

$$r = \sup_{S_k(\mathcal{S})\setminus S_{k+1}(\mathcal{S})\neq\emptyset} \left[\frac{1}{2}\dim S_k(\mathcal{S})\right] + k < \infty$$

and, in which case, $\mathcal{S}(X)$ is generated by r elements.

If $\mathcal{S}$ is a vector bundle the result can be strengthened by using Theorem 5.3:

Corollary 6.7. *Let E be a holomorphic vector bundle of rank r over a Stein space of dimension n then $\mu(E) \leq r + [n/2]$. If n is even then $\mu(E) \leq r + [n/2] - 1$ if and only if the Serge class $s_{n/2} = 0$.*

The total Serge class is by definition $s(E) = c(E)^{-1}$.

Remark 6.8. (i) Let E be a holomorphic vector bundle over a complex manifold Y (which may be identified with the zero section of E). Then E is holomorphically isomorphic to the normal bundle $N_{Y|E}$ of Y in E.

(ii) Let Y be a complex submanifold of a Stein manifold X. Then there exist an open neighborhood U of the zero section 0_N in the normal bundle $N_{Y|X}$, an open neighborhood V of Y in X and a biholomorphic map $\phi : U \to V$ such that $\phi(0_N) = Y$.

(iii) Let Y be a complex submanifold of a Stein manifold X. If Y is a complete intersection then the normal bundle $N_{Y|X}$ is trivial.

(iv) Let E be a holomorphic vector bundle over a Stein manifold Y. Then Y (identified as the zero section of E) is a complete intersection in E if and only if E is holomorphically trivial. □

Definition 6.9. A subspace Y in X is said to be an *almost complete intersection* if $\mu(I(Y)) \leq \operatorname{codim} Y + 1$ (resp. $\mu_{\text{alg}}(I(Y)) \leq \operatorname{codim} Y + 1$). □

The preceding definition is motivated by the Lemma of Serre:

Theorem 6.10. *Let I be a finitely generated ideal of a commutative ring. Then*

$$\mu(I) \leq \mu(I/I2) + 1.$$

In particular, if Y is an affine algebraic subspace of A^n then

$$\mu_{\text{alg}}(\mathcal{I}_Y) \leq \mu_{\text{alg}}(\mathcal{I}_Y/\mathcal{I}_Y^2) + 1.$$

The preceding theorem is valid on Stein spaces:

Theorem 6.11. *Let $\mathcal{I}$ be a finitely generated coherent sheaf in a Stein X then*

$$\mu(\mathcal{I}) \leq \mu(\mathcal{I}/\mathcal{I}^2) + 1.$$

In particular, this is true for $\mathcal{I} = \mathcal{I}_Y$ = the ideal sheaf of a closed complex subspace Y in X. Consequently, if Y is a local complete intersection then

$$\mu(\mathcal{I}_Y) \leq \mu(N_Y) + 1.$$

If, in addition the normal bundle N_y is trivial then Y is an almost complete intersection.

Remark 6.12. See Remark 6.4 (ix) for the last assertion. One can obviously say a little more, namely, if N_Y is trivial then $\mu(N_Y) = \text{rank } N_Y = \operatorname{codim} Y$, hence

$$\operatorname{codim} Y \leq \mu(\mathcal{I}_Y) \leq \operatorname{codim} Y + 1.$$

Thus the main question reduces to: "when is $\mu(\mathcal{I}_Y) = \mu(N_Y)$" or $\mu(\mathcal{I}_Y) = \mu(\mathcal{I}/\mathcal{I}^2)$?

Thus, for this particular problem, we are not yet able to recover the analogue of the algebraic result. There are, however, other cases for which we have stronger results in the Stein case than in the algebraic case. We begin with some technical results concerning the removal of intersections. The formulation below is due to Gromov and a more detailed version can be found in Forsternic. Suppose that $f : X \to Y$ is a holomorphic map between complex spaces and A is a closed complex subvariety of Y. If $f^{-1}(A)$ is the union of two disjoint complex subvarieties X_0 and X_1 in X. We would like to find conditions so that there is a homotopy $f_t : X \to Y, t \in [0,1]$, of continuous maps such that $f_0 = f, f_t|X_0 = f|X_0$ for all $t \in [0,1]$ and $f_1^{-1}(A) = X_0$. We have the following version of Oka principle for removing intersections.

Theorem 6.13. (Forsternic) *Let $f : X \to Y$ be a holomorphic map where X is Stein and Y is subelliptic. Let A be a complex subvariety of Y such that $Y \setminus A$ is subelliptic. Assume that $f^{-1}(A) = X_0 \cup X_1$ where X_0 and X_1 are disjoint complex subvarieties in X. Assume that there is a homotopy $\tilde{f}_t : X \to Y, t \in [0,1]$, of continuous maps such that $\tilde{f}_0 = f, \tilde{f}_1^{-1}(A) = X_0$ and $\tilde{f}_t|U = f|U$ for all $t \in [0,1]$ and for some fixed open neighborhood U of X_0. Then, for any positive integer r, there is a homotopy $f_t : X \to Y, t \in [0,1]$, of holomorphic maps such that $f_0 = f, f_1^{-1}(A) = X_0$ and f_t agrees with f on X_0 up to order r for all $t \in [0,1]$.*

This implies the Oka principle for local complete intersections of Forster and Ramspott:

Corollary 6.14. *Let X be a Stein manifold. Let $Y \subset X$ be a closed complex submanifold of codimension d which is a complete intersection in an open neighborhood U of Y. Let $f_1, ..., f_d$ be holomorphic functions on U such that $Y = \{x \in U \mid f_1(x) = ... = f_d(x) = 0\}$. If these local defining functions admit extension to global continuous functions $\tilde{f}_0, ..., \tilde{f}_d$ on X such that $Y = \{x \in X \mid \tilde{f}_1(x) = \cdots = \tilde{f}_d(x) = 0\}$ then Y is a holomorphic complete intersection in X, i.e., there exists holomorphic functions $F_1, ..., F_d$ on X such that $Y = \{x \in X \mid F_1(x) = \cdots = F_d(x) = 0\}$.*

Theorem 6.15. *Let X be an n-dimensional Stein manifold and Y be a codimension k closed complex submanifold. Then*

$$\mu(I(Y)) \leq \begin{cases} [(n+k)/2] & \text{if } k \geq 2, \\ 1 + [n/2] & \text{if } k = 1. \end{cases}$$

By Oka's principle the problem is reduced to counting continuous generators. The next two corollaries are immediate consequence of the preceding theorem:

Corollary 6.16. *If Y is a closed set of discrete points in a Stein manifold X then Y is a complete intersection.*

Corollary 6.17. *If Y is a non-singular curve in a Stein manifold X then Y is a complete intersection.*

These results can be extended (see Forster-Ramspott [12], [13]).

Theorem 6.18. *let Y be a closed submanifold of a Stein manifold X. Assume that the normal bundle of Y in X is trivial. Then Y is a complete intersection if*

$$H^{q+1}(X, Y, \pi_q(S^{2k-1})) = 0$$

for all $q \geq 2k - 1$.

Theorem 6.19. *let Y be a closed submanifold of pure dimension in a connected Stein manifold X such that* $\dim Y < \dim X/2$. *Then Y is a complete intersection if the normal bundle of Y in X is trivial.*

The last result can be improved if $X = \mathrm{C}^n$.

Theorem 6.20. *let Y be a closed submanifold of pure dimension in $\mathbf{C}^n$ such that* $\dim Y < 3(n-1)/2$.*Then Y is a complete intersection if the normal bundle of Y in X is trivial.*

In fact the preceding is valid for any contractible Stein space X and Y a local complete intersection.

Theorem 6.21. *Let Y be a k-dimension closed subspace of a n-dimension Stein space. Assume that Y is a local complete intersection and the conormal bundle is trivial. Then Y is a complete intersection if any one of the following conditions is satisfied*:

(i) $2k < n$;

(ii) *X is contractible and $3k \leq 2(n-1)$;*

(iii) *$2k = n$ and the dual class of Y in $H^{2k}(X, \mathrm{Z})$ vanishes.*

7 Embedding dimensions of Stein spaces

Let X be a Stein space of complex dimension n. The embedding dimension m_x of a point $x \in X$ is the smallest integer m such that there is an open neighborhood U of x which is biholomorphic to a closed analytic subset of the unit ball B^m in C^m. The *global embedding dimension of X* is by definition:

$$m_X = \mathrm{embdim} X = \sup_{x \in X} m_x.$$

If X is non-singular then $m_X = n$. The classical theorem of Remmert, Narisimhan, Bishop, Wiegmann [48] asserts that:

Theorem 7.1. (Embedding Theorem) *Let X be a Stein space of complex dimension n for which the local embedding dimension m is finite. Then there exists a proper embedding of*

X into $\mathbf{C}^N$ with $N = \max\{2n+1, n+m\}$. In fact the proper embeddings are dense in $\mathcal{O}_(X)^N$ where the space of holomorphic functions $\mathcal{O}_(X)$ is equipped with the topology of uniform convergence on compact sets.

Remark 7.2. If X is non-singular or more generally, reduced, then the embedding dimension N can be taken to be $2n+1$. Forster conjectured that, if X is smooth then N should be $n+[n/2]+1$. The following example (due to Forster shows that the conjectured number is sharp. Let Y be the complement of a quadric in $\mathbf{P}^2$, i.e.,

$$Y = \{(x, y, z) \in \mathbf{P}^2 \mid x^2 + y^2 + z^2 \neq 0\}$$

and

$$X = \begin{cases} Y^m, & \text{if } n = 2m, \\ qY^{[m]} \times \mathbf{C}, & \text{if } n = 2m+1. \end{cases}$$

Then X is a n-dimensional Stein manifold which can be embedded in $\mathbf{C}^{n+[n/2]+1}$ but not in $\mathbf{C}^{n+[n/2]}$. □

We shall always assume that the embedding dimension m_X is finite. For an integer $k \geq 0$ the set

$$X_k = \{x \in X \mid m_x \geq k\}$$

is an analytic subset. Let $n_k = \dim_{\mathrm{C}} X_k$ and define the invariant

$$q_X = \sup_{k \in \mathrm{N}, k \leq m_X} \{k + \left[\frac{n_k}{2}\right]\}.$$

It is known that

$$n + [n/2] \leq q_X \leq \max\{n + [n/2], m_X + \dim S(\mathcal{O}_X)/2\}$$

where $S(\mathcal{O}_X)$ is the singular locus of the structure sheaf $\mathcal{O}_X$.

Remark 7.3. It is well-known that

$$\dim S(\mathcal{O}_X) \begin{cases} \leq \dim X - 1, & \text{if } X \text{ is reduced}, \\ \leq \dim X - 2, & \text{if } X \text{ is normal}, \\ = 0, & \text{if } X \text{ has only isolated singularities}, \\ = -\infty, & \text{if } X \text{ is smooth}. \end{cases}$$

□

In general (singularities permitted) Schürmann [46] conjectured that

Conjecture 7.4. *Any Stein space X of complex dimension n and of bounded embedding dimension m_X can be properly holomorphically embedded in C^N with*

$$N = \max\{n + [n/2] + 1, q_X\}.$$

The conjecture reduces to that of Forster if X is smooth. The conjecture is solved, except for one case, by [Schürmann 46]:

Theorem 7.5. *Any Stein space X of complex dimension n and of bounded embedding dimension m_X can be properly holomorphically embedded in C^N with*

$$N = \max\{n + [n/2] + 1, q_X, 3\}.$$

Remark 7.6. The only case that is still unsettled (due to the extra entry 3 in N above) is whether a Stein curve X with $m_X \leq 2$ can be embedded in $\mathbf{C}^2$. For non-singular Stein curves (= open Riemann surfaces) the following are known to admit proper embeddings in C2:

(a) the unit disc $\Delta = \{z \in \mathbf{C} \mid |z| < 1\}$,

(b) all annuli $A = \{z \in \mathbf{C} \mid 1 < |z| < r\}$,

(c) the punctured unit disc $\Delta^* = \{z \in \mathbf{C} \mid 0 < |z| < 1\}$,

(d) (Globevnik-Stensones) all bounded finitely connected domain D in $\mathbf{C}$ whose boundary contains no isolated points. □

The proof of Theorem 7.5 is based, as usual, on (1) the Runge Approximation of the Oka Principle (see Theorem 3.6 and Theorem 4.3) and (2) the topology of a Stein space. The improvement over the previous work is to construct a *special* almost proper map. (A continuous map $f : X \to Y$ between two locally compact topological spaces is said to be *almost proper* if, for every compact set $K \subset Y$, each connected component of $f^{-1}(K)$ is compact.) The idea of using almost proper rather than proper directly was already used by Forster, Eliasberg-Gromov. Schürmann's construction is superior in that he used the subtler version (Lefschetz's Theorem for Stein spaces, due to Hamm) of the topology of a Stein space given by Theorem 1.8. This version, which we now describe, is more technical.

Let X be a reduced Stein space and let $f : X \to B(R) = \{z \in \mathrm{C}^n \mid ||z|| < R\}$ be a *finite* holomorphic map. Denote by $Z = f(X)$ the image of f. Let Z' be a closed analytic subset of Z and $X' = f^{-1}(Z')$. Assume that

(i) $Z \setminus Z'$ and $X \setminus X'$ are smooth and non-empty,

(ii) $f|_{Z\setminus Z'} : Z \setminus Z' \to X \setminus X'$ is an immersion.

Let Φ be the restriction of $||z||^2$ to Z and denote by $Z_r = \{z \in Z \mid ||z|| < r\} \cup Z'$ and $X_r = f^{-1}(Z_r)$.

