

Graduate Texts in Physics

Günter Ludyk

# Einstein in Matrix Form

Exact Derivation of the Theory of Special  
and General Relativity without Tensors

 Springer

# Graduate Texts in Physics

For further volumes:  
[www.springer.com/series/8431](http://www.springer.com/series/8431)

# Graduate Texts in Physics

Graduate Texts in Physics publishes core learning/teaching material for graduate- and advanced-level undergraduate courses on topics of current and emerging fields within physics, both pure and applied. These textbooks serve students at the MS- or PhD-level and their instructors as comprehensive sources of principles, definitions, derivations, experiments and applications (as relevant) for their mastery and teaching, respectively. International in scope and relevance, the textbooks correspond to course syllabi sufficiently to serve as required reading. Their didactic style, comprehensiveness and coverage of fundamental material also make them suitable as introductions or references for scientists entering, or requiring timely knowledge of, a research field.

## *Series Editors*

Professor William T. Rhodes

Department of Computer and Electrical Engineering and Computer Science  
Imaging Science and Technology Center  
Florida Atlantic University  
777 Glades Road SE, Room 456  
Boca Raton, FL 33431  
USA  
[wrhodes@fau.edu](mailto:wrhodes@fau.edu)

Professor H. Eugene Stanley

Center for Polymer Studies Department of Physics  
Boston University  
590 Commonwealth Avenue, Room 204B  
Boston, MA 02215  
USA  
[hes@bu.edu](mailto:hes@bu.edu)

Professor Richard Needs

Cavendish Laboratory  
JJ Thomson Avenue  
Cambridge CB3 0HE  
UK  
[rn11@cam.ac.uk](mailto:rn11@cam.ac.uk)

Professor Martin Stutzmann

Walter Schottky Institut  
TU München  
85748 Garching  
Germany  
[stutz@wsi.tu-muenchen.de](mailto:stutz@wsi.tu-muenchen.de)

Günter Ludyk

# Einstein in Matrix Form

Exact Derivation of the Theory of Special  
and General Relativity without Tensors

 Springer

Professor Dr. Günter Ludyk  
Physics and Electrical Engineering  
University of Bremen  
Bremen, Germany  
[Guenter.Ludyk@nord-com.net](mailto:Guenter.Ludyk@nord-com.net)

ISSN 1868-4513

Graduate Texts in Physics

ISBN 978-3-642-35797-8

DOI 10.1007/978-3-642-35798-5

Springer Heidelberg New York Dordrecht London

ISSN 1868-4521 (electronic)

ISBN 978-3-642-35798-5 (eBook)

Library of Congress Control Number: 2013942312

© Springer-Verlag Berlin Heidelberg 2013

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media ([www.springer.com](http://www.springer.com))

*To my grandchildren Ann-Sophie  
and  
Alexander  
Hüttermann*

# Preface

This book is an introduction to the theories of Special and General Relativity. The target audience are physicists, engineers and applied scientists who are looking for an understandable introduction to the topic—without too much new mathematics.

All necessary mathematical tools are provided either directly in the text or in the appendices. Also the appendices contain an introduction to vector or matrices: first, as a refresher of known fundamental algebra, and second, to gain new experiences, e.g. with the Kronecker-product of matrices and differentiation with respect to vectors and matrices.

The fundamental equations of Einstein's theory of Special and General Relativity are derived using matrix calculus without the help of tensors. This feature makes the book special and a valuable tool for scientists and engineers with no experience in the field of tensor calculus. But physicists are also discovering that Einstein's vacuum field equations can be expressed as a system of first-order differential-matrix equations, wherein the unknown quantity is a matrix. These matrix equations are also easy to handle when implementing numerical algorithms using standard software as, e.g. MATHEMATICA or MAPLE.

In Chap. 1, the foundations of Special Relativity are developed. Chapter 2 describes the structure and principles of General Relativity. Chapter 3 explains the Schwarzschild solution of spherical body gravity and examines the “Black Hole” phenomenon. Furthermore, two appendices summarize the basics of the matrix theory and differential geometry.

After completion of the book, I discovered the paper [37], where Einstein's equations of a similar shape are derived.

I would like to thank *Claus Ascheron* (Springer) who has made great effort towards the publication of this book. Finally, I would like to thank my wife *Renate*, without her this book would have never been published!

Bremen

Günter Ludyk

# Contents

<b>1</b>	<b>Special Relativity</b>	<b>1</b>
1.1	Galilei Transformation	1
1.1.1	Relativity Principle of Galilei	1
1.1.2	General Galilei Transformation	5
1.1.3	Maxwell's Equations and Galilei Transformation	6
1.2	Lorentz Transformation	7
1.2.1	Introduction	7
1.2.2	Determining the Components of the Transformation Matrix	8
1.2.3	Simultaneity at Different Places	12
1.2.4	Length Contraction of Moving Bodies	13
1.2.5	Time Dilation	15
1.3	Invariance of the Quadratic Form	15
1.3.1	Invariance with Respect to Lorentz Transformation	17
1.3.2	Light Cone	17
1.3.3	Proper Time	19
1.4	Relativistic Velocity Addition	20
1.4.1	Galilean Addition of Velocities	20
1.5	Lorentz Transformation of the Velocity	22
1.6	Momentum and Its Lorentz Transformation	25
1.7	Acceleration and Force	26
1.7.1	Acceleration	26
1.7.2	Equation of Motion and Force	28
1.7.3	Energy and Rest Mass	30
1.7.4	Emission of Energy	31
1.8	Relativistic Electrodynamics	32
1.8.1	Maxwell's Equations	32
1.8.2	Lorentz Transformation of the Maxwell's Equations	34
1.8.3	Electromagnetic Invariants	37
1.8.4	Electromagnetic Forces	39



1.9	The Energy–Momentum Matrix . . . . .	41
1.9.1	The Electromagnetic Energy–Momentum Matrix . . . . .	41
1.9.2	The Mechanical Energy–Momentum Matrix . . . . .	43
1.9.3	The Total Energy–Momentum Matrix . . . . .	47
1.10	The Most Important Definitions and Formulas in Special Relativity . . . . .	48
<b>2</b>	<b>Theory of General Relativity . . . . .</b>	<b>51</b>
2.1	General Relativity and Riemannian Geometry . . . . .	51
2.2	Motion in a Gravitational Field . . . . .	53
2.2.1	First Solution . . . . .	54
2.2.2	Second Solution . . . . .	55
2.2.3	Relation Between $\tilde{\Gamma}$ and $\mathbf{G}$ . . . . .	56
2.3	Geodesic Lines and Equations of Motion . . . . .	57
2.3.1	Alternative Geodesic Equation of Motion . . . . .	62
2.4	Example: Uniformly Rotating Systems . . . . .	64
2.5	General Coordinate Transformations . . . . .	67
2.5.1	Absolute Derivatives . . . . .	67
2.5.2	Transformation of the Christoffel Matrix $\tilde{\Gamma}$ . . . . .	69
2.5.3	Transformation of the Christoffel Matrix $\hat{\Gamma}$ . . . . .	71
2.5.4	Coordinate Transformation and Covariant Derivative . . . . .	72
2.6	Incidental Remark . . . . .	76
2.7	Parallel Transport . . . . .	78
2.8	Riemannian Curvature Matrix . . . . .	80
2.9	Properties of the Riemannian Curvature Matrix . . . . .	81
2.9.1	Composition of $\mathbf{R}$ and $\hat{\mathbf{R}}$ . . . . .	81
2.10	The Ricci Matrix and Its Properties . . . . .	88
2.10.1	Symmetry of the Ricci Matrix $\mathbf{R}_{\text{Ric}}$ . . . . .	90
2.10.2	The Divergence of the Ricci Matrix . . . . .	92
2.11	General Theory of Gravitation . . . . .	93
2.11.1	The Einstein’s Matrix $\mathfrak{G}$ . . . . .	93
2.11.2	Newton’s Theory of Gravity . . . . .	94
2.11.3	The Einstein’s Equation with $\mathfrak{G}$ . . . . .	97
2.12	Summary . . . . .	99
2.12.1	Covariance Principle . . . . .	99
2.12.2	Einstein’s Field Equation and Momentum . . . . .	101
2.13	Hilbert’s Action Functional . . . . .	101
2.13.1	Effects of Matter . . . . .	105
2.14	Most Important Definitions and Formulas . . . . .	106
<b>3</b>	<b>Gravitation of a Spherical Mass . . . . .</b>	<b>109</b>
3.1	Schwarzschild’s Solution . . . . .	109
3.1.1	Christoffel Matrix $\Gamma$ . . . . .	110
3.1.2	Ricci Matrix $\mathbf{R}_{\text{Ric}}$ . . . . .	112
3.1.3	The Factors $A(r)$ and $B(r)$ . . . . .	114
3.2	Influence of a Massive Body on the Environment . . . . .	116
3.2.1	Introduction . . . . .	116

- 3.2.2 Changes to Length and Time . . . . . 117
- 3.2.3 Redshift of Spectral Lines . . . . . 118
- 3.2.4 Deflection of Light . . . . . 120
- 3.3 Schwarzschild’s Inner Solution . . . . . 124
- 3.4 Black Holes . . . . . 127
  - 3.4.1 Astrophysics . . . . . 127
  - 3.4.2 Further Details about “Black Holes” . . . . . 129
  - 3.4.3 Singularities . . . . . 131
  - 3.4.4 Eddington’s Coordinates . . . . . 135
- 3.5 Rotating Masses . . . . . 138
  - 3.5.1 Ansatz for the Metric Matrix  $\mathbf{G}$  . . . . . 138
  - 3.5.2 Kerr’s Solution in Boyer–Lindquist Coordinates . . . . . 139
  - 3.5.3 The Lense–Thirring Effect . . . . . 139
- 3.6 Summary of Results for the Gravitation of a Spherical Mass . . . . . 141
- 3.7 Concluding Remark . . . . . 143
- Appendix A Vectors and Matrices . . . . . 145**
  - A.1 Vectors and Matrices . . . . . 145
  - A.2 Matrices . . . . . 147
    - A.2.1 Types of Matrices . . . . . 147
    - A.2.2 Matrix Operations . . . . . 148
    - A.2.3 Block Matrices . . . . . 152
  - A.3 The Kronecker-Product . . . . . 154
    - A.3.1 Definitions . . . . . 154
    - A.3.2 Some Theorems . . . . . 154
    - A.3.3 The Permutation Matrix  $\mathbf{U}_{p \times q}$  . . . . . 156
    - A.3.4 More Properties of the Kronecker-Product . . . . . 157
  - A.4 Derivatives of Vectors/Matrices with Respect to Vectors/Matrices . . . . . 157
    - A.4.1 Definitions . . . . . 157
    - A.4.2 Product Rule . . . . . 158
    - A.4.3 Chain Rule . . . . . 159
  - A.5 Differentiation with Respect to Time . . . . . 159
    - A.5.1 Differentiation of a Function with Respect to Time . . . . . 159
    - A.5.2 Differentiation of a Vector with Respect to Time . . . . . 160
    - A.5.3 Differentiation of a  $2 \times 3$ -Matrix with Respect to Time . . . . . 161
    - A.5.4 Differentiation of an  $n \times m$ -Matrix with Respect to Time . . . . . 161
  - A.6 Supplements to Differentiation with Respect to a Matrix . . . . . 162
- Appendix B Some Differential Geometry . . . . . 165**
  - B.1 Curvature of a Curved Line in Three Dimensions . . . . . 165
  - B.2 Curvature of a Surface in Three Dimensions . . . . . 166
    - B.2.1 Vectors in the Tangent Plane . . . . . 166
    - B.2.2 Curvature and Normal Vectors . . . . . 168
    - B.2.3 Theorema Egregium and the Inner Values  $g_{ij}$  . . . . . 169

<b>Appendix C Geodesic Deviation</b> . . . . .	179
<b>Appendix D Another Ricci-Matrix</b> . . . . .	183
<b>References</b> . . . . .	189
<b>Index</b> . . . . .	191

# Notation

*Important definitions, facts and theorems* are framed, and important *intermediate results* are double-underlined.

*Scalars* are written in normal typeface:

$$a, b, c, \alpha, \beta, \gamma, \dots;$$

*Vectors* in the 3-dimensional space ( $\mathbb{R}^3$ ) are written in small bold typeface:

$$\mathbf{x}, \mathbf{v}, \mathbf{u}, \mathbf{a}, \dots;$$

*Vectors* in the 4-dimensional spacetime ( $\mathbb{R}^4$ ) are written in small bold typeface with an arrow:

$$\vec{\mathbf{x}}, \vec{\mathbf{v}}, \vec{\mathbf{u}}, \vec{\mathbf{a}}, \dots;$$

*Matrices* are written in big bold typeface:

$$\mathbf{M}, \mathbf{G}, \mathbf{R}, \mathbf{I}, \dots$$

The identity matrix  $\mathbf{I}_n$  of size  $n \times n$  is the matrix in which all the elements on the main diagonal are equal to 1 and all other elements are equal to 0, e.g.,

$$\mathbf{I}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The *Derivative Operator*  $\nabla$  is the 3-dimensional column vector

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix},$$

and  $\nabla^\top$  is the 3-dimensional row-vector

$$\nabla^\top = \left( \frac{\partial}{\partial x} \mid \frac{\partial}{\partial y} \mid \frac{\partial}{\partial z} \right).$$

The *Derivative Operator*  $\vec{\nabla}$  is the 4-dimensional column vector

$$\vec{\nabla} = \gamma \begin{pmatrix} -\frac{1}{c} \frac{\partial}{\partial t} \\ \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix}$$

with

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

*Remark* The derivative operators  $\nabla$  and  $\vec{\nabla}$  are column vector operators and can act both on the right and on the left! Example:

$$\vec{\nabla}^\top \vec{a} = \vec{a}^\top \vec{\nabla} = \gamma \left( -\frac{1}{c} \frac{\partial a_0}{\partial t} + \frac{\partial a_1}{\partial x} + \frac{\partial a_2}{\partial y} + \frac{\partial a_3}{\partial z} \right).$$

# Chapter 1

## Special Relativity

This chapter begins with the classical theorems of Galilei and Newton and the Galilei transformation. The special theory of relativity, developed by Einstein in 1905, leads to the four-dimensional spacetime of Minkowski and the Lorentz transformation. After that the relativity of simultaneity of events, the length contraction of moving bodies and the time dilation are discussed. This is followed by the velocity-addition formula and relativistic mechanics. The next topic is the mass–energy equivalence formula  $E = mc^2$ , where  $c$  is the speed of light in a vacuum. Then relativistic electromagnetism is treated and the invariance of special forms of the equations of dynamics and Maxwell’s electrodynamic with respect to the Lorentz transformation is shown. The energy–momentum matrix is introduced and discussed.