Theorem 7.7. *Let $s < t$ be regular values of $\Phi|_{Z\setminus Z'}$. For any open set U_s such that $Z_s \subset U_s \subset Z_t$. Then there exist a continuous function $\phi : U_t \to [0, \infty)$ defined on some open neighborhood U_t of Z_t and a constant $c > 0$ such that*

(1) $Z_s \subset Z_{\phi,c} = \{z \in Z_t \mid \phi(z) < c\} \subset U_s$,

(2) *ϕ is of class $\mathcal{C}^\infty$ on $U \setminus Z_s$,*

(3) *ϕ is strictly plurisubharmonic on $U \setminus Z_{\phi,c}$,*

(4) *$\phi|_{Z_t\setminus Z_s}$ is a Morse function of $Z_i \setminus Z_s$ and c is a regular value.*

The following corollary (known as the Lefschetz Theorem for Stein space) is immediate:

Corollary 7.8. *Let $Z_{\phi,c}$ be as in the preceding theorem. Then $Z_t \setminus Z_s$ is obtained from $Z_{\phi,c} \setminus Z_s$ be adjoining a cell of real dimension $\leq \dim_{\mathbf{C}}(Z \setminus Z')$.*

The assumption that f is an immersion implies that the preceding theorem is valid also for the function $\phi' = \phi \circ f$ on X. Then $X_{\phi',c} = f^{-1}Z_{\phi,t} = \{x \in X_t \mid \phi'(x) < c\}$ where $X_t = f^{-1}(Z_t)$. Thus $X_t \setminus X_s$ is obtained from $X_{\phi,c} \setminus X_s$ be adjoining a cell of real dimension $\leq \dim_{\mathbf{C}}(X \setminus X')$.

The problem of proper map is quite difficult. We have:

Open Problem 7.9. The Oka principle for proper embedding of Stein manifolds is open. Namely, suppose that X is a Stein manifold and $f : X \to \mathbf{C}^N$ is a proper embedding of class $\mathcal{C}^\infty$. Is f homotopic to a proper holomorphic embedding?

In contrast the Oka principle for holomorphic immersion is known. The following results are due to Eliashberg and Gromov and Gromov [24]:

Theorem 7.10. *Every Stein manifold of complex dimension n admits an immersion into $\mathbf{C}^{[3n/2]}$.*

Theorem 7.11. *Let X be a Stein manifold of complex dimension n. Let $q > n$ be an integer. Then there is an immersion $f : X \to \mathbf{C}^q$ if and only if the cotangent T^*X is spanned by q global continuous 1-forms.*

Open Problem 7.12. The case $q = n$, however, remains open.

There is also a result concerning relative proper embeddings (or extension of embeddings):

Theorem 7.13. *Let Y be a closed submanifold of an n-dimensional Stein manifold. Let $f : Y \to \mathbf{C}^N, N \geq 2n+1$ be a proper embedding then there exists a proper embedding $\tilde{f} : X \to \mathbf{C}^N$ such that $\tilde{f}|Y = f$.*

There is also results on embeddings with interpolations:

Theorem 7.14. *Let X be a Stein manifold and suppose that there is a proper embedding of X in $\mathbf{C}^N, N > 1$. Then for any discrete subset Γ in C^N there exists an embedding $f : X \to \mathbf{C}^N$ such that $\Gamma \subset f(X)$. Moreover, the map f can be chosen so that for any holomorphic map $\phi : \mathbf{C}^d \to \mathbf{C}^N, d = N - \dim X$, of maximal rank the intersection $\phi(\mathbf{C}^r) \cap f(X)$ is an infinite set. If $d = 1$ then f can be chosen so that $\mathbf{C}^N \setminus f(X)$ is Kobayashi hyperbolic.*

For convenience we say that a space Y is d-hyperbolic if there does not exist any holomorphic map $\phi : \mathbf{C}^d \to Y$ of rank r. Since C^n can be properly embedded in $\mathbf{C}^N$ for any $N > n$ the theorem above implies the existence of a proper embedding $f : \mathbf{C}^n \to \mathbf{C}^{n+d}$ such that $\mathbf{C}^{n+d} \setminus f(\mathbf{C}^n)$ is d-hyperbolic. On the other hand, a result of Abyankar and Moh asserts that every polynomial embedding of $\mathbf{C}$ in $\mathbf{C}^2$ is equivalent to a linear embedding hence the complement of any such embedding is biholomorphic to $\mathbf{C} \times \mathbf{C}^*$.

Next we consider the problem of constructing holomorphic functions on a Stein manifold without critical points (note that it is clear that the set of critical points of a generic holomorphic function on a Stein manifold is a discrete set). The problem was first studied by Gunning and Narasimhan in the case of Riemann surfaces and by Forsternic:

Theorem 7.15. *Let X be a Stein manifold of complex dimension n then there exists $[(n+1)/2]$ holomorphic functions $f_1, ..., f_{[(n+1)/2]}$ on X with the property that the differentials $df_1, ..., df_{[(n+1)/2]}$ are linearly independent over $\mathbf{C}$ at any point $x \in X$.*

Gunning and Narasimhan's construction in the case of a Riemann surface X is based on the fact that T^*X is trivial hence there exists a global holomorphic non-vanishing differential 1-form ω. (The higher dimensional analogue is the result of Ramspott that, every holomorphic vector bundles on a Stein manifold admits $[(n+1)/2]$ linearly independent (over C) global holomorphic sections.) The idea is to find an antiderivative for ω, i.e., $\omega = df$ is exact for some global holomorphic function f. For this we need to find ω whose integral over every closed curve vanishes (Cauchy's condition). This is achieved by showing that there is a non-vanishing holomorphic function g on X such that $e^g\omega$ satisfies the Cauchy condition, i.e., $e^g\omega = d\gamma$ where γ is differentiable. This is done first by adjusting ω by continuous functions on an exhaustion of X be relative compact open Runge (holomorphically convex) domains and the appropriate global holomorphic function is constructed via Runge's approximation theorem.

Let f be a holomorphic function defined on a complex manifold X the critical set of $f = \{x \in X \mid df_x = 0\}$ shall be denoted by $\mathrm{Crit}(f, X)$.

The following is the Oka principle for extensions preserving critical points.

Theorem 7.16. *Let X be a Stein manifold and $Y \subset X$ a closed complex submanifold. Let f be a holomorphic function on some open neighborhood Y such that the critical set* $\mathrm{Crit}(f, U)$ *is discrete and is contained in Y. For any non-negative integer and any holomorphic function f on Y there is a holomorphic function $\tilde{f}$ on X such that f and $\tilde{f}$ agrees on Y up to order r and that* $\mathrm{Crit}(f, U) = \mathrm{Crit}(\tilde{f}, X)$.

The following is Oka's homotopy principle for extensions of submersions.

Theorem 7.17. *Let X be a Stein manifold and Y a closed complex submanifold. Let $\theta^0 = (\theta_1^0, ..., \theta_q^0), 1 \leq q < \dim X$ be a q-tuple of continuous 1-form of type $(1,0)$ on X such that $\theta_1, ..., \theta_q$ are linearly independent at every point $x \in X$ and $\theta^0 = df^0 = (df_1^0, ..., df_q^0)$ where $f_1^0, .., f_q^0$ are holomorphic functions on an open neighborhood U of Y. Then there exist a homotopy θ^t of q-tuples of continuous forms of type $(1,0)$ on X such that $\theta^t = df^t$ is holomorphic and exact on U and that $\theta^1 = (df_1^1, ..., df_q^1)$ is globally exact holomorphic and $f^1 = (f_1^1, ..., f_q^q) : X \to \mathbf{C}^q$ is a holomorphic submersion.*

Recall that (Theorem 1.10 of Androetti-Frankel) every Stein manifold of complex dimension n admits a strictly pseudoconvex exhaustion function with Morse index at most n at each critical point. This implies that

Theorem 7.18. *Let X be a Stein manifold of complex dimension n admitting an exhaustion function with Morse index at most k at each critical point. Then X admits a holomorphic submersion onto $\mathbf{C}^{n_k}$ where $n_k = \min\{n - [k/2], n-1\}$. In particular, every Stein manifold of complex dimension n admits a holomorphic submersion onto $\mathbf{C}^{[(n+1)/2}$.*

Gunning and Rossi's result is the case $\dim X = q = 1$. Note that the case $\dim X = q \geq 2$ is open.

8 Oka principle with growth condition

Among the complex subspaces of $\mathbf{C}^N$ there are those that are defined by polynomials (the common zeros of a finite number of polynomials). These are the affine algebraic subvarieties. Such a variety X can also be characterized by the condition that the closure (the term compactification and completion are also commonly used in the literature) $\overline{X}$ in $\mathbf{P}^N$ is a complex subspace (necessarily algebraic by Chow's theorem) of P^N.

In other words a Stein space X is affine algebraic if it admits a compactification $\overline{X}$ which is projective. It is well-known that a non-singular affine algebraic manifold X admits an embedding with the property that its closure $\overline{X}$ is a smooth projective variety and that $D = \overline{X} \setminus X$ is of simple normal crossings.

An n-dimensional complex space X is affine algebraic if and only if it is a *finite* ramified cover of $\mathbf{C}^n$, i.e., there exists a proper holomorphic surjection $\phi : X \to \mathbf{C}^n$ such that the preimage $\phi^{-1}(z)$ is a finite set of points in X.

A holomorphic coherent sheaf $\mathcal{S}$ on an affine variety X is said to be *algebraic* if can be extended to a holomorphic coherent sheaf on the projective variety $\overline{X}$. The condition is independent of the choices of the affine embeddings of X in $\mathbf{C}^N$. By Chow's theorem every holomorphic coherent sheaf over a projective variety is algebraic. However this is not true for coherent sheaves over non-compact algebraic varieties.

In general there are non-extendible continuous principle bundles on an affine algebraic variety X and such bundles cannot admit an algebraic structure. Denote, respectively, by $\mathcal{G}_{\mathrm{cont}}, \mathcal{G}_{\mathrm{hol}}$ and $\mathcal{G}_{\mathrm{alg}}$ the sheaf of germs of continuous, holomorphic and algebraic principle G-bundles. We have, in general:

$$H^1(X, \mathcal{G}_{\mathrm{cont}}) = H^1(X, \mathcal{G}_{\mathrm{hol}}) \neq H^1(X, \mathcal{G}_{\mathrm{alg}}).$$

(However, it is true that every algebraic principal G-bundle on C^n is algebraically trivial.) Thus the Grauert-Oka principle is not valid in the algebraic category.

In 1974 Cornalba-Griffiths [9] proposed a *finite order* Grauert-Oka principle. The concept of a holomorphic function of finite order defined on $\mathbf{C}^N$ is well known, namely a holomorphic function is of finite order if its logarithmic maximum modulus is of polynomial growth:

$$M_f(r) = \max_{||z||<r} \log |f(z)| \leq Ar^k + B$$

where k, A, B are non-negative constants. The definition above makes perfect sense for any subvariety X of $\mathbf{C}^N$ (in the definition above the maximum is taken over all $x \in X$ such that $||Z|| < r$) hence the concept holomorphic function of finite order is defined on all affine varieties (even though the definition makes sense for any Stein spaces but in general there may not be any non-trivial holomorphic functions satisfying the growth condition). The product and sum of functions of finite order are of finite order hence it has a natural structure of a ring. Note that polynomials on an affine variety are precisely those holomorphic functions satisfying the condition

$$M_f(r) = \max_{z \in X, ||z||<r} \log |f(z)| \leq A \log r + B.$$

Thus the ring of polynomials is a subring of the ring of functions of finite order.

For a non-singular affine variety the concept of finite order can be localized via open sets of its closure $\overline{X}$. Namely, on an open set U in $\overline{X}$, a holomorphic function defined on $U \setminus D$ ($D = \overline{X} \setminus X$) is said to be of finite order if

$$M_f(r) = \max_{z \in U \setminus D, ||z||<r} \log |f(z)| \leq Ar^k + B.$$

Denote by $\Gamma_{\text{f.o.}}(U)$ = holomorphic functions of finite order defined on $U \setminus D$ then the sheaf associated to the presheaf

$$U \mapsto \Gamma_{\text{f.o.}}(U), \quad U \text{ open set in } \overline{X}$$

is called the sheaf of germs of holomorphic functions of finite order on X and is denoted by $\mathcal{O}_{\text{f.o.}}$. It is important to keep in mind that $\mathcal{O}_{\text{f.o.}}$ is a sheaf on $\overline{X}$ containing $\mathcal{O}_{\overline{X}}$ as a subsheaf. On the other hand, we have $\mathcal{O}_{\text{f.o.}}|_X = \mathcal{O}_{\overline{X}}|_X = \mathcal{O}_X$. The main fact concerning the sheaf $\mathcal{O}_{\text{f.o.}}$ is the following result of Muflur-Vitter-Wong [40]:

Theorem 8.1. *Let X be an affine algebraic variety with completion $\overline{X}$. Then $\mathcal{O}_{\text{f.o.}}$ is flat over $\mathcal{O}_{\overline{X}}$, namely, for any short exact sequence of $\mathcal{O}_{\overline{X}}$-sheaves:*

$$0 \to \mathcal{A} \to \mathcal{B} \to \mathcal{C} \to 0$$

the sequence

$$0 \to \mathcal{A} \otimes \mathcal{O}_{\text{f.o.}} \to \mathcal{B} \otimes \mathcal{O}_{\text{f.o.}} \to \mathcal{C} \otimes \mathcal{O}_{\text{f.o.}} \to 0$$

is also exact.

This implies the following vanishing theorem:

Corollary 8.2. *Let X be an affine algebraic variety with completion $\overline{X}$. Then, for any coherent sheaf $\mathcal{S}$ on $\overline{X}$,*

$$H^q(\overline{X}, \mathcal{S} \otimes \mathcal{O}_{\text{f.o.}}) = 0$$

for all $q \geq 1$.

The topology of an affine manifold can also be computed using the finite order forms. Recall that on a smooth affine variety X we have a complex:

$$d_p : H^0(X, \mathcal{A}_p) \to H^0(X, \mathcal{A}_{p+1})$$

where $\mathcal{A}_p$ is the sheaf of germs of p-forms of class $\mathcal{C}^\infty$, resulting in the deRham isomorphism ($H^p_d(X, \mathcal{A}_*) = \ker d_p / \operatorname{im} d_{p-1}$ for $p \geq 1$ and $H^0_d(X, \mathcal{A}_*) = \ker d_0$):

$$H^*_d(X, \mathcal{A}_*) \cong H^*(X, \mathrm{C}).$$

Analogously we have the complex:

$$\partial : H^0(X, \Omega_p) \to H^0(X, \Omega_{p+1})$$

where Ω_p is the sheaf of germs of holomorphic p-forms, resulting in the Serre isomorphism:

$$H^*_\partial(X, \Omega_*) \cong H^*(X, \mathbf{C}).$$

Denote by $(\Omega_{\text{alg}})_p$ the sheaf of germs of rational p-forms on $\overline{X}$ holomorphic on X and poles along $D = \overline{X} \setminus X$ then we have a complex:

$$\partial : H^0(X, (\Omega_{\text{alg}})_p) \to H^0(X, \Omega_{\text{alg}})_{p+1})$$

and the Atiyah-Hodge isomorphism:

$$H^*_{\partial}(X, \Omega_{\text{alg}}) \cong H^*(X, \mathbf{C}).$$

These isomorphisms hold also in the finite order category (see [M-V-W]). We get have from the complex

$$d : H^0(X, (\mathcal{A}_{\text{f.o.}})_p) \to H^0(X, (\mathcal{A}_{\text{f.o.}})_{p+1})$$

the finite order deRham isomorphism:

$$H^*_d(X, \mathcal{A}_{\text{f.o.}}) \cong H^*(X, \mathbf{C})$$

and from the complex

$$\partial : H^0(X, (\Omega_{\text{f.o.}})_p) \to H^0(X, \Omega_{\text{f.o.}})_{p+1})$$

the finite order Atiyah-Hodge-Serre isomorphism:

$$H^*_{\partial}(X, \Omega_{\text{f.o.}}) \cong H^*(X, \mathbf{C}).$$

We now introduce the main result of Cornalba-Griffiths. Let $\mathcal{O}^*_{\text{f.o.}}$ be the sheaf of germs of non-vanishing holomorphic functions of finite order. A rank r holomorphic vector bundle E over a smooth affine manifold X is said to admit a finite order structure if there exists an open cover $\{U_i\}$ of $\overline{X}$ such that $f_i : U_i \times \mathrm{C}^r \cong E|U_i$ is trivial ($f_i(x, \xi) = (x, g_i(x))$) and that the transition functions $f_j^{-1} \circ f_i(x, \xi) = (x, g_{ij}(x)\xi)$ where

$$g_{ij} : U_i \cap U_j \to GL(r, \mathrm{C})$$

is of finite order (i.e., every component is a function of finite order on $U_i \cap U_j$). For a line bundle this means that g_{ij} is a non-vanishing holomorphic functions of finite order. Thus $H1(\overline{X}, \mathcal{O}^*_{\text{f.o.}})$ is the group of isomorphism classes of holomorphic line bundles on X. A finite order line bundle is trivial in the sense of finite order if

$$g_{ij} = g_i g_j^{-1}$$

where g_i is finite order on U_i. The following Oka principle with growth condition is due to Cornalba-Griffiths [9]:

Theorem 8.3. *Let X be an affine manifold then*

$$H^1(\overline{X}, \mathcal{O}^*_{\text{f.o.}}) \cong H^1(X, \mathcal{C}^*_X) \cong H^2(X, \mathbf{Z}).$$

In particular, every topological line bundle admits a finite order structure and a finite order line bundle is trivial in the category of finite order bundles if and only it is topologically trivial.

Cornalba-Griffiths also proposed several definitions of finite order structure for vector bundles of rank ≥ 2 over affine manifolds. *We shall always assume that X is embedded in some C^N and $\overline{X}$ is smooth and $D = \overline{X} \setminus X$ is of simple normal crossings.* Let $\tau = ||z||^2|_X$ be the induced strictly plurisubharmonic exhaustion function. We shall also use the concept of "mixed holomorphic bisectional curvature" associated to a metric on a bundle E (see Cao-Wong [5] for details). If $E = TX$ is the tangent bundle equipped with a metric on X then this is just the usual concept of "holomorphic bisectional curvature". The first definition is to impose condition on the "mixed holomorphic bisectional curvature". This is natural because this is the curvature term that appears in the Bochner-Weitzenbock formula. Thus with control on this curvature it is possible to sole the $\bar{\partial}$-problem (for E-valued forms) with growth condition.

Definition I. A holomorphic vector bundle over an affine algebraic manifold X is said to be of algebraic (resp. *finite order*) if there exists a hermitian metric h on E such that the mixed bisectional curvature Θ (which in the case of a line bundle is simply the first Chern form of the metric) satisfies the condition

$$|\Theta| \leq C dd^c \tau$$

for some positive constant $C > 0$ (resp.

$$|\Theta| \leq C \tau^{\lambda} dd^c \tau$$

for some positive constant $C > 0$ and non-negative constant $\lambda \geq 0$).

The second definition is to impose a growth condition on the classifying map. Let ω_{FS} be the Kähler form (induced by the usual Fubini-Study metric on the complex projective space via the Plücker embedding) on the Grassmann manifold $Gr(r, N)$ of r-dimensional linear subspaces in $\mathbf{C}^N$. Let $f : X \to Gr(r, N)$ be a holomorphic map. The only known way of imposing growth condition on mappings is via the theory of Nevanlinna. The characteristic function (obtained via integrating the Chern form of the appropriated bundle, hence the terminology) of f is defined by

$$T_f(r) = \int_0^r \frac{dt}{t} \int_{\tau \leq r2} f^* \omega_{FS} \wedge (dd^c \log \tau)^{n-1}.$$

It is well-known (a theorem of Stoll) that the map f is algebraic if and only if $T_f(r) = O(\log r)$. In fact, for an algebraic map

$$\lim_{r \to \infty} \frac{T_f(r)}{\log r} = \deg f.$$

The map is said to be of finite order if and only if $T_f(r) \leq O(r^{\lambda})$ for some $\lambda \geq 0$. These definitions are standard in Nevanlinna theory.

Definition II. A holomorphic vector bundle (E, h) over X is said to be algebraic (resp. of finite order) if there exists an algebraic (resp. a holomorphic map of finite order) $f : X \to \mathrm{Gr}(r, N)$ for some m such that $E = f^* \mathcal{U}$ where $\mathcal{U}$ is the universal bundle over $\mathrm{Gr}(r, N)$.

The next definition impose condition on the transition function of the bundle relative to the special coverings of $\overline{X}$.

Definition III. A holomorphic vector bundle E of rank r on X is algebraic (resp. of finite order) if there exists an open cover $\{U_i\}$ of $\overline{X}$ such that $E|_{U_i \setminus D}$ is trivial and with transition functions

$$g_{ij} \in GL(r, \mathcal{O}_{\mathrm{alg}}\big((U_i \cap U_j) \setminus D\big) \ \ (\text{resp. } GL(r, \mathcal{O}_{\mathrm{f.o.}}\big((U_i \cap U_j) \setminus D\big))).$$

Another definition introduced by Cornalba-Griffiths is to impose condition on the Schubert cycles. Let Y be a subvariety of (pure) dimension k in X. The standard way of measuring the "volume" growth of subvarieties is to use Nevanlinna theory. The counting function of Y is defined by

$$N_Y(r) = \int_0^r \frac{dt}{t^{2k-1}} \int_{Y(r)} (dd^c\tau)^k$$

where $Y(r) = \{x \in Y \mid \tau(r) < r^2\}$. It is well-known that Y is algebraic if and only if $N_Y(r) = O(\log r)$. The subvariety Y is said to be of finite order if and only if $N_Y(r) \leq Cr^\lambda$. These definitions are standard in Nevanlinna theory.

Definition IV. A holomorphic vector bundle E of rank r on X is algebraic (resp. of order at most λ) if there exists global holomorphic sections $\sigma_1, ..., \sigma_r$ of E such that $Y_q = [\sigma_1 \wedge \cdots \wedge \sigma_{r-q+1} = 0]$ is a subvariety of codimension $\min\{q, n\}$ is algebraic (resp. of finite order) for $q = 1, ..., r$.

It is relatively easy to show that Definitions I-IV are equivalent for algebraic bundles of any rank and also for line bundles in the finite order case (see [9] for details). *The case of finite order bundles of rank* ≥ 2 *remains open.* There appears to be essential difficulties to work directly with vector bundles of higher rank.

One way of dealing with a vector bundle E of higher rank is to use the Lemma of Grothendieck/Serre:

Theorem 8.4. *Let E be a holomorphic vector bundle, of rank ≥ 2, over a complex manifold X. Then*

$$H^q(X, \odot^m E \otimes \mathcal{S}) \cong H^q(\mathbf{P}(E^\vee), \mathcal{L}^m \otimes p^*\mathcal{S})$$

for all $q \geq 0$ and any sheaf $\mathcal{S}$ on X where $\mathcal{L}$ is the Serre bundle on the projectivization $\mathbf{P}(E^\vee)$, $E^\vee$ is the dual, and $\odot^m E$ the m-fold symmetric product, of E.

The Serre line bundle on $\mathbf{P}(E^\vee)$ is the line bundle whose restriction to each of the fiber ($\cong \mathbf{P}^{r-1}, r = \operatorname{rank} E$) of the projection $p : \mathbf{P}(E^\vee) \to X$ is the hyperplane line bundle over $\mathbf{P}^{r-1}$. *The preceding theorem implies that, at the cohomology level, the investigation of the behavior of vector bundles is reduced to that of line bundles over the projectivization.* The standard technique is to find a good Hermitian metric on the line bundle and use $\bar{\partial}$-theory. Now a Hermetian metric h on the Serre line bundle defines canonically (and vice versa) a Finsler metric (by abuse of notations we use the same symbol h for both) on E and the first Chern form $c_1(\mathcal{L}, h)$ is positive-definite (resp. non-negative) if and only if the mixed holomorphic bisectional curvature of (E, h) is positive (resp. non-negative). The readers are referred to Cao-Wong [5] for details. This suggest that we should extend Definition I of Cornalba-Griffiths to Finsler metrics. Equivalently, by putting assumption on the Serre line bundle and work with line bundles over $\mathrm{P}(E^\vee)$. In the thesis of my student M. Maican [39] the Finsler formulation was used and in Wong [50] the Serre bundle approach was used (*moreover, for technical reason, an integral bound was used in place of the pointwise bound used in* [40] *and* [50]):

Definition V. A holomorphic vector bundle E of rank $r \geq 2$ over an affine algebraic manifold X is said to be of finite order if there exists a Hermitian metric h along the Serre line bundle $\mathcal{L}$ on $\mathbf{P}(E^\vee)$ such that the following conditions hold:

(i) $c_1(\mathcal{L}, h) > 0$ and $c_1(\mathcal{L}, h) \geq dd^c(p^*\tau)$ where $p : \mathbf{P}(E^\vee) \to X$ is the projection, τ is the exhaustion function on X used in Definition I and $dd^c(p^*\tau)$ is the Levi form of the exhaustion function $p^*\tau$ on $\mathrm{P}(E^\vee)$;

(ii) $T((\mathcal{L}, h), r) \leq O(r^\lambda)$ for some $\lambda \geq 0$ where

$$T((\mathcal{L}, h), r) = \int_0^r \frac{dt}{t} \int_{\{p^*\tau \leq r\}} \eta \wedge dd^c(dd^c \log p^*\tau)^{n-1} \wedge c_1(\mathcal{L}, h)^r$$

(iii) $c_1(\mathcal{L}, h) - \mathrm{Ric}\, c_1(\mathcal{L}, h) \geq \epsilon c_1(\mathcal{L}, h)$ for some positive constant ϵ.

It is also convenient to replace Definition III of Cornalba-Griffiths by:

Definition VI. A holomorphic vector bundle E, of rank $r \geq 2$, over a special affine algebraic manifold X is said to be of finite order if there exists an injective holomorphic immersion $F : \mathbf{P}(E^\vee) \to \mathbf{P}^N$ such that $F^*(\mathcal{O}_{\mathbf{P}^N}(1)) = \mathcal{L}_{\mathbf{P}(E^\vee)}$ and satisfying the following estimate

$$\int_0^r \frac{dt}{t} \int_{\{p^*\tau \leq t\}} (dd^c \log p^*\tau)^{n-1} \wedge F^*\omega_{FS}^r = O(r^\lambda)$$

where ω_{FS} is the Fubini-Study metric on P^N.

With these modifications it was shown in [50] that:

Theorem 8.5. *Let E be a holomorphic vector bundle over an affine algebraic manifold X of complex dimension n then Definitions I, V and VI are equivalent.*

In [39] it was shown, with Definition V formulated in terms of Finsler metric and a pointwise bound, that $V \implies VI \implies I$.

9 Oka's principle and the Moving Lemma in hyperbolic geometry

We present in this section an application of the Oka Principle in establishing a "Moving Lemma" which is very useful in the transcendental intersection theory and in hyperbolic geometry. A complex space X is said to be *Brody hyperbolic* if every holomorphic map $f : \mathbf{C} \to X$ is constant. For compact complex space Brody hyperbolicity is equivalent to Kobayashi hyperbolic. For non-compact space Kobayashi hyperbolicity implies Brody hyperbolicity but the converse is false in general. A weaker notion is that of weak hyperbolicity. A complex space is said to be *weakly hyperbolic* if the image of every holomorphic map $f : \mathrm{C} \to X$ is contained in a subspace Y of dimension strictly less than that of X. An even weaker version is that of *measure hyperbolicity*. A complex space X of complex dimension n is said to be measure hyperbolic if every holomorphic map $f : \mathrm{C}^n \to X$ is degenerate in the sense that the Jacobian determinant vanishes, i.e.,

$$\det \begin{pmatrix} \frac{\partial f_1}{\partial z^1} & . & . & . & \frac{\partial f_1}{\partial z^n} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ \frac{\partial f_n}{\partial z^1} & . & . & . & \frac{\partial f_n}{\partial z^n} \end{pmatrix} = 0.$$

Let X be a complex manifold of dimension n. We shall denote by J^kX the k-th jet bundle of X. These bundles are defined as follows (the readers are referred to [8], [18], [32] for further details). Let $\mathcal{H}_x, x \in X$, be the sheaf of germs of holomorphic curves:

$$\{f : \Delta_r \to X \text{ is holomorphic for some } r > 0 \text{ and } f(0) = x\}$$

where Δ_r is the disc of radius r in C. Define, for $k \in$ N, an equivalence relation by designating two elements $f, g \in \mathcal{H}_x$ as k-equivalent (written $f \sim_k g$) if

$$f_j^{(p)}(0) = g_j^{(p)}(0)$$

for all $1 \leq p \leq k$, where $f_j = z_j \circ f, z_1, ..., z_n$ are local holomorphic coordinates near x and $f_j^{(p)} = \partial^p f_j/\partial\zeta$ is the p-th order derivative relative to the variable $\zeta \in \Delta_r$. The sheaf of *parameterized k-jets* is defined by:

$$J^kX = \cup_{x\in X}\mathcal{H}_x/\sim_k\ . \tag{9.1}$$

Elements of J^kX will be denoted by $j^kf(0) = (f(0), f'(0), ..., f^{(k)}(0))$.