### 1.1 Galilei Transformation

#### 1.1.1 Relativity Principle of Galilei

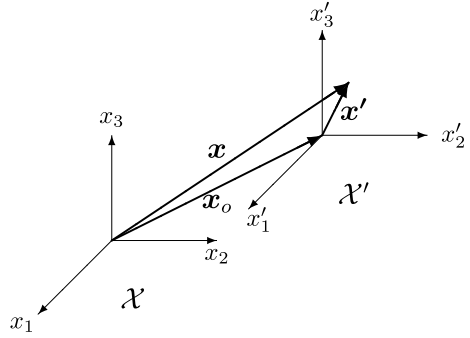
An *event* is anything that can happen in space and time, e.g. the emission of a flash of light in a room corner. Events happen at a single point. We assign to each event a set of four coordinates  $t, x_1, x_2$  and  $x_3$ , or with  $t$  and the three-dimensional vector

$$\mathbf{x} \stackrel{\text{def}}{=} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3.$$

The position vector  $\mathbf{x}$  and the time  $t$  form a *reference frame*  $\mathcal{X}$ . In this frame, Newton’s fundamental law of mechanics has the form

$$\frac{d\mathbf{p}}{dt} = \mathbf{f}$$

**Fig. 1.1** Two against each other shifted reference systems



or, if the mass  $m$  in the momentum

$$\mathbf{p} = m \frac{d\mathbf{x}}{dt}$$

is constant,

$$m \frac{d^2\mathbf{x}}{dt^2} = \mathbf{f}. \quad (1.1)$$

An observer may now execute any motion, for example, he makes an experiment in a moving train. We want to find the equation that takes the place of

$$\frac{d\mathbf{p}}{dt} = \mathbf{f}$$

for the moving observer. A coordinate system  $\mathcal{X}'$  is connected firmly with the moving observer; it should be axis-parallel to the original coordinate system  $\mathcal{X}$ .  $\mathbf{x}_o$  is the location of the origin of  $\mathcal{X}'$  measured in  $\mathcal{X}$  (Fig. 1.1). Then

$$\mathbf{x} = \mathbf{x}_o + \mathbf{x}',$$

or

$$\mathbf{x}' = \mathbf{x} - \mathbf{x}_o. \quad (1.2)$$

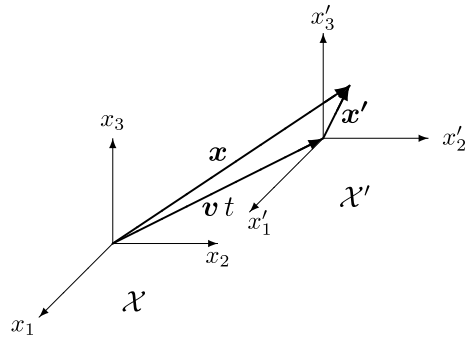
Here  $\mathbf{x}$  is the position vector of the event measured by an observer at rest in the reference system  $\mathcal{X}$ , and  $\mathbf{x}'$  is what an observer measures in the moving reference system  $\mathcal{X}'$ . Equation (1.2) differentiated with respect to time  $t$ ,

$$\frac{d\mathbf{x}'}{dt} = \frac{d\mathbf{x}}{dt} - \frac{d\mathbf{x}_o}{dt}, \quad (1.3)$$

results in the speed addition theorem of classical mechanics:

$$\mathbf{v}'(t) = \mathbf{v}(t) - \mathbf{v}_0(t).$$

**Fig. 1.2** Two reference systems moving against each other



For the acceleration one obtains

$$\frac{d^2 \mathbf{x}'}{dt^2} = \frac{d^2 \mathbf{x}}{dt^2} - \frac{d^2 \mathbf{x}_o}{dt^2}. \quad (1.4)$$

The force  $\mathbf{f}$  acting on the mass  $m$  is independent of the chosen coordinate system, so  $\mathbf{f}' = \mathbf{f}$ . This and (1.1) used in (1.4) result in

$$\mathbf{f}' - m \frac{d^2 \mathbf{x}_o}{dt^2} = m \frac{d^2 \mathbf{x}'}{dt^2}. \quad (1.5)$$

The fundamental law of mechanics has lost its validity! If the moving observer knows the external force  $\mathbf{f}'$ , he can determine with measurements in  $\mathcal{X}'$  his acceleration with respect to the rest system  $\mathcal{X}$ . However, if the motion of  $\mathcal{X}'$  with respect to  $\mathcal{X}$  is uniform and rectilinear, i.e.  $\mathbf{x}_o = \mathbf{v}_o t$  with a constant  $\mathbf{v}_o$ , then from (1.5)

$$\mathbf{f}' = m \frac{d^2 \mathbf{x}'}{dt^2} \quad (1.6)$$

and the fundamental laws of mechanics have in  $\mathcal{X}'$  the same form as in  $\mathcal{X}$ . Such a uniform and rectilinear moving coordinate system is called an *inertial system*. The moving observer has no possibility to determine his own motion with respect to the coordinate system  $\mathcal{X}$  by a mechanical experiment. Any free particle moves in a straight line with constant speed. For example, Galilei considered a uniformly moving ship in a port, whose occupants cannot decide whether the ship moves with respect to the port or whether the port moves with respect to the ship. Today as an example one would take a train in a railway station. This is the *Relativity Principle* of Galilei:

**Axiom: All natural laws are the same at every moment in all inertial systems.**  
**All coordinate systems moving uniformly linearly with respect to an inertial system are themselves inertial systems.**



If two reference systems  $\mathcal{X}$  and  $\mathcal{X}'$  move with a constant velocity  $\mathbf{v}$  against each other (Fig. 1.2), then

$$\mathbf{x}' = \mathbf{x} - \mathbf{v}t \quad \text{and} \quad t' = t. \quad (1.7)$$

With the four-dimensional column vector

$$\vec{\mathbf{x}} \stackrel{\text{def}}{=} \begin{pmatrix} t \\ \mathbf{x} \end{pmatrix} \in \mathbb{R}^4$$

the two equations (1.7) can be summarized in one equation, and we get the Galilei transformation in a matrix form

$$\vec{\mathbf{x}}' = \begin{pmatrix} 1 & \mathbf{o}^T \\ -\mathbf{v} & \mathbf{I} \end{pmatrix} \vec{\mathbf{x}} = \mathbf{T}_{\text{Galilei}} \vec{\mathbf{x}}. \quad (1.8)$$

Looking at the inverse transformation of  $\mathcal{X}'$  to  $\mathcal{X}$ , one gets

$$t = t' \quad \text{and} \quad \mathbf{x} = \mathbf{v}t + \mathbf{x}',$$

or

$$\vec{\mathbf{x}} = \begin{pmatrix} 1 & \mathbf{o}^T \\ \mathbf{v} & \mathbf{I} \end{pmatrix} \vec{\mathbf{x}}' = \mathbf{T}'_{\text{Galilei}} \vec{\mathbf{x}}'.$$

If both transformations are applied in series, the result is

$$\mathbf{T}_{\text{Galilei}} \mathbf{T}'_{\text{Galilei}} \vec{\mathbf{x}}' = \begin{pmatrix} 1 & \mathbf{o}^T \\ -\mathbf{v} & \mathbf{I} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{o}^T \\ \mathbf{v} & \mathbf{I} \end{pmatrix} \vec{\mathbf{x}}' = \begin{pmatrix} 1 & \mathbf{o}^T \\ \mathbf{o} & \mathbf{I} \end{pmatrix} \vec{\mathbf{x}}' = \vec{\mathbf{x}}',$$

i.e. the matrix  $\mathbf{T}'_{\text{Galilei}}$  is the inverse of the matrix  $\mathbf{T}_{\text{Galilei}}$ .

For the time derivatives of the four-vector  $\vec{\mathbf{x}}$  we obtain

$$\frac{d\vec{\mathbf{x}}}{dt} = \begin{pmatrix} 1 \\ \frac{d\mathbf{x}}{dt} \end{pmatrix} \quad \text{and} \quad \frac{d^2\vec{\mathbf{x}}}{dt^2} = \begin{pmatrix} 0 \\ \frac{d^2\mathbf{x}}{dt^2} \end{pmatrix}.$$

With  $\vec{\mathbf{f}} \stackrel{\text{def}}{=} \begin{pmatrix} 0 \\ \mathbf{f} \end{pmatrix}$  we can write the fundamental equation of mechanics

$$\vec{\mathbf{f}} = m \frac{d^2\vec{\mathbf{x}}}{dt^2}. \quad (1.9)$$

This equation multiplied from the left with the transformation matrix  $\mathbf{T}_{\text{Galilei}}$ , in fact, gives back the same form:

$$\underbrace{\mathbf{T}_{\text{Galilei}} \vec{\mathbf{f}}}_{\stackrel{\text{def}}{=} \vec{\mathbf{f}}'} = m \underbrace{\mathbf{T}_{\text{Galilei}} \frac{d^2\vec{\mathbf{x}}}{dt^2}}_{\stackrel{\text{def}}{=} \frac{d^2\vec{\mathbf{x}}'}{dt'^2}}, \quad (1.10)$$

$$\vec{\mathbf{f}}' = m \frac{d^2\vec{\mathbf{x}}'}{dt'^2}.$$

The fundamental equation of dynamics is thus *invariant* with respect to this Galilei transformation, i.e. it retains its shape regardless of the reference system.

### 1.1.2 General Galilei Transformation

Up to this point, the Galilei transformation was considered only under the uniform motion with the speed  $v$  of the two inertial systems against each other and a fixed initial time  $t_o = 0$ , a fixed initial point  $\mathbf{x}_0 = \mathbf{0}$  of the new coordinate system  $\mathcal{X}'$  and with no rotation of the coordinate system.

It is now generally assumed that all laws of nature remain constant, therefore, are *invariant* with respect to *time shifting*. If  $\mathbf{x}(t)$  is a solution of  $m\ddot{\mathbf{x}} = \mathbf{f}$ , then for all  $t_o \in \mathbb{R}$  is  $\mathbf{x}(t + t_o)$  also a solution.

Next, it is assumed that the considered spaces are *homogeneous*, thus the same features are available at all points. So, if  $\mathbf{x}(t)$  is again a solution of  $m\ddot{\mathbf{x}} = \mathbf{f}$ , then  $\mathbf{x}(t) + \mathbf{b}$  is also a solution, but now for the starting point  $\mathbf{x}_0 + \mathbf{b}$ .

Moreover, it is assumed that the considered spaces are *isotropic*, i.e. there is no directional dependence of properties. So again, if  $\mathbf{x}(t)$  is a solution of  $m\ddot{\mathbf{x}} = \mathbf{f}$ , then also  $D\mathbf{x}(t)$  is a solution for the initial state  $D\mathbf{x}_0$ . Here, however, also the equality of the distances

$$\rho(\mathbf{x}_1, \mathbf{x}_2) = \rho(D\mathbf{x}_1, D\mathbf{x}_2)$$

must be true. For the rotation matrix  $D$  this means that it must be *orthogonal* because

$$\begin{aligned} \rho(D\mathbf{x}_1, D\mathbf{x}_2) &= \sqrt{(D\mathbf{x}_2 - D\mathbf{x}_1)^\top (D\mathbf{x}_2 - D\mathbf{x}_1)} \\ &= \sqrt{(\mathbf{x}_2 - \mathbf{x}_1)^\top D^\top D (\mathbf{x}_2 - \mathbf{x}_1)} \stackrel{!}{=} \rho(\mathbf{x}_1, \mathbf{x}_2) \\ &= \sqrt{(\mathbf{x}_2 - \mathbf{x}_1)^\top (\mathbf{x}_2 - \mathbf{x}_1)}, \end{aligned}$$

so

$$D^\top D \stackrel{!}{=} I.$$

The time invariance, homogeneity and isotropy can be summarized in one transformation, the *general Galilei transformation*, as follows:

If  $t'$  is shifted versus  $t$  by  $t_0$ , i.e. one has

$$t' = t_0 + t, \tag{1.11}$$

furthermore, the new coordinate system relative to the old is moved by  $\mathbf{x}_0$  and rotated by the rotation matrix  $D$ , so for  $\mathbf{v} = \mathbf{0}$  one gets

$$\mathbf{x}' = D\mathbf{x} + \mathbf{x}_0, \tag{1.12}$$

and so

$$\begin{pmatrix} t' \\ \mathbf{x}' \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{o}^\top \\ \mathbf{o} & \mathbf{D} \end{pmatrix} \begin{pmatrix} t \\ \mathbf{x} \end{pmatrix} + \begin{pmatrix} t_o \\ \mathbf{x}_o \end{pmatrix}. \quad (1.13)$$

By moving the origin of the new coordinate system with the velocity  $\mathbf{v}$  as mentioned above, one finally gets the general Galilei transformation:

$$\vec{\mathbf{x}}' = \begin{pmatrix} 1 & \mathbf{o}^\top \\ -\mathbf{v} & \mathbf{D} \end{pmatrix} \vec{\mathbf{x}} + \vec{\mathbf{x}}_o, \quad (1.14)$$

with

$$\vec{\mathbf{x}} \stackrel{\text{def}}{=} \begin{pmatrix} t \\ \mathbf{x} \end{pmatrix} \quad \text{and} \quad \vec{\mathbf{x}}_o \stackrel{\text{def}}{=} \begin{pmatrix} t_o \\ \mathbf{x}_o \end{pmatrix}.$$

This is an *affine mapping*, or an *affine transformation*.

One gets a *linear transformation* by introducing the extended vector:

$$\begin{pmatrix} t \\ \mathbf{x} \\ 1 \end{pmatrix} \in \mathbb{R}^5, \quad (1.15)$$

namely

$$\begin{pmatrix} t' \\ \mathbf{x}' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{o}^\top & t_o \\ -\mathbf{v} & \mathbf{D} & \mathbf{x}_o \\ 0 & \mathbf{o}^\top & 1 \end{pmatrix} \begin{pmatrix} t \\ \mathbf{x} \\ 1 \end{pmatrix}. \quad (1.16)$$

Newton's fundamental laws are invariant also with respect to such a transformation. A general Galilei transformation is determined by 10 parameters:  $t_o$ ,  $\mathbf{x}_o \in \mathbb{R}^3$ ,  $\mathbf{v} \in \mathbb{R}^3$  and  $\mathbf{D} \in \mathbb{R}^{3 \times 3}$ . The rotation matrix  $\mathbf{D}$ , in fact, has only three main parameters since any general rotation consists of successive rotations performed around the  $x_1$ -,  $x_2$ - and  $x_3$ -axis, so the whole rotation is characterized by the three angles  $\varphi_1$ ,  $\varphi_2$  and  $\varphi_3$ , where, for example, the rotation around the  $x_1$ -axis is achieved by the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi_1 & \sin \varphi_1 \\ 0 & -\sin \varphi_1 & \cos \varphi_1 \end{pmatrix}.$$

### 1.1.3 Maxwell's Equations and Galilei Transformation

The situation is completely different with Maxwell's equations of electromagnetic dynamics. They are not invariant with respect to a Galilei transformation! Indeed, in an inertial system  $\mathcal{X}$ , a static charge  $q$  generates only a static electric field; in an inertial frame, moving with the speed  $\mathbf{v}$ ,  $q\mathbf{v}$  is an electric current which generates there a magnetic field!

In the nineteenth century, it was believed that all physical phenomena are mechanical and electromagnetic forces could be traced to the stress states of a world-aether, the Maxwell's tensions. It was assumed that even a vacuum must be filled with aether. This aether then is the carrier of the electromagnetic phenomena.

Suppose an *inertial frame* is a reference system in which Galilei's principle of inertia is valid. Then Einstein in his *general relativity principle* claims:

**The laws of nature take on the same form in all inertial systems.**

For the fundamental law of mechanics, this is derived above. The principle of relativity applies neither to electrodynamics nor to optics. How should the basic equations of electrodynamics be modified so that the relativity principle is valid? This is the content of Einstein's *Theory of Special Relativity*. He expanded it in the *Theory of General Relativity*. This theory treats how the natural laws must be modified so that they are also valid in *accelerating* or not uniformly against each other moving reference systems.