We set $J^0X = \mathcal{O}_X$. It is clear that $J^1X = TX$ but, in general, for $k \geq 2$, J^kX is not locally free. There is, however, a natural C*-action on J^kX defined via parametrization. Namely, for $\lambda \in \mathbf{C}^*$ and $f \in \mathcal{H}_x$ a map $f_\lambda \in \mathcal{H}_x$ is defined by $f_\lambda(t) = f(\lambda t)$. Then $j^kf_\lambda(0) = (f_\lambda(0), f_\lambda'(0), ..., f_\lambda^{(k)}(0)) = (f(0), \lambda f'(0), ..., \lambda^k f^{(k)}(0))$. So the $\mathbf{C}^*$-action is given by

$$\lambda \cdot j^kf(0) = (f(0), \lambda f'(0), ..., \lambda^k f^{(k)}(0)). \tag{9.2}$$

Note that even though J^kX is not a vector bundle but the zero section still makes sense.

For the tangent bundle TX we have the dual $T^*X = \Omega^1_X$ which is the sheaf associated to the presheaf

$$\Omega^1_U = \{\omega : TX|_U \to \mathbf{C} \text{ holomorphic } \mid \omega(\lambda \cdot j^1f) = \lambda\omega(j^1f), \lambda \in \mathrm{C}\}.$$

Analogously, we define for positive integers m, k, *the sheaf of germs of k-jet differentials of weight m*, denoted $\mathcal{J}_k^mX$, to be the sheaf associated to the presheaf

$$\mathcal{J}_k^mU = \{\omega : J^kX|_U \to \mathbf{C} \text{ holomorphic } \mid \omega(\lambda j^kf) = \lambda^m\omega(j^kf), \lambda \in \mathrm{C}\}. \tag{9.3}$$

Note that $\mathcal{J}_1 1X = T^*X = \Omega_X 1$. We also set $\mathcal{J}_0^mX = \mathcal{O}_X$ for all m.

We gather some basic properties of J^kX and $\mathcal{J}_k^mX$ in the next Proposition.

Proposition 9.1. *Let X and Y be complex manifolds and let $F : X \to Y$ be a holomorphic map.*

(a) *For any $\ell \leq k$ the map $p_{k\ell} : J^kX \to J^\ell X$ defined by*

$$p_{k\ell}(j^kf(0)) = j^\ell f(0)$$

is a well-defined C**-bundle map* (*the forgetting map*)*, i.e., $p_{k\ell}$ respects the* C**-action defined by* (9.2).

(b) *The k-th order induced map $J^kF : J^kX \to J^kY$, defined by*

$$J^kF(j^kf(0)) = j^k(F \circ f)(0)$$

is a well-defined C^**-bundle map.*

(c) *Given any holomorphic map $f : \Delta_r \to X$ $(0 < r \leq \infty)$, the map $($the k-th order lifting$)$ $j^kf : \Delta_{r/2} \to J^kX$ defined by*

$$j^kf(\zeta) = j^kg(0), \ \zeta \in \Delta_{r/2}$$

where $g(\xi) = f(\zeta + \xi)$ is holomorphic for $\xi \in \Delta_{r/2}$ and commutes with the projection $p_k : J^kX \to X$, i.e., $p_k \circ j^kf = f$.

(d) *The map $\delta : \mathcal{J}_k^mX \to \mathcal{J}_{k+1}^{m+1}X$ defined by $\delta f = df$ if $k = 0$ and for $k \geq 1$ and $\omega \in \mathcal{J}_k^mX$*

$$\delta\omega(j^{k+1}f) = (\omega(j^kf))^{'}$$

is a C^**-bundle map $($derivation$)$. Iteration yields a* C^**-bundle map*

$$\delta^k\phi(j^kf) = (\phi \circ f)^{(k)}.$$

(e) *For $\ell \leq k$ the natural projection $p_{kl} : J^kX \to J^lX$ induces an injection $($the dual forgetting map$)$ $p_{kl}^* : \mathcal{J}_l^mX \to \mathcal{J}_k^mX$ defined by "forgetting" those derivatives higher than l:*

$$p_{kl}^*\omega(j^kf) = \omega(p_{kl}(j^kf)) = \omega(j^lf), \ \ \omega \in \mathcal{J}_l^mX.$$

$($*We shall simply write $\omega(j^kf) = \omega(j^lf)$ if no confusion arises*$)$.

Example 9.2. A 1-jet differential is a differential 1-form $\omega = \sum_{i=1}^n a_i(z)dz_i$. Let $f = (f_1, ..., f_n) : \Delta_r \to X$ be a holomorphic map. Then

$$\omega(j^1f) = \sum_{i=1}^n a_i(f)dz_i(f^{'}) = \sum_{i=1}^n a_i(f)f_i^{'}$$

and $\delta\omega$ is a 2-jet differential of weight 2, given by

$$\delta\omega(j^2f) = (\omega(j^1f))^{'} = (\sum_{i=1}^n a_i(f)f_i^{'})^{'} = \sum_{i,j=1}^n \frac{\partial a_i}{\partial z_j}(f)f_i^{'}f_j^{'} + \sum_{i=1}^n a_i(f)f_i^{''}.$$

Analogously, $\delta2\omega$ is a 3-jet differential of weight 3, given by

$$\delta^2\omega(j^3f) = \sum_{i,j=1}^n \frac{\partial^2 a_i}{\partial z_j \partial z_k}(f)f_i^{'}f_j^{'}f_k^{'} + 3\sum_{ij=1}^n \frac{\partial a_i}{\partial z_j}(f)f_i^{''}f_j^{'} + \sum_{i=1}^n a_i(f)f_i^{'''}.$$

Example 9.3. (See [8] for details.) Let

$$X = \{[z_0, z_1, z_2] \in \mathbf{P}^2 \,|\, P(z_0, z_1, z_2) = 0\}$$

be a non-singular curve of degree $d = 4$. Using the preceding we may write down explicitly a basis for $H^0(\mathcal{J}_k^m X)$. We demonstrate via examples. For $d = 4$, it can be shown that

$$h0(\mathcal{J}_2^2 X) = h^0(\mathcal{K}_X^2) + h^0(\mathcal{K}_X) = 6 + 3 = 9$$

(where $\mathcal{K}_X$ is the canonical line bundle) and, since the genus is 3, there are 3 linearly independent 1-forms $\omega_1, \omega_2, \omega_3$ which may be taken as ([8]):

$$\omega_1 = \frac{z_0(z_0dz_1 - z_1dz_0)}{\partial P/\partial z_2}, \omega_2 = \frac{z_1(z_0dz_1 - z_1dz_0)}{\partial P/\partial z_2}, \omega_3 = \frac{z_2(z_0dz_1 - z_1dz_0)}{\partial P/\partial z_3}.$$

A basis for $H^0(\mathcal{J}_2^2 X)$ is given by

$$\omega_1^{\otimes 2}, \omega_2^{\otimes 2}, \omega_3^{\otimes 2}, \omega_1 \otimes \omega_2, \omega_1 \otimes \omega_3, \omega_2 \otimes \omega_3, \delta\omega_1, \delta\omega_2, \delta\omega_3$$

where δ is the derivation defined in Proposition 2.1. The first six of these provide a basis of $H^0(\mathcal{K}_X 2)$ and the last three may be identified with a basis of $H^0(\mathcal{K}_X)$. For $\mathcal{J}_2^3 X$ we have

$$\begin{aligned} h^0(\mathcal{J}_2^3 X) &= h^0(\mathcal{K}_X^2) + h^0(\mathcal{K}_X^3) \\ &= h^0(\mathcal{O}_X(2(d-3))) + h^0(\mathcal{O}_X(3(d-3))) \\ &= \binom{2d-4}{2} - \binom{d-4}{2} + \binom{3d-7}{2} - \binom{2d-7}{2}. \end{aligned}$$

In particular, for $d = 4, h^0(\mathcal{J}_2^3 X) = h^0(\mathcal{K}_X^2) + h^0(\mathcal{K}_X^3) = 6 + 10 = 16$. A basis for $H^0(\mathcal{J}_2^3 X)$ is given by the six elements (identified with a basis of $H^0(\mathcal{K}_X^2)$)

$$\delta\omega_1^{\otimes 2}, \delta\omega_2^{\otimes 2}, \delta\omega_3^{\otimes 2}, \delta(\omega_1 \otimes \omega_2), \delta(\omega_1 \otimes \omega_3), \delta(\omega_2 \otimes \omega_3)$$

and the 10 elements (a basis of $H^0(\mathcal{K}3_X)$):

$$\begin{aligned} &\omega_1^{\otimes 3}, \ \omega_2^{\otimes 3}, \ \omega_3^{\otimes 3}, \ \omega_1 \otimes \omega_2 \otimes \omega_3, \\ &\omega_1^{\otimes 2} \otimes \omega_2, \ \omega_1^{\otimes 2} \otimes \omega_3, \ \omega_2^{\otimes 2} \otimes \omega_1, \ \omega_2^{\otimes 2} \otimes \omega_3, \ \omega_3^{\otimes 2} \otimes \omega_1, \ \omega_3^{\otimes 2} \otimes \omega_2. \end{aligned}$$

□

Remark 9.4. The concept of jet bundles extends also to singular spaces. Let us remark on how this may be defined. One may locally embed an open set U of X as a subvariety in a smooth variety Y. By abuse of notation we write simply $U \subset Y$. At a point $x \in U$ the stalk jet $(J^k Y)_x$ is then defined (Y is smooth). The stalk $(J^k X)_x$ is defined as the subset

$$\{j^k f(0) \in (J^k Y)_x \mid f : \Delta_r \to Y \text{ is holomorphic }, f(0) = x \text{ and } f(\Delta_r) \subset U\}.$$

□

Example 9.5. Consider the embedded curve in **C**2:

$$X = \{(x, y) \mid y^2 = x^3\}$$

with a simple cusp at the origin. Differentiation yields:

$$2ydy = 3x^2 dx$$

thus J^1X is simply the variety in C4 defined by

$$J^1X = \{(x, y; x_1, y_1) \in \mathbf{C}^2 \times \mathbf{C}^2 \mid y^2 = x^3, 2yy_1 = 3x^2x_1\}.$$

Consider the projection $p_{10} : J^1X \to X$ induced by the projection $pr_2 : \mathbf{C}^2 \times \mathbf{C}^2 \to \mathbf{C}^2$ onto the first factor. Then $p_{10}^{-1}(0,0) \cong \mathbf{C}^2$ but for $(x, y) \neq (0,0)$

$$p_{10}^{-1}(x, y) \cong \{(x_1, y_1) \in \mathbf{C}^2 \mid 2yx_1 = 3x^2y_1\} \cong \mathbf{C}$$

is a line through the origin in $\mathbf{C}^2$.

Differentiate further, we get:

$$2(dy)^2 + 2y\delta^2 y = 6x(dx)^2 + 3x^2\delta^2 x.$$

Thus J^2X is the variety in $\mathbf{C}^6$ defined by

$$J^2X = \{(x, y; x_1, y_1; x_2, y_2) \in \mathbf{C}^6 \mid f = f_1 = f_2 = 0\}$$

where

$$\begin{cases} f(x, y) = y^2 - x^3, \\ qf_1(x, y; x_1, y_1) = 2yy_1 - 3x^2x_1, \\ f_2(x, y; x_1, y_1; x_2, y_2) = 2y_1^2 + 2yy_2 - 6xx_1^2 - 3x^2x_2 \end{cases}$$

For 2-jets we have two projections:
$p_{20} : J^2X \to X$ be the projection induced by the projection $\mathbf{C}^2 \times \mathbf{C}^4$ onto the first factor, and
$p_{21} : J^2X \to J^1X$ be the projection induced by the projection $\mathbf{C}^4 \times \mathbf{C}^2$ onto the first factor.

We have:
(1) $p_{21}^{-1}(0, 0; x_1, 0) \cong \mathbf{C}^2$ for any x_1
(2) $p_{21}^{-1}(0, 0; 0, y_1) = \varnothing$ if $y_1 \neq 0$,
(3) $p_{21}^{-1}(x, y; 0, 0) = \mathbf{C}$ if $x \neq 0, y \neq 0$ is a line in $\mathbf{C}^2$ through the origin,
(4) $p_{21}^{-1}(x, y; x_1, y_1) = \mathbf{C}$ if $x \neq 0, y \neq 0, x_1 \neq 0, y_1 \neq 0$ is an affine line in $\mathbf{C}^2$ not passing through the origin. □

Let $f : \mathbf{C} \to X$ be a holomorphic map such that the k-jet $j^k f : \mathbf{C} \to J^kX$ is non-trivial in the sense that the image $j^k f(\mathbf{C})$ is not entirely contained in the zero-section of J^kX then

$$[j^k f] : \mathbf{C} \to \mathbf{P}(J^kX)$$

(where $[\ ] : J^kX \to \mathbf{P}(J^kX)$ is the quotient map) is a well-defined holomorphic map. Clearly we have $p \circ [j^k f] = f$ where $p : (J^kX) \to X$ is the projection. This map shall be referred to as the *canonical lifting of f*. In the theory of holomorphic maps it is usually difficult to "move" the map f holomorphically. It is, however much easier to "move" the jet $j^k f$ in J^kX, by that we mean moving the image of $j^k f$ along the fiber of a generic point. More precisely, we look for a holomorphic map $g_k : \mathbf{C} \to \mathbf{P}(J^kX)$ such that $g_k(\zeta)$ is in the fiber $\mathbf{P}(J^k_{f(\zeta)}(X))$ for all ζ but at a prescribed point $\zeta_0, g_k(\zeta_0) \neq$

$j^k f(\zeta_0)$. This is possible basically due to the fact that $\mathbf{C}$ is contractible hence the pull-back $f^*(\mathbf{P}(J^k X))$ is a trivial bundle with fiber the weighted projective space $\mathbf{P}(Q_{k,n})$. Thus, if there is an automorphism ϕ of $\mathbf{P}(Q_{k,n})$, sending an element $(f(\zeta), v_1)$ to $(f(\zeta), v_2)$ then the automorphism can be extended to a bundle morphism of $f^*(\mathbf{P}(J^k X))$. Even though a weighted projective space admits many automorphisms it is not a homogeneous space. For example, a singular point cannot be moved to a non-singular point by an automorphism, we can always do so via rational maps.