## 1.2 Lorentz Transformation

### 1.2.1 Introduction

At the end of the nineteenth century, experiments were conceived which should determine the velocity of the Earth with respect to the resting cosmic aether. This speed relative to the aether can only be measured by an electromagnetic effect, e.g. the light wave propagation. But in the Michelson–Morley experiment in 1881 and 1887 no drift velocity was found. Einstein concluded:<sup>1</sup>

**The speed of light  $c$  is always constant.**

Independently of the movement of the light source and the observer, light has the same speed value in every inertial frame.

Assume a light pulse is produced at time  $t = t' = 0$  in the two axis-parallel reference systems  $\mathcal{X}$  and  $\mathcal{X}'$  with a common origin. If the light spreads with the speed of light  $c$  in the reference system  $\mathcal{X}'$ , then, for example, it is true that  $x'_1 = c \cdot t$ . From (1.7) it follows for the  $x_1$ -direction that if  $\mathbf{v}$  has the  $x_1$ -directional component, then

$$x_1 = x'_1 + vt = (c + v)t,$$

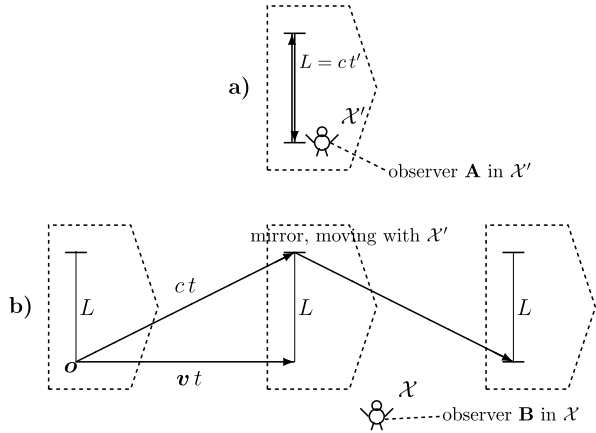
yielding, contrary to the Michelson–Morley experiment, a propagation velocity for light of  $c + v > c$ . Therefore, a transformation must be valid, different from the Galilei transformation. We try a linear transformation:

$$t = ft' + \mathbf{e}^T \mathbf{x}', \quad (1.17)$$

---

<sup>1</sup> $c$  from Latin *celeritas*: speed.

**Fig. 1.3** Different ways of light: (a) seen by an observer in  $\mathcal{X}'$ , (b) seen by an observer in  $\mathcal{X}$



$$\mathbf{x} = \mathbf{b}t' + \mathbf{A}\mathbf{x}', \quad (1.18)$$

i.e.

$$\vec{\mathbf{x}} = \hat{\mathbf{L}}'\vec{\mathbf{x}}' \in \mathbb{R}^4,$$

with

$$\hat{\mathbf{L}}' \stackrel{\text{def}}{=} \begin{pmatrix} f & \mathbf{e}^\top \\ \mathbf{b} & \mathbf{A} \end{pmatrix} \in \mathbb{R}^{4 \times 4}.$$

## 1.2.2 Determining the Components of the Transformation Matrix

That  $t'$  is different from  $t$  (according to Galilei this was not the case, he took  $t' = t$ ) is shown by the following reasoning. We consider two observers; observer **A** moves relative to the observer **B** in a spaceship with velocity  $\mathbf{v}$  (Fig. 1.3). The spaceship with observer **A** has the inertial system  $\mathcal{X}'$ , and the observer **B** on Earth has the inertial system  $\mathcal{X}$ . A beam of light moves from the origin  $\mathbf{x} = \mathbf{x}' = \mathbf{o}$  of the reference systems  $\mathcal{X}$  and  $\mathcal{X}'$  at time  $t = t' = 0$  perpendicular to the velocity  $\mathbf{v}$  and reaches after  $t'$  seconds for the observer **A** in the moving reference system  $\mathcal{X}'$  a mirror moving also with the reference system  $\mathcal{X}'$ . For the observer **B** in the stationary reference system  $\mathcal{X}$ , the light beam reaches the mirror, which has been moved in the  $\mathbf{v}$  direction a distance of  $\mathbf{v} \cdot t$  after  $t$  seconds. Since in all initial systems the speed of light is constant and equal to  $c$ , by Pythagorean theorem, one has

$$(ct)^2 = (vt)^2 + (L)^2 = (vt)^2 + (ct')^2,$$

or, after solving for  $t$ ,

$$\underline{\underline{t = \gamma t'}} \quad (1.19)$$

with

$$\gamma \stackrel{\text{def}}{=} \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}. \quad (1.20)$$

A comparison of (1.19) with (1.17) provides for  $\mathbf{x}' = \mathbf{o}$

$$\underline{\underline{f}} = \gamma. \quad (1.21)$$

For  $\mathbf{x}' = \mathbf{o}$  one has  $\mathbf{x} = \mathbf{v}t$ , and from (1.18) it follows that  $\mathbf{x} = \mathbf{b}t'$ . Thus,  $\mathbf{b}t' = \mathbf{v}t$ , i.e.  $\mathbf{b} = \mathbf{v}\frac{t}{t'}$ . From (1.19) it follows, on the other hand, that  $\frac{t}{t'} = \gamma$ , so

$$\underline{\underline{\mathbf{b}}} = \gamma \mathbf{v}. \quad (1.22)$$

Up to now, the following transformation equations were determined:

$$t = \gamma t' + \mathbf{e}^\top \mathbf{x}', \quad (1.23)$$

$$\mathbf{x} = \gamma \mathbf{v}t' + \mathbf{A}\mathbf{x}'. \quad (1.24)$$

Equations (1.23) and (1.24) deliver a transformation of  $\mathcal{X}'$  to  $\mathcal{X}$ . If one wants to invert this transformation, one must replace  $\mathbf{v}$  by  $-\mathbf{v}$ ,  $\mathbf{x}$  by  $\mathbf{x}'$ , and vice versa, and  $t$  by  $t'$  and vice versa (as  $\mathbf{A}$  and  $\mathbf{e}$  may depend on  $\mathbf{v}$ , in the following we first write  $\tilde{\mathbf{A}}$  and  $\tilde{\mathbf{e}}$ ):

$$t' = \gamma t + \tilde{\mathbf{e}}^\top \mathbf{x},$$

$$\mathbf{x}' = -\gamma \mathbf{v}t + \tilde{\mathbf{A}}\mathbf{x},$$

which can be combined to

$$\vec{\mathbf{x}}' = \begin{pmatrix} \gamma & \tilde{\mathbf{e}}^\top \\ -\gamma \mathbf{v} & \tilde{\mathbf{A}} \end{pmatrix} \vec{\mathbf{x}} \stackrel{\text{def}}{=} \hat{\mathbf{L}}\vec{\mathbf{x}}. \quad (1.25)$$

Both transformations performed one after another must result in the identity matrix:

$$\hat{\mathbf{L}}' \hat{\mathbf{L}} \stackrel{!}{=} \mathbf{I}. \quad (1.26)$$

For the top left (1, 1)-element of the matrix product  $\hat{\mathbf{L}}' \hat{\mathbf{L}}$  one gets:

$$(\gamma, \mathbf{e}^\top) \begin{pmatrix} \gamma \\ -\gamma \mathbf{v} \end{pmatrix} = \gamma^2 - \gamma \mathbf{e}^\top \mathbf{v} \stackrel{!}{=} 1.$$

Hence,

$$\gamma \mathbf{e}^\top \mathbf{v} = \gamma^2 - 1. \quad (1.27)$$

Taking for  $\mathbf{e}$

$$\mathbf{e}^\top = \alpha \mathbf{v}^\top \quad (1.28)$$

and using (1.27) yields

$$\gamma \alpha v^2 = \gamma^2 - 1,$$

and moreover,

$$\gamma \alpha = \frac{1}{v^2} \left( \frac{c^2}{c^2 - v^2} - 1 \right) = \frac{1}{c^2 - v^2} = \frac{\gamma^2}{c^2},$$

i.e.

$$\alpha = \frac{\gamma}{c^2}. \quad (1.29)$$

Equation (1.29) used in (1.28) finally yields

$$\underline{\underline{\mathbf{e}^\top = \frac{\gamma}{c^2} \mathbf{v}^\top}}. \quad (1.30)$$

Thus, till now we have calculated:

$$\hat{\mathbf{L}}' = \begin{pmatrix} \gamma & \frac{\gamma}{c^2} \mathbf{v}^\top \\ \gamma \mathbf{v} & \mathbf{A} \end{pmatrix}.$$

Obviously,

$$\tilde{\mathbf{e}}^\top = -\frac{\gamma}{c^2} \mathbf{v}^\top.$$

Suppose now that in (1.25) we have  $\tilde{\mathbf{A}} = \mathbf{A}$ . Then for the matrix element in the lower right corner of the matrix product  $\hat{\mathbf{L}}' \hat{\mathbf{L}}$  in (1.26) we obtain

$$(\gamma \mathbf{v}, \mathbf{A}) \begin{pmatrix} -\frac{\gamma}{c^2} \mathbf{v}^\top \\ \mathbf{A} \end{pmatrix} = -\frac{\gamma^2}{c^2} \mathbf{v} \mathbf{v}^\top + \mathbf{A}^2 \stackrel{!}{=} \mathbf{I},$$

i.e.

$$\mathbf{A}^2 = \mathbf{I} + \frac{\gamma^2}{c^2} \mathbf{v} \mathbf{v}^\top. \quad (1.31)$$

From (1.27), by inserting (1.30), follows

$$\frac{\gamma^2}{c^2} v^2 = \gamma^2 - 1.$$

Plugging this into (1.31) provides

$$\mathbf{A}^2 = \mathbf{I} + (\gamma^2 - 1) \frac{\mathbf{v} \mathbf{v}^\top}{v^2}. \quad (1.32)$$

Since  $(\gamma - 1)^2 = \gamma^2 - 2(\gamma - 1) - 1$ , for  $\gamma^2 - 1$  one can write

$$\gamma^2 - 1 = (\gamma - 1)^2 + 2(\gamma - 1). \quad (1.33)$$

Equation (1.33) inserted into (1.32) yields

$$\mathbf{A}^2 = \mathbf{I} + 2(\gamma - 1) \frac{\mathbf{v}\mathbf{v}^\top}{v^2} + (\gamma - 1)^2 \frac{\mathbf{v}\mathbf{v}^\top}{v^2} = \left( \mathbf{I} + (\gamma - 1) \frac{\mathbf{v}\mathbf{v}^\top}{v^2} \right)^2,$$

where

$$\frac{\mathbf{v}\mathbf{v}^\top \mathbf{v}\mathbf{v}^\top}{v^4} = \frac{\mathbf{v}(\mathbf{v}^\top \mathbf{v})\mathbf{v}^\top}{v^4} = \frac{\mathbf{v}\mathbf{v}^\top}{v^2}$$

was used. Therefore,

$$\underline{\underline{\mathbf{A} = \mathbf{I} + (\gamma - 1) \frac{\mathbf{v}\mathbf{v}^\top}{v^2}}}. \quad (1.34)$$

It is, in fact, true that  $\mathbf{A}(-\mathbf{v}) = \mathbf{A}(\mathbf{v})$ , i.e. the above assumption that  $\tilde{\mathbf{A}} = \mathbf{A}$  is correct.

Thus, the matrix  $\hat{\mathbf{L}}$  of the Lorentz transformation is determined completely as

$$\hat{\mathbf{L}} = \left( \begin{array}{c|c} \gamma & -\frac{\gamma}{c^2} \mathbf{v}^\top \\ \hline -\gamma \mathbf{v} & \mathbf{I} + (\gamma - 1) \frac{\mathbf{v}\mathbf{v}^\top}{v^2} \end{array} \right). \quad (1.35)$$

For  $c \rightarrow \infty$  we get  $\gamma = 1$  and the Lorentz transformation turns into the Galilei transformation with  $\mathbf{T}_{\text{Galilei}}$ .

In the often in the textbooks treated special case when the velocity  $\mathbf{v}$  is towards the  $x_1$ -axis, i.e.

$$\mathbf{v} = \begin{pmatrix} v \\ 0 \\ 0 \end{pmatrix},$$

one obtains

$$\hat{\mathbf{L}} = \left( \begin{array}{c|c} \gamma & -\frac{\gamma}{c^2} (v, 0, 0) \\ \hline -\gamma \begin{pmatrix} v \\ 0 \\ 0 \end{pmatrix} & \mathbf{I} + \frac{(\gamma-1)}{v^2} \begin{pmatrix} v^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{array} \right) = \begin{pmatrix} \gamma & -\frac{\gamma v}{c^2} & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

so

$$\begin{cases} t' = \gamma t - \frac{\gamma}{c^2} v x_1, \\ x'_1 = -\gamma v t + \gamma x_1, \\ x'_2 = x_2, \\ x'_3 = x_3. \end{cases} \quad (1.36)$$



Introducing as the time component in the vector  $\vec{x}$  the time multiplied with the speed of light  $c$ , i.e.  $x_0 = ct$ , we obtain from (1.35) the transformation

$$\vec{x}' \stackrel{\text{def}}{=} \begin{pmatrix} ct' \\ \mathbf{x}' \end{pmatrix} = \left( \begin{array}{c|c} \gamma & -\frac{\gamma}{c} \mathbf{v}^\top \\ \hline -\frac{\gamma}{c} \mathbf{v} & \mathbf{I} + (\gamma - 1) \frac{\mathbf{v} \mathbf{v}^\top}{v^2} \end{array} \right) \begin{pmatrix} ct \\ \mathbf{x} \end{pmatrix}, \quad (1.37)$$

i.e. the new transformation matrix

$$\mathbf{L}(\mathbf{v}) \stackrel{\text{def}}{=} \left( \begin{array}{c|c} \gamma & -\frac{\gamma}{c} \mathbf{v}^\top \\ \hline -\frac{\gamma}{c} \mathbf{v} & \mathbf{I} + (\gamma - 1) \frac{\mathbf{v} \mathbf{v}^\top}{v^2} \end{array} \right) \quad (1.38)$$

is now a *symmetric matrix*, and so

$$\begin{aligned} ct' &= \gamma ct - \frac{\gamma}{c} \mathbf{v}^\top \mathbf{x}, \\ \mathbf{x}' &= \mathbf{x} + (\gamma - 1) \frac{\mathbf{v}^\top \mathbf{x}}{v^2} \mathbf{v} - \gamma \mathbf{v} t. \end{aligned} \quad (1.39)$$

### 1.2.3 Simultaneity at Different Places

It will be seen that events at different places which are simultaneous for an observer in  $\mathcal{X}$  are not in general simultaneous for a moving observer in  $\mathcal{X}'$ . This is caused by the finite speed of light. The reference system  $\mathcal{X}'$  may move towards the stationary reference system  $\mathcal{X}$  with velocity  $\mathbf{v}$ . If the two events  $E_1$  and  $E_2$  in  $\mathcal{X}$  have the coordinates  $\vec{x}_1$  and  $\vec{x}_2$ , then they are *simultaneous* if  $t_1 = t_2$ . Do these two events then happen also at the same time for an observer in the reference system  $\mathcal{X}'$ ? Due to (1.37),

$$ct'_1 = \gamma ct_1 - \frac{\gamma}{c} \mathbf{v}^\top \mathbf{x}_1$$

and

$$ct'_2 = \gamma ct_2 - \frac{\gamma}{c} \mathbf{v}^\top \mathbf{x}_2.$$

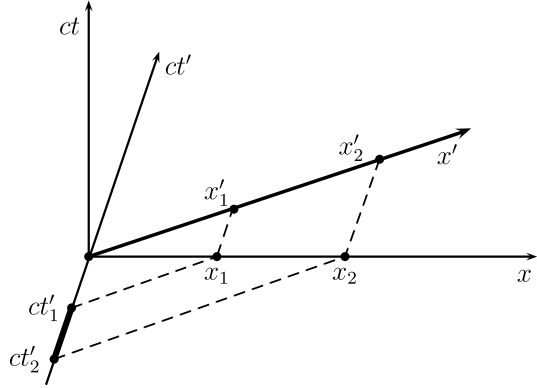
Dividing both equations by  $c$  and subtracting one from another yields

$$t'_1 - t'_2 = \gamma(t_1 - t_2) - \frac{\gamma}{c^2} \mathbf{v}^\top (\mathbf{x}_1 - \mathbf{x}_2),$$

so

$$\underline{\underline{t'_2 - t'_1 = \frac{\gamma}{c^2} \mathbf{v}^\top (\mathbf{x}_1 - \mathbf{x}_2)}}.$$

The events  $\vec{x}'_1$  and  $\vec{x}'_2$  are also simultaneous only if the velocity  $\mathbf{v}$  is perpendicular to the local difference  $\mathbf{x}_1 - \mathbf{x}_2$ . Conclusion:

**Fig. 1.4** Simultaneity

Events at different locations, which in the reference system  $\mathcal{X}$  are simultaneous, need *not* be simultaneous when seen from the reference system  $\mathcal{X}'$ .

*Example* For the special case  $\mathbf{v} = [v, 0, 0]^T$  one only needs to consider the  $ct$ - and the  $x$ -coordinate (Minkowski diagram, Fig. 1.4). The Lorentz transformation does not change the  $x_2$ - and  $x_3$ -component. It is therefore possible to restrict the consideration to the two-dimensional transformations:

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \gamma & -\frac{\gamma}{c}v \\ -\frac{\gamma}{c}v & \gamma \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}.$$

The event  $\begin{pmatrix} 0 \\ x_1 \end{pmatrix}$  is transformed to

$$\begin{pmatrix} ct'_1 \\ x'_1 \end{pmatrix} = \begin{pmatrix} \gamma & -\frac{\gamma}{c}v \\ -\frac{\gamma}{c}v & \gamma \end{pmatrix} \begin{pmatrix} 0 \\ x_1 \end{pmatrix} = \begin{pmatrix} -\frac{\gamma}{c}vx_1 \\ \gamma x_1 \end{pmatrix}$$

and accordingly the event  $\begin{pmatrix} 0 \\ x_2 \end{pmatrix}$ , simultaneous in  $\mathcal{X}$ , is transformed to  $\begin{pmatrix} -\frac{\gamma}{c}vx_2 \\ \gamma x_2 \end{pmatrix}$  in  $\mathcal{X}'$ . In Fig. 1.4, the difference on the  $ct'$ -axis is then

$$ct'_2 - ct'_1 = \frac{\gamma}{c}v(x_1 - x_2) \neq 0,$$

if  $v \neq 0$  and  $x_1 \neq x_2$ , i.e. in the moving reference system  $\mathcal{X}'$  the two events  $E_1$  and  $E_2$  are no longer simultaneous.

### 1.2.4 Length Contraction of Moving Bodies

Einstein (1905) was the first who completely demonstrated that length contraction is an effect due to the change in the notions of space, time and simultaneity brought about by special relativity. Length contraction can simply be derived from

the Lorentz transformation. Let the reference system  $\mathcal{X}$  be stationary and the reference system  $\mathcal{X}'$  moving towards it with the speed  $\mathbf{v}$ . A yardstick has in the stationary system the two endpoints  $\mathbf{x}_1$  and  $\mathbf{x}_2$ . Its resting length is

$$\mathbf{l}_0 = \mathbf{x}_2 - \mathbf{x}_1, \quad (1.40)$$

i.e.

$$l_0^2 = (\mathbf{x}_2 - \mathbf{x}_1)^\top (\mathbf{x}_2 - \mathbf{x}_1). \quad (1.41)$$

At time  $t'$ , the endpoints of the yardstick in the moving reference system  $\mathcal{X}'$  have the coordinates  $\mathbf{x}'_1$  and  $\mathbf{x}'_2$ . With (1.38) we obtain

$$\mathbf{l}_0 = \mathbf{x}_2 - \mathbf{x}_1 = \mathbf{A}(\mathbf{x}'_2 - \mathbf{x}'_1) \stackrel{\text{def}}{=} \mathbf{A}\mathbf{l}'. \quad (1.42)$$

It follows that

$$l_0^2 = \mathbf{l}'^\top \mathbf{l}_0 = \mathbf{l}'^\top \mathbf{A}^2 \mathbf{l}'. \quad (1.43)$$

With (1.32) one gets

$$l_0^2 = \mathbf{l}'^\top \left( \mathbf{I} + (\gamma^2 - 1) \frac{\mathbf{v}\mathbf{v}^\top}{v^2} \right) \mathbf{l}' = \mathbf{l}'^\top \mathbf{l}' + (\gamma^2 - 1) \frac{(\mathbf{v}^\top \mathbf{l}')^2}{v^2}. \quad (1.44)$$

In the product  $\mathbf{v}^\top \mathbf{l}'$ , only the component  $l'_\parallel$  of  $\mathbf{l}'$  parallel to the velocity  $\mathbf{v}$  comes into effect, i.e. it is true that  $\mathbf{v}^\top \mathbf{l}' = v^\top l'_\parallel$ . Thus,

$$l_0^2 = l'^2 + (\gamma^2 - 1) \frac{(\mathbf{v}^\top \mathbf{l}'_\parallel)^2}{v^2} = l'^2 + (\gamma^2 - 1) l'^2_\parallel,$$

so

$$\underline{\underline{l'^2 = l_0^2 - (\gamma^2 - 1) l'^2_\parallel}}. \quad (1.45)$$

Since always  $\gamma^2 - 1 \geq 0$ , it follows from (1.45) that

$$\underline{\underline{l' \leq l_0}}. \quad (1.46)$$

The result of the yardstick length measurement therefore depends on the reference system in which the length measurement was made.

If the yardstick is parallel to the velocity  $\mathbf{v}$ , then  $\mathbf{l}' = l'_\parallel$  and (1.45) becomes  $\gamma l' = l_0$ , or

$$\boxed{l' = \frac{l_0}{\gamma} = l_0 \sqrt{1 - \frac{v^2}{c^2}}}, \quad (1.47)$$

i.e. as  $v \rightarrow c$  one gets  $l' \rightarrow 0$ . For example, if the velocity is  $v = 0.8c$  (i.e. 80 % of the speed of light), then the length is  $l' = 0.6l_0$ .

### 1.2.5 Time Dilation

Time dilation is an actual difference of elapsed time between two events as measured by observers moving relative to each other. Using (1.19), it is

$$\boxed{\Delta t = \gamma \Delta t'}, \quad (1.48)$$

where

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \geq 1.$$

Let  $t'$  be the time of a light clock, measured by an observer in the moving inertial frame  $\mathcal{X}'$  after covering the distance between the mirrors. Then  $t$  is the time required for the light to cover the distance between the two mirrors of the *moving* clock, which is measured by an observer at rest in  $\mathcal{X}$ . The faster the clock moves, i.e. the greater  $v$  is, the longer this time. If, for example, for the moving observer  $t' = 1$  second is elapsed, for the stationary observer  $t = \gamma \geq 1$  seconds have passed. If  $v$  so large that  $\gamma = 20$ , then, for example, for the stationary observer  $t = 20$  years have passed, and for the moving observer only  $t' = 1$  year is gone!

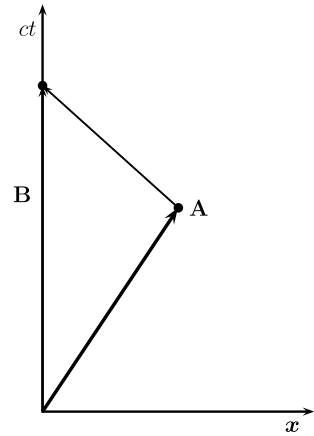
**Twin Paradox** There are two twins **A** and **B** on Earth. **A** starts in a rocket and flies away with high speed, while **B** stays on Earth. During the flight, **A** is aging more slowly than **B**. After some time the rocket is slowed down and returns with **A** at high speed back to Earth. During the flight **A** has aged less than **B**, which remained at rest on Earth, Fig. 1.5. Now comes the paradox: The velocities are relative! One could take also the twin **A** in its entrained coordinate system as stationary and consider **B** as moving with large speeds. That's right; but one key difference is that the twin **A** is not always in the same uniformly moving inertial frame since, at the turning point where the return begins, the inertial system changes! This is not the case for **B**. He always stays in the same inertial frame. Therefore, there is no paradox.

## 1.3 Invariance of the Quadratic Form

The Michelson–Morley experiment says that in any reference system the light propagates in all directions at the same speed  $c$ . If at the origin  $\mathbf{x} = \mathbf{o}$  of  $\mathcal{X}$  a flash of light is ignited, it propagates with the speed of light  $c$  spherically. After the time  $t$ , the light signal reaches all points of the sphere of radius  $ct$ . To points on the sphere applies:

$$x_1^2 + x_2^2 + x_3^2 = (ct)^2, \quad \text{i.e.} \quad (ct)^2 - x_1^2 - x_2^2 - x_3^2 = 0. \quad (1.49)$$

If the origins of the two reference systems  $\mathcal{X}$  and  $\mathcal{X}'$  at the ignition time  $t = t' = 0$  of the light flash are at the same space point  $\mathbf{x}(t = 0) = \mathbf{X}'(t' = 0) = \mathbf{o}$ , the light in

**Fig. 1.5** Twin paradox

the reference system  $\mathcal{X}'$  is also spreading according to the following law:

$$(ct')^2 - x_1'^2 - x_2'^2 - x_3'^2 = 0. \quad (1.50)$$

So this quantity is *invariant*.

One can imagine (1.49) generated by the quadratic form

$$\boxed{\vec{x}^\top \mathbf{M} \vec{x} = 0}, \quad (1.51)$$

with the Minkowski matrix

$$\mathbf{M} \stackrel{\text{def}}{=} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

and the *four-dimensional* vector

$$\vec{x} \stackrel{\text{def}}{=} \begin{pmatrix} ct \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Minkowski was the first who represented Einstein's Special Theory of Relativity using four-dimensional *spacetime* vectors.

### 1.3.1 Invariance with Respect to Lorentz Transformation

Now the invariance of the quadratic form (1.51) to a Lorentz transformation is examined. If  $\vec{x}' = L\vec{x}$ , then

$$\vec{x}'^T M \vec{x}' = \vec{x}^T L^T M L \vec{x}. \quad (1.52)$$

Then for the matrix product  $L^T M L$  with the help of (1.31) and (1.34) one obtains

$$\begin{aligned} L^T M L &= \left( \begin{array}{c|c} \gamma & -\frac{\gamma}{c} \mathbf{v}^T \\ \hline -\frac{\gamma}{c} \mathbf{v} & A \end{array} \right) \left( \begin{array}{c|c} \gamma & -\frac{\gamma}{c} \mathbf{v}^T \\ \hline \frac{\gamma}{c} \mathbf{v} & -A \end{array} \right) \\ &= \left( \begin{array}{c|c} \gamma^2 - \frac{\gamma^2}{c^2} \mathbf{v}^T \mathbf{v} & \frac{\gamma}{c} \mathbf{v}^T A - \frac{\gamma^2}{c} \mathbf{v}^T \\ \hline \frac{\gamma}{c} A \mathbf{v} - \frac{\gamma^2}{c} \mathbf{v} & \frac{\gamma^2}{c^2} \mathbf{v} \mathbf{v}^T - A^2 \end{array} \right) = \begin{pmatrix} 1 & \mathbf{o}^T \\ \mathbf{o} & -I \end{pmatrix} = M. \end{aligned}$$

So for the quadratic forms, in fact, the following is valid:

$$\vec{x}'^T M \vec{x}' = \vec{x}^T M \vec{x},$$

i.e. they are invariant with respect to a Lorentz transformation!

When dealing with the quadratic form, the propagation of light was considered till now. For this the quadratic form  $\vec{x}^T M \vec{x}$  is equal to zero. Considering, however, the movement of a particle, the light will spread faster than the particle, i.e. it will always be true that

$$(ct)^2 > \mathbf{x}^T \mathbf{x}$$

or

$$(ct)^2 - \mathbf{x}^T \mathbf{x} = \vec{x}^T M \vec{x} > 0.$$

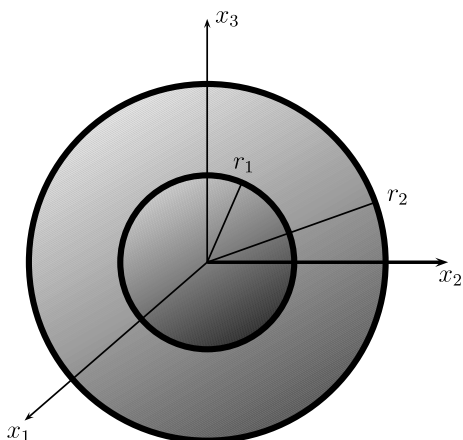
If we denote by  $\Delta \mathbf{x}$  the travelled way between two events and the elapsed time by  $\Delta t$ , we will obtain the four-dimensional spacetime interval  $\Delta s$  from

$$\Delta s^2 = \Delta \vec{x}^T M \Delta \vec{x}. \quad (1.53)$$

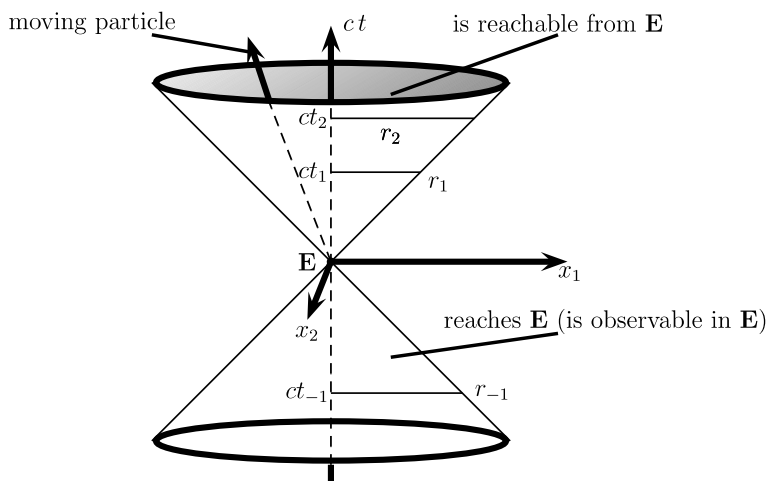
The “distance”  $\Delta s$  between the two events is an *invariant interval* in the four-dimensional spacetime. Since the right-hand side of (1.53) is invariant with respect to a Lorentz transformation,  $\Delta s$  is independent of the chosen inertial frame, it has always the same length. So the theory of relativity does not relativise everything!  $\Delta s^2$  is negative when the distance  $\Delta \vec{x}$  is so big that no light signal can traverse the distance in finite time. This possibility will be examined in the following section.