Consider the following commutative diagram:

$$\begin{array}{ccc} f^*(J^k X) & \xrightarrow{f_*} & J^k X \\ [\,]\downarrow & & \downarrow [\,] \\ f^*\mathrm{P}(J^k X) & \xrightarrow{f_*} & \mathrm{P}(J^k X) \\ \rho\downarrow \quad [j^k f] \nearrow & & \downarrow p \\ \mathrm{C} & \xrightarrow{f} & X \end{array} \tag{9.4}$$

Since $\mathbf{C}$ is contractible and Stein, Oka Principle implies that the bundle $f^*(J^k X) \to \mathbf{C}$ is holomorphically trivial. The same is then also true for $f^*\mathbf{P}(J^k X) \to \mathbf{C}$. Thus we have a commutative diagram

$$\begin{array}{ccccc} \mathbf{C} \times \mathbf{P}(Q_{k,n}) & \xrightarrow{\Phi} & f^*\mathbf{P}(J^k X) & \xrightarrow{f_*} & \mathbf{P}(J^k X) \\ p_1 \downarrow & & \rho \downarrow [j^k f] \nearrow & & \downarrow p \\ \mathbf{C} & \xrightarrow{\mathrm{id}} & \mathbf{C} & \xrightarrow{f} & X \end{array} \tag{9.5}$$

where $\mathbf{P}(Q_{k,n})$ is the weighted projective space which is isomorphic to any of the fibers of $\mathbf{P}(J^k X)$, Φ is the trivialization map and p_1 is the projection onto the first factor. From the definition (2.1.3) of the $\mathbf{C}^*$-action, the fiber $\mathbf{P}(J^k X)_x$ is a weighted projective space of type

$$Q_{k,n} = (1, \cdots, 1, 2, \cdots, 2, ..., k, \cdots, k)$$

(each integer is repeated n-times).

In general for any $Q = (q_0, ..., q_r)$ be a $(r+1)$-tuples of positive integers and (C^{r+1}, Q) be the $(r+1)$-dimensional complex vector space such that the variable $z_i, 0 \leq i \leq r$ is assigned the *weight* (or *degree*) q_i. A $\mathbf{C}^*$-action is defined on $(\mathbf{C}^{r+1}, Q)$ by:

$$\lambda \cdot (z_0, ..., z_r) = (\lambda^{q_0} z_0, ..., \lambda^{q_r} z_r), \ \lambda \in \mathrm{C}^*.$$

The quotient space, $\mathbf{P}(Q) = (\mathbf{C}^{r+1}, Q)/\mathbf{C}^*$, is the *weighted projective space* of type Q. (Note that for the fiber $\mathbf{P}(Q_{k,n})$ of the projectivized k-jet bundle over an n-dimensional manifold X, $r = nk - 1$.) The equivalence class of an element $(z_0, ..., z_r)$ is denoted by $[z_0, ..., z_r]_Q$. For $Q = (1, ..., 1) = \mathbf{1}, \mathbf{P}(Q) = \mathbf{P}^r$ is the usual complex projective space of dimension r and an element of P^r is denoted simply by $[z_0, ..., z_r]$. For a tuple Q define a map $\psi_Q : (\mathbf{C}^{r+1}, \mathbf{1}) \to (\mathbf{C}^{r+1}, Q)$ by

$$\psi_Q(z_0, ..., z_r) = (z_0^{q_0}, ..., z_r^{q_r}).$$

It is easily seen that ρ_Q is compatible with the respective C^*-actions and hence descends to a well-defined morphism:

$$[\psi_Q] : \mathbf{P}^r \to \mathbf{P}(Q), \;\; [\psi_Q]([z_0, ..., z_r]) = [z_0^{q_0}, ..., z_r^{q_r}]_Q.$$

The weighted projective space can be alternatively described as follows. Denote by Θ_{q_i} the group consisting of all q_i-th roots of unity then the group $\Theta_Q = \oplus_{i=0}^r \Theta_{q_i}$ acts on P^r by coordinatewise multiplication:

$$(\theta_0, ..., \theta_r).[z_0, .., z_r] = [\theta_0 z_0, ..., \theta_r z_r], \; \theta_i \in \Theta_{q_i}.$$

The the quotient space P^r/Θ_Q is isomorphic to $\mathrm{P}(Q)$.

Theorem 9.6. *The weighted projective space $\mathbf{P}(Q)$ is isomorphic to the quotient $\mathbf{P}_r/\Theta_Q$. In particular, (Q) is irreducible and normal (the singularities are cyclic quotients and hence rational).*

The singular set can be described explicitly as follows.

Theorem 9.7. *Let $Q = (q_0, ..., q_r)$ be q $(r+1)$-tuple of positive integers such that* $\gcd\{q_0, .., q_r\} = 1$ *and* $\gcd\{q_0, ..., \hat{q}_i, ..., q_r\} = 1$ *for $i = 0, ..., r$. Then the a point $[z_0, ..., z_r]_Q$ is a singular points of $\mathbf{P}(Q) =$ if and only if* $\gcd\{q_i \mid z_i \neq 0\} > 1$. *The set of regular points $\mathbf{P}(Q)_{\mathrm{reg}}$ is simply connected.*

Example 9.8. Let $Q_{k,2} = (1, 1; 2, 2; ..., k, k)$ then the singular points are of the form:

$$\begin{array}{l}[0, 0; *, *; 0, 0; 0, 0; ...; 0, 0]_Q \\ [0, 0; 0, 0; *, *; 0, 0; ...; 0, 0]_Q \\ \quad \cdots \\ \quad \cdots \\ [0, 0; 0, 0; 0, 0; ...; 0, 0; *, *]_Q\end{array}$$

where $*$ represents non-zero complex numbers. The weighted projective space $\mathbf{P}(Q_{k,2})$ is of dimension $2k-1$. It is non-singular if and only if $k = 1$, and for $k \geq 2$,

$$\dim \mathbf{P}(Q_{k,2})_{\mathrm{sing}} = 1.$$

The projectivized jet bundle $\mathbf{P}(J^k X)$ over a surface X is of dimension $2k+1$. It is non-singular if and only if $k = 1$ and

$$\dim \mathbf{P}(J^k X)_{\mathrm{sing}} = 3.$$

In general, the weighted projective space $\mathbf{P}(Q_{k,n})$ is of dimension $nk-1$. It is non-singular if and only if $k = 1$ and

$$\dim \qquad mathbf{P}(Q_{k,2})_{\mathrm{sing}} = n - 1.$$

The projectivized jet bundle $\mathbf{P}(J^k X)$ over a manifold X of dimension n is of dimension $n(k+1)-1$. It is non-singular if and only if $k = 1$ and

$$\dim \mathbf{P}(J^k X)_{\mathrm{sing}} = 2n - 1.$$

□

The following notion is standard in algebraic geometry:

Definition 9.9. A coherent sheaf $\mathcal{S}$ of rank r over a complex space X of dimension n is said to be *big* if

$$\dim H^0(X, \mathcal{S}^{\otimes m}) = O(m^{n+r-1}).$$

□

Remark 9.10. (i) It is well-known that $\mathcal{S}$ is big if and only if the Serre line sheaf $\mathcal{L}_{\mathbf{P}(J^k X)}$, over the projectivization $\mathbf{P}(J^k X)$, is big. The Serre line bundle is the line sheaf on $\mathbf{P}(J^k X)$ whose restriction to each fiber is the sheaf $\mathcal{O}_{\mathbf{P}(Q_{n,k})}(1)$ of the weighted projective space $\mathbf{P}(Q_{n,k})$.

(ii) It is well-known that the space of sections

$$H0(X, \odot^m \mathcal{S} \otimes [D]) \text{ and } H^0(\mathbf{P}(\mathcal{S}^\vee), \mathcal{L}^{\otimes m}_{\mathbf{P}(\mathcal{S}^\vee)} \otimes p^*[D])$$

are canonically isomorphic where p is the projection map and D is a divisor on X. In particular,

$$H0(X, \odot^m \mathcal{J}^m_k X \otimes [D]) \text{ and } H^0(\mathbf{P}(J^k X), \mathcal{L}^{\otimes m}_{\mathbf{P}(J^k X)} \otimes p^*[D]).$$

(iii) It is well known that, if a coherent $\mathcal{S}$ over X is big then, for any effective divisor D in X

$$\dim H^0(X, \mathcal{S}^{\otimes m}(-D)) = O(m^{n+r-1}), \qquad n = \dim X.$$

This follows from the exact sequence:

$$0 \to \mathcal{S}(-D) \xrightarrow{\otimes\sigma} \mathcal{S} \xrightarrow{\rho|_D} \mathcal{S}^{\otimes m}|_D \to 0$$

where $[\sigma = 0] = D$ and $\rho|_D$ is the restriction map. The associate long exact sequence starts with

$$0 \to H^0(X, \mathcal{S}^{\otimes m}(-D)) \xrightarrow{\otimes\sigma} H^0(X, \mathcal{S}^{\otimes m}) \xrightarrow{\rho|_D} H^0(D, \mathcal{S}^{\otimes m}|_D) \to ...$$

The middle term is $O(m^{n+r-1})$ and it is well-known that $H^0(D, \mathcal{S}^{\otimes m}|_D)$ is of lower order because $\dim D = n - 1$. □

We now return to the diagrams (9.4) and (9.5). By definition

$$f^*\mathbf{P}(J^k X) = \{(\zeta, v) \in \mathbf{C} \times \mathbf{P}(J^k X) \mid f(\zeta) = p(v)\}$$

and the commutativity of the diagram implies that $[j^k f](\zeta) = f_*(\zeta, [j^k f](\zeta))$. Triviality of the bundle then implies that there is a holomorphic section $\phi : \mathbf{C} \to \mathbf{C} \times \mathbf{P}(Q_{k,n})$ ($p_1 \circ \phi = \text{id}$) such that

$$f_* \circ \Phi \circ \phi = [j^k f].$$

We may write $\phi(\zeta) = (\zeta, \psi(\zeta))$ where $\psi : \mathbf{C} \to \mathbf{P}(Q_{k,n})$ is a holomorphic map. Thus we can move $[j^k f]$ by moving the section ϕ (which amounts to moving ψ).

More precisely, let $u_0 = \psi(\zeta_0) \in \mathrm{P}(J^kX)_{\zeta_0} = \mathbf{P}(Q_{k,n})$ (*observe that there is no loss of generality in assuming that* $f_i(\zeta_0) = 1$ *for all* i), where ζ_0 is a fixed point in $\mathbf{C}$, and an arbitrary point $[v] \in \mathbf{P}(Q_{k,n})$ there exists a *rational* map

$$\gamma : \mathbf{P}(Q_{k,n}) \to \mathbf{P}(Q_{k,n})$$

such that $\gamma([u_0]) = [v]$ inducing a *rational* map

$$G : \mathbf{C} \times \mathbf{P}(Q_{k,n}) \to \mathbf{C} \times \mathbf{P}(Q_{k,n}), \quad G(\zeta, [\xi]) = (\zeta, \gamma([\xi])). \tag{9.6}$$

Then we get a section

$$G \circ \phi : \mathbf{C} \to \mathbf{C} \times \mathbf{P}(Q_{k,n}), \quad G \circ \phi(\zeta) = (\zeta, \gamma([\xi])) \tag{9.7}$$

which is *holomorphic* and $G \circ \phi(\zeta_0) = (\zeta_0, \gamma([u_0])) = (\zeta_0, [v])$. (For example, in the case $k = 1, \mathbf{P}(Q_{1,n})$ is just the usual projective space $\mathbf{P}^{n-1}$, and is a homogeneous manifold. Thus we can simply take γ to be an automorphism sending $[u_0]$ to $[v]$.) In any case γ may be chosen to be of the form

$$\begin{aligned} &\gamma\big(x_1 = (x_1^1, ..., x_1^n); ...; x_k = (x_k^1, ..., x_k^n)\big) \\ &= \Big(\big(P_1^1(x_1), ..., P_1^n(x_1)\big); ...; \big(..., P_k^1(x_1, ..., x_k), ..., P_k^n(x_1, ..., x_k)\big)\Big), \end{aligned} \tag{9.8}$$

where $\big((..., x_1^j, ...); ...; (..., x_k^j, ..)\big), 1 \leq j \leq n$, are the *homogeneous coordinates* on $\mathbf{P}(Q_{k,n})$ and each $P_i^\ell, 1 \leq i \leq n, 1 \leq \ell \leq k$ is a weighted homogeneous polynomial of degree dj (with coefficients in $\mathbf{C}$). *The map* γ *is said to be a weighted homogeneous map of weight* d. The map

$$[g_k] = f_* \circ \Phi \circ G \circ \phi : \mathbf{C} \to \mathbf{P}(J^kX) \tag{9.9}$$

is holomorphic (a meromorphic map from a curve into a projective variety is necessarily holomorphic) and is a lifting of f, i.e., $p \circ [g_k] = f$ with the property that

$$[g_k(\zeta_0)] = f_* \circ \Phi(\zeta, [v]).$$

On the other hand we have:

$$[j^k f(\zeta_0)] = f_* \circ \Phi \circ \phi(\zeta_0) = f_* \circ \Phi(\zeta_0, [u_0]).$$

We have thus successfully move $[j^k f]$ to $[g_k]$. Observe that

$$g_k(\zeta) = (f(\zeta), \Big(\big(P_1^1(j^1 f)), ..., P_1^n(j^1 f)\big); ...; \big(P_k^1(j^k f), ..., P_k^n(j^k f)\big)\Big). \tag{9.10}$$

We summarize the construction above in the following lemma:

Lemma 9.11. (Moving Lemma) *Let* $f : \mathbf{C} \to X$ *be a holomorphic map into a projective manifold. Assume that the image of* $j^k f$ *is not contained entirely in the zero section of* J^kX. *Let* ζ_0 *be an arbitrary point of* $\mathbf{C}$ *and* $[\xi_0] \in \mathbf{P}(J^kX)_{f(\zeta_0)} \cong \mathbf{P}(Q_{k,n})$. *Then there exists a holomorphic map*

$$[g_k] : \mathbf{C} \to \mathbf{P}(J^kX)$$

such that $(i)\, p \circ [g_k] = f$, $(ii)\, [g_k](\zeta_0) = [\xi_0]$, *and* $(iii)\ [g_k] = (f, [P_1, ..., P_k])$ *where each* $[P_1, ..., P_k]$ *is a weighted homogeneous polynomial map of degree* d *in* $j^k f$.