### 1.3.2 Light Cone

A flash of light at time  $t_0 = 0$  is spreading spherically in the three-dimensional space with the speed of light  $c$ . At the time  $t_1 > t_0$ , the light has reached the surface of a

**Fig. 1.6** Light propagation

sphere of radius  $r_1 = ct_1$ . At time  $t_2 > t_1$ , a spherical surface with a larger radius  $r_2 = ct_2$  is reached, and so forth as shown in Fig. 1.6. In this two-dimensional image, the light wave through space is illustrated by a circle that expands with the speed of light. One can implement this movement of light waves in a space–time diagram in which the time coordinate  $ct$  is vertical and two of the three spatial coordinates, e.g.  $x_1$  and  $x_2$ , are displayed horizontally (Fig. 1.7). Due to the fact that the time coordinate is represented by the time  $t$  multiplied by the speed of light  $c$ , the photons move in straight lines on this diagram, which are sloped at  $45^\circ$ . For photons the possible paths are in an open top cone whose walls have a slope of  $45^\circ$ . The speed of a moving particle is always less than the speed of light, therefore the path must always be run within the light cone with a slope which is always less than  $45^\circ$  to the time axis.

**Fig. 1.7** Light cone

The photons which reach the point **E** at a time before  $t_0 = 0$ , e.g. in the time interval of  $t_{-1} < t_0$ , all come from a ball of radius  $ct_{-1} = r_{-1}$ , etc. Altogether we obtain again a cone as in Fig. 1.7, the cone of the events which can reach the event **E**, i.e. these events are observable from **E**.

In the Minkowski spacetime of special relativity, in any event the light cones are aligned in parallel; the central axes of all light cones are parallel to the time axis. In the general theory of relativity, this is no longer always the case due to the curvature of space, i.e. the central axes of the light cones are not always parallel to the time axis.

### 1.3.3 Proper Time

If one looks at a completely arbitrarily moving clock from any inertial frame, then one can interpret this movement as uniform at any instant of time. If we introduce a coordinate system which is permanently connected to the clock, then this is again an inertial frame. In the infinitesimal time interval  $dt$ , measured with the clock of the observer in the inertial system  $\mathcal{X}$ , the moving clock covers a distance of  $(dx_1^2 + dx_2^2 + dx_3^2)^{1/2}$ . In the inertial system  $\mathcal{X}'$ , permanently connected to the clock, the clock does not move, so  $dx'_1 = dx'_2 = dx'_3 = 0$ , but the elapsed displayed time is  $dt'$ . The invariant quadratic form

$$ds^2 = d\vec{x}^\top M d\vec{x} = d\vec{x}'^\top M d\vec{x}'$$

computed for

$$d\vec{x}^\top = [c dt, dx_1, dx_2, dx_3] \quad \text{and} \quad d\vec{x}'^\top = [c dt', 0, 0, 0]$$

gives

$$ds^2 = c^2 dt^2 - dx_1^2 - dx_2^2 - dx_3^2 = c^2 dt'^2, \quad (1.54)$$

so

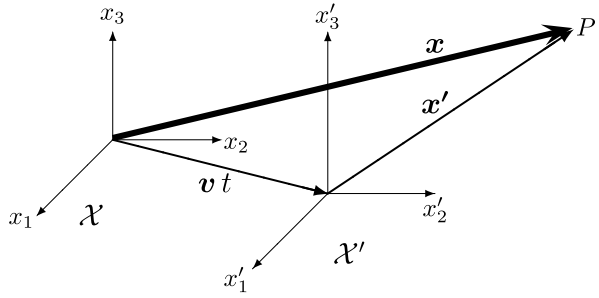
$$\begin{aligned} dt' &= \frac{1}{c} ds = \frac{1}{c} \sqrt{c^2 dt^2 - dx_1^2 - dx_2^2 - dx_3^2} \\ &= dt \sqrt{1 - \frac{dx_1^2 + dx_2^2 + dx_3^2}{c^2 dt^2}}, \end{aligned}$$

and taking

$$v^2 = \frac{dx_1^2 + dx_2^2 + dx_3^2}{dt^2},$$



**Fig. 1.8** Relativistic velocity addition



where  $v$  is the velocity of the moving clock relative to the observer, finally yields the relationship (1.19)

$$dt' = dt \sqrt{1 - \frac{v^2}{c^2}}. \quad (1.55)$$

So, if the stationary clock of the observer shows the time interval  $t_2 - t_1$ , then the moving clock shows the interval  $t'_2 - t'_1$  of the *proper time*,

$$t'_2 - t'_1 = \int_{t_1}^{t_2} \sqrt{1 - \frac{v(t)^2}{c^2}} dt. \quad (1.56)$$

The proper time interval of a moving mass is due to (1.55) and (1.56) always smaller than the time interval in the stationary system. In general, the proper time of a moving mass is called  $\tau$  instead of  $t'$ . Due to (1.54), the proper time is thus

$$\underline{\underline{d\tau = ds/c}}. \quad (1.57)$$

## 1.4 Relativistic Velocity Addition

### 1.4.1 Galilean Addition of Velocities

Galilei observed that if a ship is moving relative to the shore at velocity  $v$ , and a sailor is moving with velocity  $u$  measured on the ship, calculating the velocity of the sailor measured on the shore is what is meant by the addition of the velocities  $v$  and  $u$ . When both the sailor and the ship are moving slowly compared to light, it is accurate enough to use the vector sum

$$\mathbf{w} = \mathbf{v} + \mathbf{u}$$

where  $\mathbf{w}$  is the velocity of the sailor relative to the shore.

Consider the two inertial systems  $\mathcal{X}$  and  $\mathcal{X}'$  in Fig. 1.8. The reference system  $\mathcal{X}'$  is moving with respect to the other reference system  $\mathcal{X}$  with the velocity  $\mathbf{v}$ . Suppose

that the vector  $\mathbf{x}'$  describes the motion of a point  $P$  in the inertial system  $\mathcal{X}'$ , and this point is moving with respect to the reference system  $\mathcal{X}$  with the velocity  $\mathbf{u} = \frac{d\mathbf{x}'}{dt'}$ . What is the speed  $\mathbf{w} \stackrel{\text{def}}{=} \frac{d\mathbf{x}}{dt}$  of the point  $P$  relative to the reference system  $\mathcal{X}$ ?

It is

$$\mathbf{w} = \frac{d\mathbf{x}}{dt} = \frac{d\mathbf{x}}{dt'} \left( \frac{dt}{dt'} \right)^{-1}. \quad (1.58)$$

Between  $\vec{\mathbf{x}}$  and  $\vec{\mathbf{x}}'$  the relationship  $\vec{\mathbf{x}}' = \mathbf{L}(v)\vec{\mathbf{x}}$  holds, or, solving for  $\vec{\mathbf{x}}$ ,

$$\vec{\mathbf{x}} = \mathbf{L}^{-1}(v)\vec{\mathbf{x}}'. \quad (1.59)$$

In detail,

$$t = \gamma_v t' + \frac{\gamma_v}{c^2} \mathbf{v}^\top \mathbf{x}', \quad (1.60)$$

$$\mathbf{x} = \mathbf{x}' + (\gamma_v - 1) \frac{\mathbf{v}^\top \mathbf{x}'}{v^2} \mathbf{v} + \gamma_v \mathbf{v} t'. \quad (1.61)$$

From (1.60) follows

$$\frac{dt}{dt'} = \gamma_v + \frac{\gamma_v}{c^2} \mathbf{v}^\top \frac{d\mathbf{x}'}{dt'} = \gamma_v \left( 1 + \frac{\mathbf{v}^\top \mathbf{u}}{c^2} \right), \quad (1.62)$$

and from (1.61)

$$\begin{aligned} \frac{d\mathbf{x}}{dt'} &= \frac{d\mathbf{x}'}{dt'} + (\gamma_v - 1) \frac{\mathbf{v} \mathbf{v}^\top}{v^2} \frac{d\mathbf{x}'}{dt'} + \gamma_v \mathbf{v} \\ &= \mathbf{u} + (\gamma_v - 1) \frac{\mathbf{v}^\top \mathbf{u}}{v^2} \mathbf{v} + \gamma_v \mathbf{v}. \end{aligned} \quad (1.63)$$

Equations (1.62) and (1.63) used in (1.58) provide

$$\mathbf{w} = \frac{\mathbf{v} + \frac{1}{\gamma_v} \mathbf{u} + (1 - \frac{1}{\gamma_v}) \frac{\mathbf{v}^\top \mathbf{u}}{v^2} \mathbf{v}}{1 + \frac{\mathbf{v}^\top \mathbf{u}}{c^2}}, \quad (1.64)$$

or, after addition of  $\mathbf{u} - \mathbf{u} = \mathbf{o}$  in the numerator,

$$\mathbf{w} = \frac{\mathbf{v} + \mathbf{u} + (\frac{1}{\gamma_v} - 1)(\mathbf{u} - \frac{\mathbf{v}^\top \mathbf{u}}{v^2} \mathbf{v})}{1 + \frac{\mathbf{v}^\top \mathbf{u}}{c^2}}. \quad (1.65)$$

So this is the speed of the point  $P$  with respect to the reference system  $\mathcal{X}$ . For a double vectorial product one has

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a}^\top \mathbf{c}) \mathbf{b} - (\mathbf{a}^\top \mathbf{b}) \mathbf{c}. \quad (1.66)$$

This allows one to transform the last parenthesis in the numerator of (1.65) as follows:

$$\mathbf{u} - \frac{\mathbf{v}^\top \mathbf{u}}{v^2} \mathbf{v} = \frac{\mathbf{v}^\top \mathbf{v}}{v^2} \mathbf{u} - \frac{\mathbf{v}^\top \mathbf{u}}{v^2} \mathbf{v} = \frac{1}{v^2} (\mathbf{v} \times (\mathbf{u} \times \mathbf{v})),$$

i.e. instead of (1.65) one can also write

$$\mathbf{w} = \frac{\mathbf{v} + \mathbf{u} + \frac{1}{v^2} (\frac{1}{\gamma_v} - 1) (\mathbf{v} \times (\mathbf{u} \times \mathbf{v}))}{1 + \frac{\mathbf{v}^\top \mathbf{u}}{c^2}}. \quad (1.67)$$

If the two velocities  $\mathbf{v}$  and  $\mathbf{u}$  are parallel, then  $\mathbf{v} \times \mathbf{u} = \mathbf{o}$  and from (1.67) it is clear that the sum of the two velocities is

$$\mathbf{w} = \frac{\mathbf{v} + \mathbf{u}}{1 + \frac{\mathbf{v}^\top \mathbf{u}}{c^2}}. \quad (1.68)$$

If, in contrast, the two velocities  $\mathbf{v}$  and  $\mathbf{u}$  are perpendicular to each other, then  $\mathbf{v}^\top \mathbf{u} = \mathbf{o}$  and from (1.65) it follows that the sum of the two velocities is

$$\mathbf{w} = \mathbf{v} + \frac{1}{\gamma_v} \mathbf{u}. \quad (1.69)$$

In a further special case which is almost exclusively treated in textbooks and in which both the vector  $\mathbf{u}$  and the vector  $\mathbf{v}$  have only one component in the  $x_1$ -direction, i.e.

$$\mathbf{u} = \begin{pmatrix} u_1 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{v} = \begin{pmatrix} v_1 \\ 0 \\ 0 \end{pmatrix},$$

one has  $w_2 = w_3 = 0$  and

$$w_1 = \frac{v_1 + u_1}{1 + \frac{u_1 v_1}{c^2}}. \quad (1.70)$$

## 1.5 Lorentz Transformation of the Velocity

How does one write the basic laws of mechanics so that they remain invariant under a Lorentz transformation? We start from the transformation equation

$$\vec{\mathbf{x}}' = L \vec{\mathbf{x}} \quad (1.71)$$

with the four-dimensional spacetime vector  $\vec{\mathbf{x}} \stackrel{\text{def}}{=} \begin{pmatrix} ct \\ \mathbf{x} \end{pmatrix}$  and the transformation matrix (1.38)

$$L = \left( \begin{array}{c|c} \gamma & -\frac{\gamma}{c} \mathbf{v} \\ \hline -\frac{\gamma}{c} \mathbf{v}^\top & \mathbf{I} + (\gamma - 1) \frac{\mathbf{v} \mathbf{v}^\top}{v^2} \end{array} \right).$$

Differentiating (1.71) with respect to time  $t'$ , we obtain

$$\frac{d\vec{x}'}{dt'} = \begin{pmatrix} c \\ \frac{d\mathbf{x}'}{dt'} \end{pmatrix} = \mathbf{L} \frac{d\vec{x}}{dt} \frac{dt}{dt'}.$$

Using (1.35) with  $\mathbf{u} \stackrel{\text{def}}{=} \frac{d\mathbf{x}}{dt}$  (which now has a different  $\mathbf{u}$  than in the previous section),

$$\frac{dt'}{dt} = \frac{d}{dt} \left( -\frac{\gamma}{c^2} \mathbf{v}^\top \mathbf{x} + \gamma t \right) = -\frac{\gamma}{c^2} \mathbf{v}^\top \mathbf{u} + \gamma. \quad (1.72)$$

Thus we obtain

$$\begin{pmatrix} c \\ \mathbf{u}' \end{pmatrix} = \mathbf{L} \begin{pmatrix} c \\ \mathbf{u} \end{pmatrix} \frac{1}{\gamma} \frac{1}{1 - \frac{\mathbf{v}^\top \mathbf{u}}{c^2}}. \quad (1.73)$$

In (1.73), we see that the velocity vector  $\begin{pmatrix} c \\ \mathbf{u}' \end{pmatrix}$  is *not* transformed into the velocity vector  $\begin{pmatrix} c \\ \mathbf{u} \end{pmatrix}$  using a Lorentz transformation matrix! For this to be the case, the definition of the velocity has to be modified. For this purpose, a short interim statement:

The second block row of (1.73) provides

$$\mathbf{u}' = \frac{1}{\gamma(1 - \frac{\mathbf{v}^\top \mathbf{u}}{c^2})} (\mathbf{A}\mathbf{u} - \gamma \mathbf{v}).$$

Taking the scalar product of this vector with itself, we obtain

$$u'^2 \stackrel{\text{def}}{=} \mathbf{u}'^\top \mathbf{u}' = \left( \frac{1}{\gamma(1 - \frac{\mathbf{v}^\top \mathbf{u}}{c^2})} \right)^2 (\mathbf{u}^\top \mathbf{A}^\top - \gamma \mathbf{v}^\top) (\mathbf{A}\mathbf{u} - \gamma \mathbf{v}).$$

Since  $\mathbf{A} = \mathbf{A}^\top$  and  $\mathbf{A}^2 = \mathbf{I} + \frac{\gamma^2}{c^2} \mathbf{v} \mathbf{v}^\top$ , the above is

$$\begin{aligned} u'^2 &= \frac{1}{\gamma^2(1 - \frac{\mathbf{v}^\top \mathbf{u}}{c^2})^2} \left( u^2 - \frac{\gamma^2}{c^2} (\mathbf{u}^\top \mathbf{v})^2 - 2\gamma^2 (\mathbf{u}^\top \mathbf{v})^2 + \gamma^2 v^2 \right) \\ &= \frac{1}{(1 - \frac{\mathbf{v}^\top \mathbf{u}}{c^2})^2} \left( \frac{1}{\gamma^2} u^2 - \frac{1}{c^2} (\mathbf{u}^\top \mathbf{v})^2 - 2(\mathbf{u}^\top \mathbf{v})^2 + v^2 \right). \end{aligned} \quad (1.74)$$

With

$$\frac{u^2}{\gamma^2} = \left( 1 - \frac{v^2}{c^2} \right) u^2 = u^2 - \frac{v^2 u^2}{c^2}$$

from (1.74) (without  $\gamma$ ) one gets

$$u'^2 = \frac{1}{(1 - \frac{\mathbf{v}^\top \mathbf{u}}{c^2})^2} \left( u^2 + \frac{-(\mathbf{u}^\top \mathbf{v})^2 - v^2 u^2}{c^2} - 2(\mathbf{u}^\top \mathbf{v})^2 + v^2 \right). \quad (1.75)$$

With the help of (1.75) one obtains

$$\begin{aligned} 1 - \frac{u'^2}{c^2} &= \frac{1}{\left(1 - \frac{\mathbf{v}^\top \mathbf{u}}{c^2}\right)^2} \left(1 - \frac{u^2}{c^2} - \frac{v^2}{c^2} + \frac{v^2 u^2}{c^4}\right) \\ &= \frac{1}{\left(1 - \frac{\mathbf{v}^\top \mathbf{u}}{c^2}\right)^2} \left(1 - \frac{u^2}{c^2}\right) \left(1 - \frac{v^2}{c^2}\right), \end{aligned}$$

and this implies

$$\sqrt{1 - \frac{u'^2}{c^2}} = \sqrt{1 - \frac{u^2}{c^2}} \sqrt{1 - \frac{v^2}{c^2}} / \left(1 - \frac{\mathbf{v}^\top \mathbf{u}}{c^2}\right),$$

or with

$$\gamma_u \stackrel{\text{def}}{=} \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}}$$

and

$$\gamma_{u'} \stackrel{\text{def}}{=} \frac{1}{\sqrt{1 - \frac{u'^2}{c^2}}},$$

we finally get

$$\underline{\underline{\gamma \left(1 - \frac{\mathbf{v}^\top \mathbf{u}}{c^2}\right) = \frac{\gamma_{u'}}{\gamma_u}}}. \quad (1.76)$$

Inserting (1.76) into (1.73), we obtain

$$\begin{pmatrix} c \\ \mathbf{u}' \end{pmatrix} = \mathbf{L} \begin{pmatrix} c \\ \mathbf{u} \end{pmatrix} \frac{\gamma_u}{\gamma_{u'}}, \quad (1.77)$$

or

$$\boxed{\begin{pmatrix} \gamma_{u'} c \\ \gamma_{u'} \mathbf{u}' \end{pmatrix} = \mathbf{L} \begin{pmatrix} \gamma_u c \\ \gamma_u \mathbf{u} \end{pmatrix}}. \quad (1.78)$$

The thus modified new velocity vector

$$\bar{\mathbf{u}} \stackrel{\text{def}}{=} \gamma_u \begin{pmatrix} c \\ \mathbf{u} \end{pmatrix} \quad (1.79)$$

now is transformed by a Lorentz transformation  $\mathbf{L}$  into the velocity vector  $\bar{\mathbf{u}}'$ :

$$\boxed{\bar{\mathbf{u}}' = \mathbf{L} \bar{\mathbf{u}}}. \quad (1.80)$$

The so-defined velocity  $\bar{\mathbf{u}}$  is much better suited for the formulation of physics laws, as they have the same shape in every inertial frame. With  $\mathbf{u}$  this would not be the

case. For clarification we re-emphasize that a particle or a focal point moves with the velocity  $\mathbf{u}$  in an inertial system  $\mathcal{X}'$ , which itself moves or can move at the speed  $\mathbf{v}$  with respect to another reference system. That is the difference between  $\mathbf{u}$  and  $\mathbf{v}$ !

By the way, the quadratic form for the velocity

$$\bar{\mathbf{u}}^\top \mathbf{M} \bar{\mathbf{u}} = \gamma_u^2 c^2 - \gamma_u^2 \mathbf{u}^\top \mathbf{u} = \frac{c^4}{c^2 - u^2} - \frac{c^2 u^2}{c^2 - u^2} = c^2 \quad (1.81)$$

is, of course, invariant with respect to a Lorentz transformation because—as one can easily show—even  $\bar{\mathbf{u}}'^\top \mathbf{M} \bar{\mathbf{u}}' = c^2$  holds.

In (1.57), the proper time  $d\tau = \frac{1}{c} ds$  was introduced. We have

$$ds^2 = \bar{\mathbf{x}}^\top \mathbf{M} \bar{\mathbf{x}} = c^2 dt^2 - \mathbf{x}^\top \mathbf{x} = c^2 dt^2 \left( 1 - \frac{1}{c^2} \frac{d\mathbf{x}^\top}{dt} \cdot \frac{d\mathbf{x}}{dt} \right),$$

so with  $\frac{d\mathbf{x}}{dt} = \mathbf{u}$

$$d\tau = dt \left( 1 - \frac{1}{c^2} \mathbf{u}^\top \mathbf{u} \right)^{\frac{1}{2}} = dt \left( 1 - \frac{u^2}{c^2} \right)^{\frac{1}{2}},$$

or with  $\gamma_u = \left( 1 - \frac{u^2}{c^2} \right)^{-\frac{1}{2}}$

$$\underline{\underline{dt}} = \gamma_u d\tau. \quad (1.82)$$

This is the same relationship as in the time dilation in (1.73).  $\tau$  is therefore the time that a comoving clock displays, while  $t$  is the time that an observer at rest measures. However, the moving clock must no longer move rectilinearly and uniformly!

In (1.79), one has  $\mathbf{u} = \frac{d\mathbf{x}}{dt}$ . Replacing  $dt$  by (1.82) in it gives  $\mathbf{u} = \frac{1}{\gamma_u} \frac{d\mathbf{x}}{d\tau}$ , or equivalently,  $\gamma_u \mathbf{u} = \frac{d\mathbf{x}}{d\tau}$ . Furthermore,

$$d\bar{\mathbf{x}} = \begin{pmatrix} c dt \\ d\mathbf{x} \end{pmatrix} = \begin{pmatrix} c \gamma_u d\tau \\ d\mathbf{x} \end{pmatrix},$$

so

$$\underline{\underline{\frac{d\bar{\mathbf{x}}}{d\tau}}} = \begin{pmatrix} \gamma_u c \\ \gamma_u \frac{d\mathbf{x}}{d\tau} \end{pmatrix} = \begin{pmatrix} \gamma_u c \\ \gamma_u \mathbf{u} \end{pmatrix} = \underline{\underline{\bar{\mathbf{u}}}}. \quad (1.83)$$

When the trajectory in spacetime is parameterized by the proper time  $\tau$ ,  $\bar{\mathbf{x}} = \bar{\mathbf{x}}(\tau)$ , then  $\bar{\mathbf{u}} = \frac{d\bar{\mathbf{x}}}{d\tau}$  is the four-velocity along the trajectory.

## 1.6 Momentum and Its Lorentz Transformation

Multiplying the equation  $\bar{\mathbf{u}}' = \mathbf{L} \bar{\mathbf{u}}$  with the rest mass  $m_0$ , we obtain

$$\begin{pmatrix} m_0 \gamma_{u'} c \\ m_0 \gamma_{u'} \mathbf{u}' \end{pmatrix} = \mathbf{L} \begin{pmatrix} m_0 \gamma_u c \\ m_0 \gamma_u \mathbf{u} \end{pmatrix}. \quad (1.84)$$

Herein we define as usual the *momentum*

$$\mathbf{p} \stackrel{\text{def}}{=} m_0 \gamma_u \mathbf{u} = m_u \mathbf{u} = m_u \frac{d\mathbf{x}}{dt} \quad (1.85)$$

by

$$m_u \stackrel{\text{def}}{=} m_0 \gamma_u = \frac{m_0}{\sqrt{1 - \frac{u^2}{c^2}}}.$$

The momentum vector

$$\vec{\mathbf{p}} \stackrel{\text{def}}{=} \begin{pmatrix} m_u c \\ \mathbf{p} \end{pmatrix} = m_0 \vec{\mathbf{u}} = m_0 \gamma_u \begin{pmatrix} c \\ \mathbf{u} \end{pmatrix} \quad (1.86)$$

is due to (1.84) transformed as

$$\boxed{\vec{\mathbf{p}}' = \mathbf{L} \vec{\mathbf{p}}.} \quad (1.87)$$

Also, the quadratic form associated with the momentum vector

$$\vec{\mathbf{p}}^\top \mathbf{M} \vec{\mathbf{p}} = m_0^2 \vec{\mathbf{u}}^\top \mathbf{M} \vec{\mathbf{u}} = m_0^2 c^2 \quad (1.88)$$

is invariant with respect to a Lorentz transformation because also  $\vec{\mathbf{p}}'^\top \mathbf{M} \vec{\mathbf{p}}' = m_0^2 c^2$ .

## 1.7 Acceleration and Force

### 1.7.1 Acceleration

The acceleration is generally defined as the time derivative of speed. Differentiating the modified velocity vector  $\vec{\mathbf{u}} = \begin{pmatrix} \gamma_u c \\ \gamma_u \mathbf{u} \end{pmatrix} \in \mathbb{R}^4$  with respect to time  $t$ , one receives as the derivative of the second component in this vector

$$\frac{d}{dt}(\gamma_u \mathbf{u}) = \frac{d\gamma_u}{dt} \mathbf{u} + \gamma_u \frac{d\mathbf{u}}{dt}. \quad (1.89)$$

In particular, for  $\frac{d\gamma_u}{dt}$  we obtain with

$$\mathbf{a} \stackrel{\text{def}}{=} \frac{d\mathbf{u}}{dt} \in \mathbb{R}^3$$

and

$$\frac{d\mathbf{u}^2}{dt} = \frac{d\mathbf{u}^\top \mathbf{u}}{dt} = \frac{d\mathbf{u}^\top}{dt} \mathbf{u} + \mathbf{u}^\top \frac{d\mathbf{u}}{dt} = 2\mathbf{u}^\top \mathbf{a}$$

the result

$$\begin{aligned}\frac{d\gamma_u}{dt} &= \frac{d}{dt} \left(1 - \frac{u^2}{c^2}\right)^{-1/2} = -\frac{1}{2} \left(1 - \frac{u^2}{c^2}\right)^{-3/2} \cdot \frac{d}{dt} \left(1 - \frac{u^2}{c^2}\right) \\ &= -\frac{1}{2} \gamma_u^3 \cdot \frac{-2\mathbf{u}^\top}{c^2} \cdot \mathbf{a},\end{aligned}\tag{1.90}$$

i.e.

$$\frac{d\gamma_u}{dt} = \frac{\gamma_u^3}{c^2} \cdot \mathbf{u}^\top \mathbf{a}.\tag{1.91}$$

With (1.91) one obtains for (1.89)

$$\frac{d}{dt}(\gamma_u \mathbf{u}) = \frac{\gamma_u^3}{c^2} \cdot \mathbf{u}^\top \mathbf{a} \cdot \mathbf{u} + \gamma_u \cdot \mathbf{a}.\tag{1.92}$$

Differentiating the velocity transformation equation (1.78),

$$\begin{pmatrix} \gamma_{u'} c \\ \gamma_{u'} \mathbf{u}' \end{pmatrix} = \mathbf{L} \begin{pmatrix} \gamma_u c \\ \gamma_u \mathbf{u} \end{pmatrix},$$

with respect to time  $t'$ , we obtain

$$\frac{d}{dt'} \begin{pmatrix} \gamma_{u'} c \\ \gamma_{u'} \mathbf{u}' \end{pmatrix} = \mathbf{L} \cdot \frac{d}{dt} \begin{pmatrix} \gamma_u c \\ \gamma_u \mathbf{u} \end{pmatrix} \cdot \frac{dt}{dt'}.\tag{1.93}$$

From (1.72) and (1.76) follows

$$\frac{dt}{dt'} = \frac{\gamma_u}{\gamma_{u'}}.\tag{1.94}$$

This, used in (1.93), results for the newly defined four-dimensional acceleration vector

$$\boxed{\vec{\mathbf{a}} \stackrel{\text{def}}{=} \gamma_u \cdot \frac{d}{dt} \vec{\mathbf{u}} \in \mathbb{R}^4}\tag{1.95}$$

in the Lorentz transformation of the acceleration vector  $\vec{\mathbf{a}}$

$$\boxed{\vec{\mathbf{a}}' = \mathbf{L} \vec{\mathbf{a}}.}\tag{1.96}$$

The acceleration vector defined in (1.95),

$$\vec{\mathbf{a}} = \gamma_u \cdot \frac{d}{dt} \begin{pmatrix} \gamma_u c \\ \gamma_u \mathbf{u} \end{pmatrix},\tag{1.97}$$



is therefore suitable to formulate physical laws in relativistic form! With  $dt = \gamma_u d\tau$  and  $\vec{u} = d\vec{x}/d\tau$  one can write for (1.97)

$$\boxed{\vec{a} = \frac{d\vec{u}}{d\tau} = \frac{d^2\vec{x}}{d\tau^2}.} \quad (1.98)$$

The vector  $\vec{a}$  is also obtained using (1.91) and (1.92) as

$$\boxed{\vec{a} = \begin{pmatrix} \frac{\gamma_u^4}{c} \cdot \mathbf{u}^\top \mathbf{a} \\ \frac{\gamma_u^4}{c^2} \cdot \mathbf{u}^\top \mathbf{a} \cdot \mathbf{u} + \gamma_u^2 \cdot \mathbf{a} \end{pmatrix}.} \quad (1.99)$$

If  $\vec{u}^\top \mathbf{M} \vec{u} = c^2$  is differentiated with respect to the proper time  $\tau$ , we obtain

$$\frac{d\vec{u}^\top}{d\tau} \mathbf{M} \vec{u} + \vec{u}^\top \mathbf{M} \frac{d\vec{u}}{d\tau} = 2\vec{u}^\top \mathbf{M} \frac{d\vec{u}}{d\tau} = 2\vec{u}^\top \mathbf{M} \vec{a} = 0,$$

i.e. the two four-dimensional vectors  $\vec{u}$  and  $\mathbf{M} \vec{a}$  are orthogonal in  $\mathbb{R}^4$ !