The proof of the Moving Schwarz Lemma is based on the Lemma of Logarithmic Derivatives in Nevanlinna Theory. We shall need the following variation of Theorem 6.1 in [8]:

Lemma 9.12. (Logarithmic Derivatives : Moving Version) *Let* X *be a projective variety. Given*

(i) *a non-constant holomorphic map* $f : \mathbf{C} \to X$,

(ii) *let* $[g_k] : \mathbf{C} \to \mathbf{P}(J^k X)$ *be a holomorphic lifting of* f *constructed in the preceding lemma,*

(iii) $0 \not\equiv \omega \in H^0(X, \mathcal{J}_k^m X)$ (*or* $H^0(X, \mathcal{J}_k^m X(\log D))$ *where* D *is an effective divisor with simple normal crossings*) *such that* $\omega \circ g_k \not\equiv 0$. *Then*

$$T_{\omega \circ g_k}(r) = \int_0^{2\pi} \log^+ \left| \omega(g_k(re^{\sqrt{-1}\theta})) \right| \frac{d\theta}{2\pi} \le O\big(\log T_f(r)\big).$$

Proof. As in the proof of Theorem 6.1 in [8] there exist a finite number of rational functions $t_1, ..., t_q$ on X such that:

($\dagger$) *the logarithmic jet differentials* $\{(d^{(j)} t_i / t_i)^{m/j} \mid 1 \le i \le q, 1 \le j \le k\}$ *span the fibers of* $\mathcal{J}_k^m X(\log D)$ (*resp.* $\mathcal{J}_k^m X$) *over every point of* X.

To obtain the estimate of the theorem, we proceed analogously as in [8], by considering the function,

$$\rho : J^k X(-\log D) \to [0, \infty],$$

defined by

$$\rho(\xi) = \sum_{i=1}^{N} \sum_{j=1}^{k} |(d^{(j)} \tau_i / \tau_i)^{md/j}(\xi)|^2,\ \xi \in J^k X(-\log D)\ \ (\text{resp. } J^k X) \qquad (\dagger\dagger)$$

is continuous in the extended sense; ρ is *strictly positive* (possibly $+\infty$) outside the zero-section of $J^k X$ (see the proof in [8] for more details, because the jet differentials in ($\dagger$) span). Fix a fiber $F_x = J_x^k X, x \in X$ and let $P_i, 1 \le i \le k$, be weighted homogeneous polynomials of degree d in ξ. Denote by $P(\xi) = (P_1(\xi), .., P_k(\xi))$ and define,

$$R_{x,P}(\xi) = \frac{|\omega(P(\xi))|^2}{\rho(\xi)} : (F_x)_* \to [0, \infty).$$

Observe that:

$$\frac{|\omega(P(\lambda\xi))|^2}{\rho(\lambda\xi)} = \frac{|\lambda|^{2md} |\omega(P(\xi))|^2}{|\lambda|^{2md} \rho(\xi)} = \frac{|\omega(P(\xi))|^2}{\rho(\xi)}$$

for all $\lambda \in \mathbf{C}$ we see that R_x descends to a well-defined continuous function on the projectivization $\mathbf{P}_x(J^k X)$ and so, by compactness, there exists a constant

$$c_{x,P} = \max_{[\xi] \in \mathbf{P}_x(J^k X)} R([\xi])$$

with the property that

$$|\omega(P(\xi))|^2 \le c_{x,P}\,\rho(\xi)$$

for all ξ and P. Since the space of weighted homogeneous polynomials of a fixed degree may be identified with some projective variety, we get

$$|\omega(P(\xi))|^2 \le c_x\rho(\xi), \quad c_x = \max_{[Q]|\deg Q=d} c_{x,[Q]} < \infty.$$

Obviously, c_x is a continuous function of x, hence (as X is compact) we may replace c_x by $C = \max_{x\in X} c_x$ in the inequality above. This implies that

$$\begin{aligned} T_{\omega\circ g_k}(r) &= \int_0^{2\pi} \log^+ |\omega(g_k(re^{\sqrt{-1}\theta}))| \frac{d\theta}{2\pi} \\ &\le \int_0^{2\pi} \log^+ |\rho(j^k f(re^{\sqrt{-1}\theta}))| \frac{d\theta}{2\pi} + O(1). \end{aligned}$$

The classical lemma of logarithmic derivatives for meromorphic functions and (††) implies the following estimate:

$$\begin{aligned} &\int_0^{2\pi} \log^+ |\rho(g_k(re^{\sqrt{-1}\theta}))| \frac{\theta}{2\pi} \\ &\le O\Big(\int_0^{2\pi} \sum_{i=1}^{N} \log^+ |(t_i\circ f)^{(j)}/t_i\circ f| \frac{d\theta}{2\pi}\Big) \\ &\le O\big(\log r(T_f(r))\big) \end{aligned}$$

as claimed. □

By remark 9.10 we can identify a jet differential $\omega \in H0(X, \mathcal{J}_k^m \otimes [-H])$ with an element of $\sigma_\omega \in H^0(\mathbf{P}(J^kX), \mathcal{L}^m \otimes p^*[-H])$. The canonical isomorphism is such that $\omega(g_k) = \sigma_\omega([g_k])$. *We shall abuse notation and write ω for σ_ω as well.* The following is a very useful consequence of the Moving Lemma:

Theorem 9.13. (Moving Schwarz Lemma) *Let X be a compact complex manifold, $f : \mathbf{C} \to X$ be a non-constant holomorphic map and $[g_k] : \mathbf{C} \to \mathbf{P}(J^kX)$ a lifting as constructed in the Moving Lemma (Lemma 9.11). Let H be an effective generic ample divisor in X and $\omega \in H^0(Y, \mathcal{L}^m_{\mathbf{P}(J^kX)} \otimes p^*[-H])$ be a non-trivial section (where $\mathcal{L}_{\mathbf{P}(J^kX)}$ is the Serre line bundle and $p : \mathrm{P}(J^kX) \to X$ is the projection map). Then $\omega([g_k]) \equiv 0$.*

Proof. By Lemma 9.12, if $\omega([g_k]) \not\equiv 0$ then

$$T_{\omega\circ g_k)}(r) = O\big(\log T_f(r)\big).$$

On the other hand, there is no loss of generality that, the ample divisor H is generic. From Nevanlinna Theory we know that $N_f(H,r) = O(T_f(r))$ for generic ample divisors. Since $[g_k]$ is a lifting of f the intersection number of $[g_k]$ with $p^*(-H)$ is not smaller than the intersection number of f with H. This implies that

$$N_{\omega\circ g_k}(p^*H, r) \ge N_f(H,r) = O(T_f(r)),$$

thus

$$O(T_f(r)) \leq N_{\omega \circ g}(p^*H, r) \leq T_{\omega \circ g_k}(r) \leq O(\log r T_f(r)).$$

This is impossible, as we may assume without loss of generality that f is transcendental (so $\log r / T_f(r) \to 0$). This means that $\omega(g_k) \equiv 0$ as claimed. □

Let X be a projective variety and

$$S = \{x \in f(\mathbf{C}) \text{ for some non-constant holomorphic curve } f : \mathbf{C} \to X\}.$$

Let $\mathcal{L} = \mathcal{L}_{\mathbf{P}(J^k X)}$ be the Serre line bundle over $\mathbf{P}(J^k(X))$. Define, for a very ample divisor H in X,

$$B_H = \cap_{m \geq 0} \{\xi \in \mathbf{P}(J^k(X)) \mid \sigma(\xi) = 0, \sigma \in H^0(\mathbf{P}(J^k(X), \mathcal{L}^m \otimes p^*(-H))\}$$

and

$$B = \cap_H B_H$$

where the intersection is taken over all very ample divisor of X.

Theorem 9.14. (Structure Theorem for Base Locus) *Let X be a projective variety and assume that the sheaf of germs of k-jet differentials $\mathcal{J}_k X$ is big. Then $p^{-1}(S)$ is contained in B where $p : \mathbf{P}(J^k X) \to X$ is the projection map.*

Proof. For any $x_0 = f(\zeta_0) \in f(\mathbf{C})$ where $f : \mathbf{C}_p \to X$ is analytic. Let $(x_0), v)$ be any point in the fiber $\mathbf{P}(J^k X)_{x_0}$, we get via the Moving Lemma the existence of a lifting, $[g_k] : \mathbf{C} \to \mathbf{P}(J^k X)$, of f such that $[g_k](\zeta_0) = (f(\zeta_0), v) = (x_0, v)$ and satisfies the growth condition $T_{[g_k]}(r) = T_f(r)$. By the moving Schwarz Lemma, $[g_k](\mathbf{C}) \subset B_H$ for any very ample divisor H. Thus $(f(\zeta), v)$ is in B. Since (x_0, v) is an arbitrary point in the fiber $\mathbf{P}(J^k X)_{x_0}$, the entire fiber $\mathbf{P}(J^k X)_{x_0}$ is in B. Since x_0 is an arbitrary point in S we have $p^{-1}(S) \subset B$ as claimed. □

If X is a surface, the base locus is the union of the preimages of all rational curves and elliptic curves in X.

Corollary 9.15. *Let X be a complex projective variety and assume that the sheaf of germs of k-jet differentials is big. Then X is Brody hyperbolic, i.e., the image of every holomorphic map $f : \mathbf{C} \to X$ is contained in a subvariety Y in X of strictly lower dimension.*

Proof. Let $f : \mathbf{C} \to X$ be a non-constant entire holomorphic curve in X and Y be the Zariski closure of $f(\mathbf{C})$ in X. Since $f(\mathbf{C}) \subset S$ hence $p^{-1}(f(\mathbf{P})) \subset B$ which is a subvariety of $\mathbf{P}(J^k X)$. Thus $p^{-1}(Y)$ is also contained in B, hence is contained in B. If the image $f(\mathbf{C})$ is not contained in a proper subvariety of X then $Y = X$, but then $B = p^{-1}(Y) = p^{-1}(X) = \mathbf{P}(J^k X)$. This is absurd, thus $f(\mathbf{C})$ must be contained in a proper subvariety of X. □

Remark 9.16. The preceding (Lemma 9.11 - Corollary 9.15) are valid for logarithmic jets $\mathcal{J}_k^m X(\log D)$ (with the condition that $f : \mathbf{C} \to X \setminus D$) instead of $\mathcal{J}_k^m X$. □

The next result can be found in [8] (see also [18]):

Theorem 9.17. *The sheaf of k-jet differential $\mathcal{J}_k X$ of a hypersurface X in $\mathbf{P}^3$ is big if* $\deg X \geq 5$.

By Corollary 9.15, every entire holomorphic curve $f : \mathrm{C} \to X$ is contained in an algebraic curve, necessarily, rational or elliptic. By a result of Xu [52], a generic hypersurface X in P3 of degree ≥ 5 does not contain any rational or elliptic curve thus (see [8]):

Corollary 9.18. A generic hypersurface X of degree $d \geq 5$ in $\mathbf{P}^3$ is Brody hyperbolic.

10 The algebraic version of Oka's principle

In Serre's celebrated paper on coherent sheaves, it was shown that an $\mathcal{O}_X$-coherent (resp. locally free) sheaf $\mathcal{M}$ over an affine algebraic variety X (defined over an algebraically closed field $\mathbf{C}$ of characteristic zero) may be identified with a finitely presentable (resp. projective) module M over the coordinate ring $\mathcal{O}_X(X)$. In the special case where X is the affine space $\mathbf{K}^n$, $\mathcal{O}_X(X)$ is just the polynomial ring $\mathbf{K}[t_1, ..., t_n]$ and in general, the coordinate ring of an affine algebraic variety is the quotient $\mathbf{K}[t_1, ..., t_N]/I$ for some ideal I. The famous *Serre's problem* is the following:

"Is a projective module over a polynomial ring $\mathbf{K}[t_1, ..., t_n]$ free?"

Geometrically this is equivalent to asking

"Is every algebraic vector bundle over $\mathbf{K}^n$ trivial?"

In the case $\mathbf{K} = \mathbf{C}$ this is a special case of the Oka Principle and the answer is affirmative because $\mathbf{C}^n$ is contractible so every topological bundle is topological trivial, hence also holomorphically trivial. Serre's algebraic Oka Principle was established affirmatively by Quillen and Suslin, independently, in 1976.