For each point of time of an arbitrarily accelerated motion, one can always specify a reference system  $\mathcal{X}'$  which is an inertial system, named “local inertial system”. We obtain the corresponding Lorentz transformation by selecting  $\mathbf{L}(\mathbf{u})$  as the transformation matrix. Then, with

$$\mathbf{A}(\mathbf{u}) \stackrel{\text{def}}{=} \mathbf{I} + (\gamma_u - 1) \frac{\mathbf{u} \mathbf{u}^\top}{u^2},$$

it is indeed true that

$$\begin{aligned} \vec{a}' &= \mathbf{L}(\mathbf{u}) \vec{a} = \begin{pmatrix} \gamma_u & -\frac{\gamma_u}{c} \mathbf{u}^\top \\ -\frac{\gamma_u}{c} \mathbf{u} & \mathbf{A}(\mathbf{u}) \end{pmatrix} \begin{pmatrix} \frac{\gamma_u^4}{c} \cdot \mathbf{u}^\top \mathbf{a} \\ \gamma_u^2 \cdot \mathbf{a} + \frac{\gamma_u^4}{c^2} \cdot \mathbf{u}^\top \mathbf{a} \cdot \mathbf{u} \end{pmatrix} \\ &= \begin{pmatrix} -\gamma_u^3 \frac{\mathbf{u}^\top \mathbf{a}}{c} - \gamma_u^5 \frac{(\mathbf{u}^\top \mathbf{a})(\mathbf{u}^\top \mathbf{u})}{c^3} + \gamma_u^5 \frac{\mathbf{u}^\top \mathbf{a}}{c} \\ \gamma_u^2 \mathbf{A}(\mathbf{u}) \mathbf{a} + \gamma_u^4 \frac{\mathbf{u}(\mathbf{u}^\top \mathbf{a})}{c^2} + (\gamma_u - 1) \gamma_u^4 \frac{\mathbf{u}(\mathbf{u}^\top \mathbf{u})(\mathbf{u}^\top \mathbf{a})}{c^2 u^2} - \gamma_u^5 \frac{(\mathbf{u}^\top \mathbf{a}) \mathbf{u}}{c^2} \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ \gamma_u^2 \mathbf{A}(\mathbf{u}) \mathbf{a} \end{pmatrix} = \begin{pmatrix} 0 \\ \gamma_u^2 \mathbf{a} + \gamma_u^2 (\gamma_u - 1) \frac{\mathbf{u}^\top \mathbf{a}}{u^2} \mathbf{u} \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{a}' \end{pmatrix}. \end{aligned}$$

## 1.7.2 Equation of Motion and Force

The relativistic equation of motion for a particle has to be Lorentz-invariant, and, in the inertial system of the considered particle, Newton’s equation of motion has to be true:

$$m_0 \frac{d\mathbf{u}}{dt} = \mathbf{f} \in \mathbb{R}^3. \quad (1.100)$$

Let the accompanying inertial system be  $\mathcal{X}$ . Furthermore, suppose  $\mathcal{X}'$  is the inertial system which moves relative to  $\mathcal{X}$  with the constant speed  $\mathbf{u}(t_0)$ . The particle rests momentarily at time  $t = t_0$  in  $\mathcal{X}'$ . The equation of motion (1.100) refers to a point of time and its neighbourhood. For this neighbourhood,  $t = t_0 \pm dt$  is the desired arbitrary small speed in  $\mathcal{X}'$ . For speeds  $v \ll c$  we have (1.100). Hence in  $\mathcal{X}'$

$$m_0 \frac{d\mathbf{u}'}{dt'} = \mathbf{f}' \in \mathbb{R}^3 \quad (1.101)$$

also holds exactly. From (1.101), the relativistic equations of motion in an arbitrary reference system may be derived. In (1.101),  $m_0$  is the rest mass and  $\mathbf{f}'$  the three-dimensional force in  $\mathcal{X}'$ . Expand the vector  $\mathbf{f}'$  in (1.101) to a four-vector and call the result  $\vec{\mathbf{f}}'$ :

$$m_0 \frac{d}{dt'} \begin{pmatrix} c \\ \mathbf{u}' \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{f}' \end{pmatrix} \stackrel{\text{def}}{=} \vec{\mathbf{f}}'. \quad (1.102)$$

Thus  $\vec{\mathbf{f}}'$  is specified in the resting system  $\mathcal{X}'$ . In the inertial system  $\mathcal{X}$ , in which the mass particle moves with the velocity  $\mathbf{u}$ ,  $\vec{\mathbf{f}}$  is obtained by a Lorentz transformation with  $\mathbf{L}(-\mathbf{u})$ :

$$\vec{\mathbf{f}} = \mathbf{L}(-\mathbf{u}) \begin{pmatrix} 0 \\ \mathbf{f}' \end{pmatrix} = \begin{pmatrix} \frac{\gamma_u}{c} \mathbf{u}^\top \mathbf{f}' \\ \mathbf{A}(\mathbf{u}) \mathbf{f}' \end{pmatrix} \stackrel{\text{def}}{=} \begin{pmatrix} f_0 \\ \mathbf{f} \end{pmatrix}. \quad (1.103)$$

The equation

$$m_0 \gamma \frac{d}{dt} \begin{pmatrix} \gamma c \\ \gamma \mathbf{u} \end{pmatrix} = \vec{\mathbf{f}} = \begin{pmatrix} f_0 \\ \mathbf{f} \end{pmatrix},$$

i.e.

$$\boxed{m_0 \vec{\mathbf{a}} = \vec{\mathbf{f}}} \quad (1.104)$$

possesses all the desired properties! The four-vectors  $\vec{\mathbf{a}}$  and  $\vec{\mathbf{f}}$  are Lorentz-invariant and, in the inertial frame of the particle, this equation is reduced to Newton's equation of motion

$$m_0 \begin{pmatrix} 0 \\ \frac{d\mathbf{u}'}{dt'} \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{f}' \end{pmatrix}.$$

For the last three components of the equation of motion (1.104),

$$\frac{d(m_u \mathbf{u})}{dt} = \frac{1}{\gamma_u} \mathbf{f} \quad (1.105)$$

with the velocity-dependent mass

$$m_u \stackrel{\text{def}}{=} \gamma_u m_0. \quad (1.106)$$

In the theory of relativity, the time derivative of the momentum  $m_u \mathbf{u}$  is also interpreted as force. The components  $f_i$  of the relativistic equation of motion are thus,

according to (1.103) and (1.104),

$$\underline{\underline{f_0 = \gamma_u \frac{d}{dt}(m_u c) = \frac{\gamma_u}{c} \mathbf{u}^\top \mathbf{f}'}} \quad (1.107)$$

and

$$\underline{\underline{\mathbf{f} = \gamma_u \frac{d}{dt}(m_u \mathbf{u}) = A(\mathbf{u}) \mathbf{f}'}} \quad (1.108)$$

### 1.7.3 Energy and Rest Mass

Equation (1.107) multiplied with  $c/\gamma_u$  provides

$$\frac{d}{dt}(m_u c^2) = \mathbf{u}^\top \mathbf{f}. \quad (1.109)$$

In (1.109),  $\mathbf{u}^\top \mathbf{f}$  is the instantaneous power. This is the work per unit of time, done by the force  $\mathbf{f}$ . So the left-hand side of (1.109) must be the temporal change of energy, i.e.  $m_u c^2 = \gamma_u m_0 c^2$  is energy. We obtain for the *relativistic energy* the most renowned formula of the theory of relativity:

$$\boxed{E = m_u c^2.} \quad (1.110)$$

When  $\mathbf{u} = \mathbf{o}$ , i.e. when the particle is at rest,  $\gamma_u = 1$ , and

$$\boxed{E_0 = m_0 c^2} \quad (1.111)$$

is the “rest energy”, Einstein’s famous formula. Weinberg says in [36] about the content of this formula: “If some mass is destroyed (as in radioactive decay or fusion, or fission), then very large quantities of kinetic energy will be liberated, with consequences of well-known importance.”

The four-dimensional momentum vector  $\vec{\mathbf{p}}$  is then recognised as the combination of energy and momentum:

$$\vec{\mathbf{p}} = \begin{pmatrix} E/c \\ \mathbf{p} \end{pmatrix}. \quad (1.112)$$

For the quadratic form in (1.88)

$$\vec{\mathbf{p}}^\top \mathbf{M} \vec{\mathbf{p}} = m_0^2 c^2$$

we now obtain

$$\vec{\mathbf{p}}^\top \mathbf{M} \vec{\mathbf{p}} = (E/c, \mathbf{p}^\top) \begin{pmatrix} E/c \\ -\mathbf{p} \end{pmatrix} = E^2/c^2 - p^2 = m_0^2 c^2,$$

i.e.

$$E = \sqrt{(m_0c^2)^2 + p^2c^2}. \quad (1.113)$$

For high velocities the momentum term  $p^2c^2$  dominates here:  $E = pc$  as for neutrinos and particles in accelerators (CERN). For small velocities  $u \ll c$ , we may use the approximation

$$E \approx m_0c^2 + \frac{1}{2}m_0u^2.$$

The first term is the rest energy  $E_0$ . The second term is the classical kinetic energy  $E_{\text{kin}}$ . The relativistic kinetic energy is

$$E_{\text{kin}} = E - m_0c^2 = (\gamma_u - 1)m_0c^2. \quad (1.114)$$

If  $v \rightarrow c$ , then  $E_{\text{kin}} \rightarrow \infty$ . Equation (1.114) gives a reason why mass particles cannot be accelerated up to the velocity of light!

### 1.7.4 Emission of Energy

A body, at rest in  $\mathcal{X}'$  with the rest mass  $m_{0,\text{before}}$ , radiates at a certain time the energy  $E'_{\text{Emission}}$  in form of light or heat radiation. This radiation is emitted symmetrically so that the total momentum of the radiated energy in  $\mathcal{X}'$  is zero, i.e. the body remains at rest during the radiation process. The energy–momentum vector of the radiation in  $\mathcal{X}'$  is therefore

$$\begin{pmatrix} \frac{1}{c}E'_{\text{Emission}} \\ 0 \\ 0 \\ 0 \end{pmatrix}. \quad (1.115)$$

Compared with the inertial system  $\mathcal{X}$ , the body is moving with the velocity  $\mathbf{v}$ . In this inertial system, the total momentum is equal to the momentum of the body  $\mathbf{p}_{\text{before}} = \gamma_u m_{0,\text{before}} \mathbf{u}$ . After radiation, the body has the momentum  $\mathbf{p}_{\text{after}} = \gamma_u m_{0,\text{after}} \mathbf{u}$ , and the momentum of radiation is calculated using (1.115) with the Lorentz transformation  $\mathbf{L}(-\mathbf{u})$  from  $\mathcal{X}'$  to  $\mathcal{X}$ :

$$\vec{\mathbf{p}} = \mathbf{L}(-\mathbf{u}) \begin{pmatrix} \frac{1}{c}E'_{\text{Emission}} \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma_u \frac{1}{c}E'_{\text{Emission}} \\ \gamma_u \mathbf{u} \frac{1}{c^2}E'_{\text{Emission}} \end{pmatrix}.$$

During the emission of radiation, the body has given away the momentum  $\gamma_u \mathbf{u} \frac{1}{c^2}E'_{\text{Emission}}$ , without changing its speed. This is only possible because the body

has changed its rest mass! Because of the momentum conservation law, one has

$$\gamma_u m_{0,\text{before}} \mathbf{u} = \gamma_u m_{0,\text{after}} \mathbf{u} + \gamma_u \mathbf{u} \frac{1}{c^2} E'_{\text{Emission}}.$$

It follows that

$$m_{0,\text{after}} = m_{0,\text{before}} - \frac{1}{c^2} E'_{\text{Emission}}. \quad (1.116)$$

In Einstein's words: "If a body emits the energy  $E'$  in the form of radiation, its mass is reduced by  $E'/c^2$ ."

## 1.8 Relativistic Electrodynamics

### 1.8.1 Maxwell's Equations

The magnetic field associated with the induction  $\mathbf{b}$  and the electric field with the field strength<sup>2</sup>  $\mathbf{e}$ , the sources of the charge  $q$  and the current  $\mathbf{j}$  satisfy the Maxwell's equations:<sup>3</sup>

$$\nabla \times \mathbf{b} = \frac{1}{c} \left( \frac{\partial \mathbf{e}}{\partial t} + \mathbf{j} \right), \quad (1.117)$$

$$\nabla^\top \mathbf{e} = \rho, \quad (1.118)$$

$$\nabla \times \mathbf{e} = -\frac{1}{c} \frac{\partial \mathbf{b}}{\partial t}, \quad (1.119)$$

$$\nabla^\top \mathbf{b} = 0. \quad (1.120)$$

The *Derivative Operator*  $\nabla$  is the 3-dimensional column vector

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix}$$

and  $\nabla^\top$  is the 3-dimensional row-vector

$$\nabla^\top = \left( \frac{\partial}{\partial x} \mid \frac{\partial}{\partial y} \mid \frac{\partial}{\partial z} \right),$$

<sup>2</sup>Since the magnetic induction and the electric field strength are *vectors*, they are marked by bold small letters  $\mathbf{b}$  and  $\mathbf{e}$ .

<sup>3</sup>In other symbols,  $\nabla \times \mathbf{b} = \text{curl } \mathbf{b}$  and  $\nabla^\top \mathbf{e} = \text{div } \mathbf{e}$ .

so that

$$\nabla^\top \mathbf{e} = \frac{\partial e_x}{\partial x} + \frac{\partial e_y}{\partial y} + \frac{\partial e_z}{\partial z}, \quad (1.121)$$

and

$$\nabla \times \mathbf{b} = -\mathbf{B}_\times \nabla \stackrel{\text{def}}{=} \begin{pmatrix} \frac{\partial b_z}{\partial y} - \frac{\partial b_y}{\partial z} \\ -\frac{\partial b_z}{\partial x} + \frac{\partial b_x}{\partial z} \\ \frac{\partial b_y}{\partial x} - \frac{\partial b_x}{\partial y} \end{pmatrix}, \quad (1.122)$$

with<sup>4</sup>

$$\mathbf{B}_\times \stackrel{\text{def}}{=} \begin{pmatrix} 0 & -b_z & b_y \\ b_z & 0 & -b_x \\ -b_y & b_x & 0 \end{pmatrix}. \quad (1.123)$$

Equations (1.117) and (1.118) can be summed up after minor changes in the following four equations:

$$\begin{aligned} \frac{\partial e_x}{\partial x} + \frac{\partial e_y}{\partial y} + \frac{\partial e_z}{\partial z} &= \rho, \\ -\frac{1}{c} \frac{\partial e_x}{\partial t} - \frac{\partial b_z}{\partial y} + \frac{\partial b_y}{\partial z} &= \frac{1}{c} j_1, \\ -\frac{1}{c} \frac{\partial e_y}{\partial t} + \frac{\partial b_z}{\partial x} - \frac{\partial b_x}{\partial z} &= \frac{1}{c} j_2, \\ -\frac{1}{c} \frac{\partial e_z}{\partial t} - \frac{\partial b_y}{\partial x} + \frac{\partial b_x}{\partial y} &= \frac{1}{c} j_3, \end{aligned} \quad (1.124)$$

or in matrix form,<sup>5</sup>

$$\underbrace{\begin{pmatrix} 0 & \mathbf{e}^\top \\ -\mathbf{e} & \mathbf{B}_\times \end{pmatrix}}_{\stackrel{\text{def}}{=} \mathbf{F}_{B,e}} \gamma \underbrace{\begin{pmatrix} -\frac{1}{c} \frac{\partial}{\partial t} \\ \nabla \end{pmatrix}}_{\stackrel{\text{def}}{=} \vec{\nabla}} = \frac{1}{c} \gamma \underbrace{\begin{pmatrix} c\rho \\ -\mathbf{j} \end{pmatrix}}_{\stackrel{\text{def}}{=} \vec{j}} \in \mathbb{R}^4. \quad (1.125)$$

The skew-symmetric matrix  $\mathbf{F}_{B,e}$ , composed of  $\mathbf{B}_\times$  and  $\mathbf{e}$ , is called Faraday's matrix. A similarly structured field strength matrix is obtained if (1.119) and (1.120)

<sup>4</sup>An equation  $\mathbf{c} = \mathbf{a} \times \mathbf{b}$  cannot in this form be linearly transformed with an invertible transformation matrix  $T$ . But with  $\mathbf{A}_\times \stackrel{\text{def}}{=} \begin{pmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{pmatrix}$  this is possible in the form of the equation  $\mathbf{c} = \mathbf{A}_\times \mathbf{b}$ :  
 $(T\mathbf{c}) = (T\mathbf{A}_\times T^{-1})(T\mathbf{b})$ .