The analogue of Stein varieties, on which the Oka Principle is establised, can be analogously defined over the p-adic number field $\mathbf{C}_p$ with the p-adic absolute value $|\ |_p$ replacing the usual Archimedean absolute on $\mathbf{C}^n$. Analytic functions can be defined as convergent power series. Since the p-adic absolute value is non-Archimedean it is necessary to work in the categories of *affinoid varieties* and *rigid analytic varieties* as introduced by Tate. It will take too much space here to properly define these concepts so we shall be contented with a very brief description and refer to [2] for the precise definitions (and further references). The algebra of convergent power series (defined over $\mathbf{C}_p$) in n variables and with radius of convergent $r > 0$ is known as the *Tate algebras* and will be denoted by $T_{n,r}$. The maximal spectrum (i.e. the space of all maximal ideals) $\mathrm{Max}(T_{n,r})$ can be identified with the space

$$\mathrm{Max}(T_{n,r}) = \overline{\Delta}^n(r) = \{(x_1, ..., x_n) \in \mathbf{C}_p^n \mid \max_{1 \leq i \leq n} |x_i|_p \leq r\}.$$

In general an *affinoid variety* is by definition a pair $\mathrm{Max}(R), R)$ where R is an *affinoid algebra*, that is, $R = T_{N,r}/I$ for some N and for some ideal I. In particular, $(\overline{\Delta}^n(r), T_{n,r})$ is an affinoid variety. Very roughly speaking, a rigid analytic variety is a space which admits an open cover consisting of affinoid varieties satisfying the obvious compatibility conditions. With these the analogue of a Stein variety can be defined:

Definition 10.1. A rigid analytic variety X is said to be *quasi-Stein* if there exists an exhaustion

$$X_1 \subset X_2 \subset \cdots \subset \cup_{i=1}^{\infty} X_i = X$$

where each X_i is an open subdomain which is affinoid. □

The definition is meant to imitate the fact that a complex Stein manifolds admits an exhaustion by pseudoconvex subdomains. We stress here that the p-adic topology (the metric topology defined by the p-adic absolute value) behaves very differently from the Archimedean absolute value on C and is not suitable for defining sheaves and sheaf co-homologies. It is therefore necessarily to work with the associated *Grothendieck topology* (abbrev. G-topology, which is, however, not a topology as the name suggested) and is included in the technical definition of rigid analytic varieties.

The following theorem (see [2]) shows that the quasi-Stein rigid varieties share many of the properties of complex Stein manifolds (see section 1 of this article)

Theorem 10.2. *Let X a quasi-Stein rigid analytic variety. Then*

(i) *$\mathcal{O}_{X,x}$ is a Noetherian local ring,*

(ii) *$\mathcal{O}_X(X)|_{X_i}$ is dense in $\mathcal{O}_{X_i}$,*

(iii) *(Theorem A) for each point $x \in X$ the evaluation map $H^0(X, \mathcal{S}) \to \mathcal{S}_x$ is surjective, in other words, global sections generate each stalk.*

(iv) *(Theorem B) $H^q(X, \mathcal{S}) = 0$ for all $q \geq 1$ for any coherent $\mathcal{O}_X$-sheaf $\mathcal{S}$, in particular, $H^q(X, \mathcal{O}_X) = 0$ for all $q \geq 1$ and $H0(X, \mathcal{O}_X) = \mathcal{O}_X(X)$.*

Theorem A and B in complex theory are known as Cartan's Theorem A and B. Serre's identification of locally free sheaves with projective modules for affine algebraic varieties is also valid in the category of quasi-Stein varieties. More precisely, a locally free $\mathcal{O})_X$-sheaf over a quasi-Stein rigid analytic variety is identified with a projective module over the sheaf of $\mathcal{O}_X(X)$-module. The Serre problem can also be formulated:

"Is a projective module over the ring $T_{n,\infty}$, of convergent power series with radius of convergence $r = \infty$, free?"

The Serre's problem is also valid in this form hence, we have:

"A locally free rigid analytic sheaf over the rigid analytic variety $\mathbf{C}_p^n$ is trivial.

Note that $\mathbf{C}_p$ admits two structures, the affine algebraic structure and the rigid analytic structure. The situation is similar to the complex case, $\mathbf{C}^n$ has the structure of an algebraic variety and also an analytic structure.

The results in section 9 can be extended to the p-adic case. In fact, some of the results in the p-adic case is stronger than the complex case. For example, a p-adic abelian varieties (in particular an elliptic curve, i.e. a curve of genus 1) is p-adic hyperbolic. The corresponding result is obviously false in the complex case. We refer to [2] for precise statements and details.

Bibliography

[1] A. Andreotti and T. Franke: The Lefschetz theorem on hyperplane sections *Ann. Math.* **69** (1959) 713–717

[2] T. T. H. An, J. T.-Y. Wang and P.-M. Wong: On p-adic hyperbolicity *Preprint* (2006)

[3] C. Banica and O. Forster: Complete intersections in Stein manifolds *Manus. Math.* **37** (1982) 343–356

[4] M. Boratynski: A note on set-theoretic complete intersection *J. Alg.* **54** (1978) 1–5

[5] J.-K. Cao and P.-M. Wong: Finsler Geometry of Projectivized Vector Bundles *J. Math. Kyoto Uni.* **43** (2003) 383–424

[6] H. Cartan: Espaces fibrés analytiques *In: Sympos. Internat. de topologia algebraica* (Mexico, 1958) 97–121; (reprinted in Oeuvres II, Sringer-Verlag 1979)

[7] K. Chandler and P.-M. Wong: On the holomorphic sectional and bisectional curvatures in complex Finsler geometry *Periodica Mathematica Hungarica* **48** (2004) 93–123

[8] K. Chandler and P.-M. Wong: Finsler Geometry of Holomorphic Jet Bundles *In: A Sampler in Riemann/Finsler Geometr* (Cambridge University press, 2004) 107–196

[9] M. Cornalba and P. Griffiths: Analytic cycles and vector bundles on non-compact algebraic varieties *Invent. Math.* **28** (1975) 1–106

[10] Y. Eliashberg: *Introduction to the h-principle* (Grad. Studies im Math. 48, Providence, R.I., AMS, 2002)

[11] Y. Eliashberg: Topological characterization of Stein manifolds of dimension $n > 2$ *Internat. J. of Math.* **1** (1990) 29–46

[12] O. Forster and K.J. Ramspott: Oka Paar von Garben nich-abelscher Gruppen *Invent.* **1** (1966) 260–286

[13] O. Forster and K.J. Ramspott: Analytische Modulgarben und Endromisbundel *Invent.* **2** (1966) 145–170

[14] F. Forstneric: On complete intersections *Ann. Inst. Fourier* **51** (2001) 497–512

[15] F. Forstneric: The homotopy principle in complex analysis: A survey *In: Explorations in complex and Riemannian geometry* (A volume dedicated to Robert E. Greene. Providence, RI: American Mathematical Society (AMS). Contemp. Math. 332, 2003) 73–99

[16] F. Forstneric and J. Prezelj: Oka's principle for for holomorphic fiber bundles with sprays *Math. Ann.* **317** (2000) 117–154

[17] M. Goresky and R. MacPherson: *Stratified Morse Theory* (Ergebnese der Math. 14, Springer, 1988)

[18] M. Green and P. Griffiths: Two applications of algebraic geometry to entire holomorphic mappings *In: Differential Geometry* (Proc. int. Chern Symp. Chern Symp., Berkeley, Calif., 1979, Springer-Verlag, New York, 1980) 41–74

[19] H. Grauert: Approximationssätze für holomorphe Funktionen mit Werten in komplexen Räumen *Math. Ann.* **133** (Globevnik-Stensones), (1957) 139–159

[20] H. Grauert: Holomorphe Funktionen mit Werten in komplexen Liesechen Gruppen *Math. Ann.* **133** (1957) 450–472

[21] M. Gromov: Pseudoholomorphic curves in symplectic manifolds *Invent. Math.* **82** (1985) 307–347

[22] M. Gromov: *Partial differential relations* (Ergeb. der Math. und ihrer Grenzgeb. 9, Springer-Verlag, Berlin-New York, 1986)

[23] M. Gromov: Oka's Principle for holomorphic sections of ellepltic bundles *JAMS* **2** (1989) 851–897

[24] A. Grothendieck: Sur la classification des fibrès holomorphes sur la sphère de Riemann *Am. J. Math.* **79** (1957) 121–138

[25] R. Gunning: *Introduction to Holomorphic Functions of Several Variables. Vol. I, II and III.* (Wadsworth & Brooks/Cole Mathematics Series, Brooks/Cole Publishing Company, 1990)

[26] H. A. Hamm: Zum Homotopietyp Steinscher Räume *J. Reine Angew. Math.* **338** (1983) 121–135

[27] G. Henkin and F. Leiterer: Oka-Grauert's principle without induction over the basis dimension *Math. Ann.* **311** (1988) 71–93

[28] F. Hirzebruch: *Topological Methods in Algebraic Geometry* (Springer-Verlag, 1966)

[29] D. Husemoller: *Fiber Bundles* (McGraw-Hill, 1966)

[30] J. Kajiwara: Equivalence of Steinness and validity of Oka's principle for subdomains with continuous boundaries of a Stein manifold *Mem. Fac. Sci. Kyushu Univ. Ser. A* **33** (1979) 83–93

[31] J. Kajiwara and M. Nishihara: Charakterisierung der Steinschen Teilgebiete durch das Okasches prinzip in zweidimensionaler Steinscher Mannigfaltigkeit *Mem. Fac. Sci. Kyushu Univ. Ser. A* **33** (1979) 71–76

[32] D. Krupka and M. Krupka: Jets and contact elements *In: Proceedings of the Seminar on Differential Geometry* (Math. Publ. vol. 2, 2000) 39–85

[33] N. M. Kumar: Complete intersections *J. Math. Koyoto Univ.* **17** (1977) 533–538

[34] N. M. Kumar: On Two Conjectures About Polynomial Rings *Invent. Math.* **46** (1978) 225–236

[35] E. Kunz: *Introduction to commutative algebra and algebraic geometry* (Birkhaüser, 1980)

[36] F. Larrusson: Excision for simplicial sheaves on the Stein site and Gromov's Oka Principle *International J. of Math.* **14** (2003) 191–209

[37] F. Larrusson: Model structures and the Oka Principle *J. Pure and Appl. Alg.* **192** (2004) 203–223

[38] J. Leiterer: Equivalence of Steinness and validity of Oka's principle for subdomains of Stein manifolds *Math. Nachr.* **89** (1979) 181–183

[39] M. Maican: *Vector bundles of finite order on affine manifolds* (Notre Dame Thesis, Univ. of Notre Dame, 2005)

[40] J. Mulflur, A. Vitter and P.-M. Wong: Holomorphic functions of finite order on affine varieties *Duke Math. J.* **48** (1981) 493–502

[41] K. Oka: Sur les fonctions analytiques de plusieurs variables III. Deuxiéme probléme de Cousin *J. Sci. Hiroshima Univ., Ser A* **9** (1939) 7–19

[42] G. Patrizio and P.-M. Wong: Stein manifolds with compact symmetric center *Math. Ann.* **289** (1991) 355–382

[43] F. P. Peterson: Some remarks on Chern classes *Ann. Math.* **69** (1959) 414–420

[44] H. Röhrl: Das Riemann-Hilbertsche Problem der Theorie der linearen Differentialgleichungen *Math. Ann.* **133** (1957) 1–25

[45] M. Schneider: On the number of equations needed to describe a variety *In: Proc. Symp. Pure Math. AMS* (41, 1982) 163–180

[46] J. Schürmann: Embeddings of Stein spaces into affine spaces of minimal dimensions *Math. Ann.* **301** (1997) 381–399

[47] K. Stein: Analytische Funktionen mehrerer komplexer Veränderichen zu vergegebenen Periodizitätsmoduln und das zweite Cousinsche Problem *Math. Ann.* **123** (1951) 201–222

[48] J. Wiegmann: Einbettungen kompxler Räume in Zahlenräume *Invent. Math.* **1** (1966) 229–242

[49] J. Winkelmann: The Oka principle for mappings between Riemann surfaces *l'Enseignment Math.* **39** (1993) 143–151

[50] P.-M. Wong: Applications of Nevanlinna Theory to Geometric Problems *In: Proc. ICCM* (edited by S.-T. Yau, to appear, 2005)

[51] P.-M. Wong and Philip P.-W. Wong: Bisectional Curvature of Complements of Curves In **P**2 *J. Math. Kyoto University* (to appear), (2005)

[52] G. Xu: Subvarieties of general hypersurfaces in projective space *J. Diff. Geom.* **39** (1994) 139–172

Pit-Mann Wong

Department of Mathematics, University of Notre Dame, Notre Dame, Indiana 46556 USA

E-mail: wong.2@nd.edu

Demeter Krupka and David Saunders

Abstracts

Global aspects of Finsler geometry
Tadashi Aikou and László Kozma

Finsler geometry originates from the calculus of variations, started in the twenties of the last century. The first essential movement towards global aspects of Finsler metrics on manifolds was due to L. Auslander in 1955. Then, the development of global (principal) connection theory, applied for Finslerian structures by M. Matsumoto in the sixties, opened the door to discuss the generalization of Riemannian global results for Finsler geometry. This was getting through by Chern's flourishing school in the last 15 years.

In this report we intend to sketch the state of today in this respect without the purpose of completeness. First the basics in Finsler Geometry are described: the fundamental function, the Chern connection, torsions and curvatures, the flag curvature. Then the properties of geodesics, the exponential map, the minimality of geodesics, and the Hopf-Rinow theorem are given. We discuss the first and second variation formulae, Jacobi fields, conjugate points, injectivity radius and related topics, such as Cartan-Hadamard and Bonnet-Myers theorems. We quote the fundamental comparison theorems of Finsler geometry: some depends on bounds of flag curvature, while others depend on bounds of the Ricci scalar curvature. Several rigidity theorems are presented, which give assumptions on the Finsler structure for the reduction to a locally Minkowski or Riemannian space. Then, applying Morse's theory for the energy functional, some results on the length and multiplicity of closed geodesics of Finsler manifolds, and the sphere theorem are reviewed. Finally, we report how the Gauss-Bonnet theorem have been extended for Finsler manifolds.

Morse theory and nonlinear differential equations
Thomas Bartsch, Andrzej Szulkin and Michel Willem

In this survey we treat Morse theory on Hilbert manifolds for functions with degenerate critical points. We describe the global and the local aspects of the theory, in particular the Morse inequalities, the Morse Lemma and critical groups. We consider applications to semi-linear elliptic problems and to closed geodesics on a riemannian manifold. The theory is extended to strongly indefinite functionals. Applications are given to periodic solutions of first order hamiltonian systems.

Index theory
David Bleecker

Early history of index theory is given, covering events leading to the discovery of the Atiyah-Singer Index Theorem. Basic facts are proven about Fredholm operators and the behavior of the index under perturbations, composition, etc.. The connection between families of Fredholm operators and K-theory is made in the Atiyah-Janich Theorem. Elliptic pseudo-differential operators on compact manifolds are shown to yield Fredholm operators. Outlines are given of the embedding proof of the Atiyah-Singer Index Theorem and the heat kernel proof for twisted Dirac operators. Statements are provided for twisted versions of the classical index theorems; e.g., the Hirzebruch-Signature, Chern-Gauss-Bonnet, and Hirzebruch-Riemann-Roch Theorems. Brief treatments of G-index theory and the Atiyah-Patodi-Singer Theorem are included.

Partial differential equations on closed and open manifolds
Jürgen Eichhorn

We present a survey of some important classes of partial differential equations on manifolds and of methods for solving them. This concerns questions of spectral theory, the heat equation and the heat kernel, the wave equation, Huygens' principle and the Hamiltonian approach, index theory on open manifolds, the continuity method and a choice of nonlinear equations important in geometry and mathematical physics. The spaces under consideration are linear and non-linear Sobolev structures which we briefly define at the beginning of our contribution.

The spectral geometry of operators of Dirac and Laplace type
Peter Gilkey

We survey results concerning asymptotic formulae in spectral geometry. We give explicit combinatorial formulas for both the heat trace asymptotics and for the heat content asymptotics in the context of smooth manifolds with boundary for the realization with respect to a variety of elliptic boundary conditions of an operator of Laplace type. We relate these formulaes to questions in spectral geometry and to the index theorem.

Lagrangian formalism on Grassmann manifolds
D. R. Grigore

The Lagrangian formalism on a arbitrary non-fibrating manifold is described from the kinematical point of view by (higher-order) Grassmann manifolds; such manifolds are obtained by factorization of the regular velocity manifold to the action of the differential group. Here we present the basic concepts of the Lagrangian formalism, as Lagrange, Euler-Lagrange and Helmholtz-Sonin forms, relevant in this context. These objects come in pairs, namely we have homogeneous objects (defined on the regular velocity manifold) and non-homogeneous objects (defined on the Grassmann manifold) and we give the connection between them. As a result the generic expressions for a variationally trivial Lagrangian and for a locally variational differential equation remain the same as in the fibrating case. Finally, we concentrate on the case of second order Grassmann bundles which are relevant for many physically interesting cases.

Sobolev spaces on manifolds
Emmanuel Hebey and Frédéric Robert

Sobolev spaces are important tools in several branches of mathematics. We discuss various aspects of Sobolev spaces on manifolds. While Sobolev spaces are well understood in Euclidean space, surprises occur in the context of Riemannian manifolds. Starting from the very first definition of such spaces, we discuss existence and nonexistence of Sobolev embeddings, different types of Sobolev inequalities, including the isoperimetric and the Nash inequality, and the difficult question of getting sharp constants in Sobolev inequalities.

Harmonic maps
Frédéric Hélein and John C. Wood

Harmonic maps are maps between Riemannian manifolds which extremize a natural energy functional or 'Dirichlet integral'. They include harmonic functions between Euclidean spaces, geodesics, minimal immersions, and harmonic morphisms (maps which preserve Laplace's equation). The Euler-Lagrange equations satisfied by a harmonic map form a semi-linear elliptic system of partial differential equations of second order.

We concentrate on the key questions of *existence*, *uniqueness* and *regularity* of harmonic maps between given manifolds. We survey some of the main methods of global analysis for answering these questions, together with the approach using twistor theory and integrable systems.

Topology of differentiable mappings
Kevin Houston

To study the topology of a differentiable mapping one can consider its image or its fibres. A proportion of this survey paper looks at how the latter can be studied in the case of singular complex analytic maps. An important aspect of this is study of the local case, the primary object of interest of which is the Milnor Fibre. More generally, Stratified Morse Theory is used to investigate the topology of singular spaces. In the complex case we can use rectified homotopical depth to generalize the Lefschetz Hyperplane Theorem. In the less studied case of images of maps we describe a powerful spectral sequence that can be used to investigate the homology of the image of a finite and proper map using the alternating homology of the multiple point spaces of the map.

Group actions and Hilbert's fifth problem
Sören Illman

In the first section of the article we consider Hilbert's fifth problem concerning Lie's theory of transformation groups. In his fifth problem Hilbert asks the following. Given a continuous action of a locally euclidean group G on a locally euclidean space M, can one choose coordinates in G and M so that the action is real analytic? We discuss the affirmative solutions given in Theorems 1.1 and 1.2, and also present known counterexamples to the general question posed by Hilbert. Theorem 1.1 is the celebrated result from 1952, due to Gleason, Montgomery and Zippin, which says that every locally euclidean group is a Lie group. Theorem 1.2 is a more recent result, due to the author, which says that every Cartan (thus in particular, every proper) C^s differentiable action, $1 \leq s \leq \infty$, of a Lie group G is equivalent to a real analytic action.

The remaining part of the article, Sections 2–18, is then used to give a complete, and to a very large extent selfcontained, proof of Theorem 1.2. This tour brings us into many different topics within the theory of transformation groups.

Exterior differential systems
Niky Kamran

We review the main existence theorems for integral manifolds of exterior differential sytems, with a special emphasis the Cartan-Kähler Theorem for involutive analytic exterior differential systems. These theorems are illustrated on a number of classical problems in differential geometry and contact geometry of differential equations. We also give an introduction to the Cartan-Kuranishi Prolongation Theorem, and to the characteristic cohomology of exterior differential systems.

Weil bundles as generalized jet spaces
Ivan Kolář

We generalize the concept of higher order velocity and we interpret a Weil bundle as the space of A-velocities for an arbitrary Weil algebra A. Using a recent identification of every product preserving bundle functor on manifolds with a Weil functor T^A, we deduce geometric results on T^A-prolongations of various geometric objects. We characterize every fiber product preserving bundle functor F on fibered manifolds in a jet-like manner and we study F-prolongations of several geometric structures.

Distributions, vector distributions, and immersions of manifolds in Euclidean spaces
Július Korbaš

Our main topics are *Schwartzian distributions* (including generalized sections of vector bundles), *vector distributions* (including the vector field problem), and *immersions* of smooth manifolds in Euclidean spaces (including isometric immersions). The Atiyah-Singer index theorem (covered by D. Bleecker's contribution in this Handbook) is a kind of node at which the three topics are joined.

Geometry of differential equations
Boris Kruglikov and Valentin Lychagin

In this contribution a review of geometric and algebraic methods for investigations of systems of partial differential equations is given. We begin with detailed description of the geometry of jet spaces, Lie transformations and pseudogroups. Investigation of topology of integral Grassmanians allows to study singularities of integral manifolds of Cartan distributions in jets and PDEs.

We consider the most general regular systems of different order PDEs and construct the corresponding algebra of non-linear differential operators and module of C-differential operators, as well as linearization and evolutionary differentiations. Then we discuss the bracket and multi-bracket formalism, which is applied to the theory of formal integrability. Here we mainly discuss computation of Spencer δ-cohomology and compatibility conditions, role of characteristics and various reductions of symbolic systems.

Finally we give a brief panorama of local solvability and integrability methods, describe formal approach to solution spaces, and discuss symmetries, auxiliary integrals and differential invariants. Spencer D-cohomology links the local aspects to the global theory and some methods for their evaluations together with examples are provided.

Part of the story we present is already classical, but we also contribute modern approach and recent results. The exposition is dense though this is compensated with quite a few references.

Global variational theory in fibred spaces
D. Krupka

We survey recent developments in the general theory of higher order, global integral variational functionals on fibred manifolds. First we study differential forms on jet prolongations of fibred manifolds. Then we introduce basic global concepts of the theory of Lagrange structures, such as the Lagrangian, the Lepage form, the Euler-Lagrange form, and characterize their properties in terms of differentiation and integration theory on manifolds. We study properties of the Euler-Lagrange mapping, and variational functionals, invariant with respect to automorphisms of underlying fibred manifolds. We finally discuss selected topics: examples of Lepage forms, the variational sequence and its consequences, possible formulations of the Hamilton theory, lifting functors and natural bundles, and properties of natural variational functionals on natural bundles. Remarks on the proofs of basic assertions are presented.

Second Order Ordinary Differential Equations in Jet Bundles and the Inverse Problem of the Calculus of Variations
O. Krupková and G. E. Prince

This article is ostensibly concerned with the inverse problem in the calculus of variations for a single independent variable: "when does a system of second order ordinary differential equations admit an equivalent variational formulation as a set of Euler-Lagrange equations?"

Because efforts to solve this famous 120 year old problem over the last 3 decades have involved the development of many entirely new geometric frameworks for second order ordinary differential equations, we firstly describe this underlying geometry, covering topics such as calculus on jet bundle prolongations of fibred manifolds, the geometric description of second order ordinary differential equations by both forms and vector fields, and of variational structures, globally and locally. We then turn to the inverse problem in both its covariant and contravariant forms and derive and discuss the famous Helmholtz conditions, being the necessary and sufficient conditions for the existence of a Lagrangian. We give the various geometric versions of the renowned work of Douglas who solved the problem for two degrees of freedom and review the latest progress on the problem for arbitrary degrees of freedom using exterior differential systems theory. The article is intended to provide a comprehensive introduction to the various aspects of current research.

Elements of noncommutative geometry
Giovanni Landi

We give an introduction to noncommutative geometry and its use. In the presented approach, a geometric space is given a spectral description as a triple $(\mathcal{A}, \mathcal{H}, D)$ consisting of a *-algebra $\mathcal{A}$ represented on a Hilbert space $\mathcal{H}$ together with an unbounded self-adjoint operator D interacting with the algebra in a bounded manner. The aim is to carry geometrical concepts over to a new class of spaces for which the algebra of functions $\mathcal{A}$ is noncommutative in general. We supplement the general theory with examples which include toric noncommutative spaces and spaces coming from quantum groups.

De Rham cohomology
M. A. Malakhaltsev

The aim of the paper is to give a brief introduction to the de Rham cohomology theory and to expose some relevant results in differential geometry. It includes the following topics: 1) De Rham complex. De Rham cohomology; 2) Integration and de Rham cohomology. De Rham currents. Harmonic forms; 3) Generalizations of the de Rham complex; 4) Equivariant de Rham cohomology; 5) Complexes of differential forms associated to differential geometric structures. No proofs are given, however the main statements are supplied with references to literature where the reader can find detailed exposition including proofs. The bibliography: 97 titles.

Topology of manifolds with corners
J. Margalef-Roig and E. Outerelo Domínguez

The study of manifolds with corners was originally developed by J. Cerf and A. Douady as a natural generalization of the concept of finite-dimensional manifold with smooth boundary, and applications of this type of manifolds in differential topology arose immediately after its definition. In the setting of Global Analysis a very natural task is to extend the results of finite-dimensional manifolds with corners to infinite dimensional manifolds. Thus, in this article, we survey the main features of the manifolds with corners modeled on Banach spaces or on larger categories of spaces as can be the normed spaces, the locally convex vector spaces and the convenient vector spaces, that have arisen as very important, in the last years.

Jet manifolds and natural bundles
D. J. Saunders

We introduce manifolds of jets, including jets of sections, jets of immersed submanifolds and higher-order velocities, and describe some of the geometrical structures which are canonically associated with these manifolds. As applications, we give brief introductions to the integrability theory for differential equations using Spencer cohomology, and to the calculus of variations and the associated variational complexes. We finally describe how jets may be used to characterise those operators having chart-independent coordinate representations by using the concepts of natural bundle and natural operator.

Some aspects of differential theories
József Szilasi and Rezső L. Lovas

As Serge Lang wrote, it is possible to lay down the foundations (and more beyond) for manifolds modeled on Banach or Hilbert spaces rather than finite dimensional spaces *at no extra cost*. In this article we briefly outline how the theory works if the model space is a more general locally convex (real) topological vector space, and the underlying differential calculus is the infinite-dimensional calculus initiated by A. D. Michal and A. Bastiani. In order to present at least one essential application of basic techniques of functional analysis, we discuss in detail a coordinate-free characterization of differential operators due to J. Peetre. In the last part we consider the covariant derivative operator discovered by S. S. Chern and H. Rund (independently), and which became an indispensable tool for present day Finsler geometry. We show that the Chern–Rund derivative can be interpreted as a differential operator on the *base manifold*.

Variational sequences
R. Vitolo

Variational sequences are complexes of modules or sheaf sequences in which one of the operations is the Euler-Lagrange operator, *i.e.*, the differential operator taking a Lagrangian into its Euler-Lagrange form, whose kernel is the Euler-Lagrange equation.

In this paper we present the most common approaches to variational sequences and discuss some directions of the current research on the topic.

The Oka-Grauert-Gromov principle for holomorphic bundles
Pit-Mann Wong

The Oka-Grauert-Gromov principle is a very powerful tool in the theory of holomorphic fiber bundles, more generally subelliptic bundles, over Stein spaces. The basic form asserts that such bundles must be holomorphically trivial if it is topologically trivial. This is particular useful in the study of holomorphic mappings from a Stein space into a projective variety. The situation is particularly nice in the case of hyperbolic geometry where the domain is the complex Euclidean space. The Oka-Grauert-Gromov principle in this case says that every subelliptic bundle is trivial. For example, the pull-back of the arc spaces or the parametrized jet bundles of any order are trivial and information of derivatives of any order can be dealt with using only classical function theory. The theory can even be extended to the case of Cartesian spaces defined over p-adic number fields. The basic form of Oka-Grauert-Gromov principle now is equivalent to the assertion that every projective module, over a ring p-adic convergent power series, is free. This is an analogue of the classical Serre's problem that every projective module, over a polynomial ring with coefficients in a field, is free. In the complex case, the principle can be applied to many important problems: submersion, immersion, embedding problems; extension problems; theory of complete intersections; homotopy theory, to name a few. The analogue of these more sophisticated problems of the principle are, as yet, to be explored in the p-adic case.

This page intentionally left blank

Index

This page intentionally left blank