<sup>5</sup>The factor  $\gamma$  has been added on both sides, so that no difficulties arise later in the invariance of this equation with respect to a Lorentz transformation.

are gathered and summarized in the following equations:

$$\begin{aligned}
 & -\frac{\partial b_x}{\partial x} - \frac{\partial b_y}{\partial y} - \frac{\partial b_z}{\partial z} = 0, \\
 & +\frac{1}{c} \frac{\partial b_x}{\partial t} - \frac{\partial e_z}{\partial y} + \frac{\partial e_y}{\partial z} = 0, \\
 & +\frac{1}{c} \frac{\partial b_y}{\partial t} + \frac{\partial e_z}{\partial x} - \frac{\partial e_x}{\partial z} = 0, \\
 & +\frac{1}{c} \frac{\partial b_z}{\partial t} - \frac{\partial e_y}{\partial x} + \frac{\partial e_x}{\partial y} = 0,
 \end{aligned} \tag{1.126}$$

or in matrix form,

$$\underbrace{\begin{pmatrix} 0 & -\mathbf{b}^\top \\ \mathbf{b} & \mathbf{E}_\times \end{pmatrix}}_{\stackrel{\text{def}}{=} \mathbf{F}_{E,b}} \gamma \underbrace{\begin{pmatrix} -\frac{1}{c} \frac{\partial}{\partial t} \\ \nabla \end{pmatrix}}_{\vec{\nabla}} = \mathbf{o} \in \mathbb{R}^4. \tag{1.127}$$

The matrix  $\mathbf{F}_{E,b}$  is called the Maxwell's matrix.

So (1.125) and (1.127) contain the Maxwell's equations in a new four-dimensional form:

$$\boxed{\mathbf{F}_{B,e} \vec{\nabla} = \frac{1}{c} \vec{\mathbf{j}} \quad \text{and} \quad \mathbf{F}_{E,b} \vec{\nabla} = \mathbf{o}.} \tag{1.128}$$

This form has a great advantage of being invariant when transitioning to another reference system by a Lorentz transformation, i.e. in every inertial system it keeps the same external form. But the quantities appearing in it take different values in each reference system. This is shown in the following.

## 1.8.2 Lorentz Transformation of the Maxwell's Equations

Multiplying (1.125) from the left with  $L^{-1}$  and inserting  $L^{-1}L = \mathbf{I}$ , we obtain

$$\underbrace{L^{-1} \mathbf{F}_{B,e} L^{-1}}_{\mathbf{F}'_{B',e'}} L \vec{\nabla} = \frac{1}{c} \underbrace{L^{-1} \vec{\mathbf{j}}}_{\vec{\mathbf{j}}'}. \tag{1.129}$$

For the new Faraday's matrix  $\mathbf{F}'_{B',e'}$  we get

$$\begin{aligned}
 \mathbf{F}'_{B',e'} &= \begin{pmatrix} 0 & \mathbf{e}'^\top \\ -\mathbf{e}' & \mathbf{B}'_\times \end{pmatrix} \\
 &= \begin{pmatrix} \gamma_v & \frac{\gamma_v}{c} \mathbf{v}^\top \\ \frac{\gamma_v}{c} \mathbf{v} & \mathbf{A}(\mathbf{v}) \end{pmatrix} \begin{pmatrix} 0 & \mathbf{e}^\top \\ -\mathbf{e} & \mathbf{B}_\times \end{pmatrix} \begin{pmatrix} \gamma_v & \frac{\gamma_v}{c} \mathbf{v}^\top \\ \frac{\gamma_v}{c} \mathbf{v} & \mathbf{A}(\mathbf{v}) \end{pmatrix}.
 \end{aligned} \tag{1.130}$$

After multiplying the three matrices, we obtain for the bottom left corner vector  $\mathbf{e}'$  with  $\mathbf{A} = \mathbf{A}(\mathbf{v})$  and  $\gamma = \gamma_v$ :

$$\begin{aligned}
 -\mathbf{e}' &= \frac{\gamma}{c} \mathbf{A} \mathbf{B}_\times \mathbf{v} + \frac{\gamma^2}{c^2} \mathbf{v} \mathbf{e}^\top \mathbf{v} - \gamma \mathbf{A} \mathbf{e} \\
 &= \frac{\gamma}{c} \left( \mathbf{B}_\times \mathbf{v} + (\gamma - 1) \frac{\overbrace{\mathbf{v} \mathbf{v}^\top \mathbf{B}_\times \mathbf{v}}^0}{v^2} \right) + \frac{\gamma^2}{c^2} \mathbf{v} \mathbf{e}^\top \mathbf{v} - \gamma \mathbf{e} - \frac{\gamma(\gamma - 1)}{v^2} \mathbf{v} \mathbf{v}^\top \mathbf{e} \\
 &= \frac{\gamma}{c} \mathbf{B}_\times \mathbf{v} - \underbrace{\left( -\frac{\gamma^2}{c^2} + \frac{\gamma^2}{v^2} - \frac{\gamma}{v^2} \right)}_{\frac{1}{v^2}} \mathbf{v}^\top \mathbf{e} \mathbf{v} - \gamma \mathbf{e}, \\
 \mathbf{e}' &= \gamma \left( \mathbf{e} - \frac{1}{c} \mathbf{B}_\times \mathbf{v} \right) + \frac{(1 - \gamma)}{v^2} (\mathbf{v}^\top \mathbf{e}) \mathbf{v}, \tag{1.131}
 \end{aligned}$$

i.e.

$$\boxed{\mathbf{e}' = \gamma \left( \mathbf{e} + \frac{1}{c} \mathbf{v} \times \mathbf{b} \right) + (1 - \gamma) \frac{\mathbf{v}^\top \mathbf{e}}{v^2} \mathbf{v}.} \tag{1.132}$$

In the matrix  $\mathbf{F}'_{B',e'}$  in (1.130), for the right bottom  $3 \times 3$  sub-matrix  $\mathbf{B}'_\times$  after performing the matrix multiplication we obtain

$$\mathbf{B}'_\times = \mathbf{B}_\times + \frac{(\gamma - 1)}{v^2} (\mathbf{B}_\times \mathbf{v} \mathbf{v}^\top + \mathbf{v} \mathbf{v}^\top \mathbf{B}_\times) + \frac{\gamma}{c} (\mathbf{v} \mathbf{e}^\top - \mathbf{e} \mathbf{v}^\top). \tag{1.133}$$

From this matrix equation, the components of the magnetic induction vector  $\mathbf{b}'$  can be filtered out with the help of the row or column vectors of the unit matrix ( $\mathbf{i}_j$  is the  $j$ th column of the  $3 \times 3$  identity matrix  $\mathbf{I}_3$ ) as follows:

$$\mathbf{B}'_\times \stackrel{\text{def}}{=} \begin{pmatrix} 0 & -b'_z & b'_y \\ b'_z & 0 & -b'_x \\ -b'_y & b'_x & 0 \end{pmatrix}. \tag{1.134}$$

We thus obtain

$$b'_x = \mathbf{i}_3^\top \mathbf{B}'_\times \mathbf{i}_2, \tag{1.135}$$

$$b'_y = \mathbf{i}_1^\top \mathbf{B}'_\times \mathbf{i}_3, \tag{1.136}$$

$$b'_z = \mathbf{i}_2^\top \mathbf{B}'_\times \mathbf{i}_1. \tag{1.137}$$

Using (1.135) for (1.133) results in

$$\begin{aligned}
 b'_x &= \mathbf{i}_3^\top \mathbf{B}'_\times \mathbf{i}_2 = b_x + \frac{(\gamma - 1)}{v^2} [(-b_y v_1 + b_x v_2) v_2 + v_3 (-b_z v_1 + b_x v_3)] \\
 &\quad + \frac{\gamma}{c} (v_3 e_y - e_z v_2). \tag{1.138}
 \end{aligned}$$



Accordingly, one obtains for the two remaining components

$$\begin{aligned} b'_y &= \mathbf{i}_1^\top \mathbf{B}'_\times \mathbf{i}_3 = b_y + \frac{(\gamma - 1)}{v^2} [(-b_z v_2 + b_y v_3) v_3 + v_1 (b_y v_1 - b_x v_2)] \\ &\quad + \frac{\gamma}{c} (v_1 e_z - e_x v_3) \end{aligned} \quad (1.139)$$

and

$$b'_z = \mathbf{i}_2^\top \mathbf{B}'_\times \mathbf{i}_1 = b_z + \frac{(\gamma - 1)}{v^2} [(b_z v_1 - b_x v_3) v_1 + v_2 (b_z v_2 - b_y v_3)] + \frac{\gamma}{c} (v_2 e_x - e_y v_1). \quad (1.140)$$

Summarizing the last terms in (1.138), (1.139) and (1.140) results in the vector product

$$\frac{\gamma}{c} \mathbf{e} \times \mathbf{v}. \quad (1.141)$$

For the second to last summands one obtains the vector

$$\frac{(\gamma - 1)}{v^2} \begin{pmatrix} (-b_y v_1 + b_x v_2) v_2 + v_3 (-b_z v_1 + b_x v_3) \\ (-b_z v_2 + b_y v_3) v_3 + v_1 (b_y v_1 - b_x v_2) \\ (b_z v_1 - b_x v_3) v_1 + v_2 (b_z v_2 - b_y v_3) \end{pmatrix},$$

and using  $0 = b_x v_1^2 - b_x v_1^2$  in the first component,  $0 = b_y v_2^2 - b_y v_2^2$  in the second, and  $0 = b_z v_3^2 - b_z v_3^2$  in the third,

$$\begin{aligned} &\frac{(\gamma - 1)}{v^2} \begin{pmatrix} (-b_x v_1^2 - b_y v_1 v_2 - b_z v_1 v_3) + (b_x v_1^2 + b_x v_2^2 + b_x v_3^2) \\ (-b_x v_1 v_2 - b_y v_2^2 - b_z v_2 v_3) + (b_y v_1^2 + b_y v_2^2 + b_y v_3^2) \\ (-b_x v_1 v_3 - b_y v_2 v_3 - b_z v_3^2) + (b_z v_1^2 + b_z v_2^2 + b_z v_3^2) \end{pmatrix} \\ &= \frac{(\gamma - 1)}{v^2} (-\mathbf{b}^\top \mathbf{v}) \mathbf{v} + v^2 \mathbf{b}. \end{aligned} \quad (1.142)$$

Equations (1.141) and (1.142), together with (1.138), (1.139) and (1.140), yield the final result

$$\boxed{\mathbf{b}' = \gamma \left( \mathbf{b} - \frac{1}{c} \mathbf{v} \times \mathbf{e} \right) + (1 - \gamma) \frac{\mathbf{v}^\top \mathbf{b}}{v^2} \mathbf{v}.} \quad (1.143)$$

One gets the same result for  $\mathbf{e}'$  and  $\mathbf{b}'$  by a Lorentz transformation of  $\mathbf{F}_{E,b}$ !

The formulas (1.132) and (1.143) are simplified considerably for small velocities  $v \ll c$  because then  $\gamma \approx 1$  and one obtains

$$\mathbf{e}' = \mathbf{e} + \frac{1}{c} \mathbf{v} \times \mathbf{b}$$

and

$$\mathbf{b}' = \mathbf{b} - \frac{1}{c} \mathbf{v} \times \mathbf{e}.$$

The decomposition of the electromagnetic field in an electric and magnetic field has no absolute significance. If there exists, for example, in a reference system  $\mathcal{X}$  a purely electrostatic field with  $\mathbf{b} = \mathbf{o}$ , then, due to (1.143), a magnetic field  $\mathbf{b}' = \frac{\gamma}{c} \mathbf{v} \times \mathbf{e} \neq \mathbf{o}$  exists in a reference system  $\mathcal{X}'$  that moves with the speed  $\mathbf{v}$  relative to the reference system  $\mathcal{X}$ . Physically this means that all charges rest in  $\mathcal{X}$ . But these charges move relative to  $\mathcal{X}'$  with the velocity  $\mathbf{v}$ . So there exists a current in  $\mathcal{X}'$  which generates a magnetic field in  $\mathcal{X}'$ .

The equations (1.132) and (1.143) can be summarized in one equation:

$$\boxed{\begin{pmatrix} \mathbf{b}' \\ \mathbf{e}' \end{pmatrix} = \underbrace{\begin{pmatrix} \gamma \mathbf{I} + (1 - \gamma) \frac{\mathbf{v} \mathbf{v}^\top}{v^2} & -\frac{\gamma}{c} \mathbf{V}_\times \\ \frac{\gamma}{c} \mathbf{V}_\times & \gamma \mathbf{I} + (1 - \gamma) \frac{\mathbf{v} \mathbf{v}^\top}{v^2} \end{pmatrix}}_{\mathbf{P}(\mathbf{v})} \begin{pmatrix} \mathbf{b} \\ \mathbf{e} \end{pmatrix} \in \mathbb{R}^6,} \quad (1.144)$$

with

$$\mathbf{V}_\times \stackrel{\text{def}}{=} \begin{pmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{pmatrix}.$$

The symmetric  $6 \times 6$ -matrix  $\mathbf{P}(\mathbf{v})$  occurring in (1.144) has a formal similarity with the Lorentz matrix  $\mathbf{L}$ !

It is, of course, true that  $\mathbf{P}(-\mathbf{v}) = \mathbf{P}^{-1}(\mathbf{v})$ , for, with

$$\mathbf{V}_\times^2 = \mathbf{v} \mathbf{v}^\top - v^2 \mathbf{I}, \quad (1.145)$$

one obtains easily  $\mathbf{P}(\mathbf{v})\mathbf{P}(-\mathbf{v}) = \mathbf{I}$ .

In the special case when only one component of the velocity vector in the  $x$ -direction is nonzero, i.e.  $\mathbf{v} = [v, 0, 0]^\top$ , one obtains the matrix

$$\mathbf{P}(\mathbf{v}) = \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \gamma & 0 & 0 & 0 & \frac{\gamma v}{c} \\ 0 & 0 & \gamma & 0 & -\frac{\gamma v}{c} & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{\gamma v}{c} & 0 & \gamma & 0 \\ 0 & \frac{\gamma v}{c} & 0 & 0 & 0 & \gamma \end{array} \right). \quad (1.146)$$

This special symmetric  $6 \times 6$ -matrix can also be found in [33].

### 1.8.3 Electromagnetic Invariants

For the electromagnetic field quantities  $\mathbf{e}$  and  $\mathbf{b}$  one can form *invariants* which do not change during the transition to another inertial system. For the scalar product of

the electric field strength  $\mathbf{e}$  and the magnetic induction  $\mathbf{b}$  we obtain

$$\begin{aligned} \mathbf{e}'^\top \mathbf{b}' &= \left( \gamma \left( \mathbf{e}^\top - \frac{1}{c} (\mathbf{v} \times \mathbf{b})^\top \right) + (1 - \gamma) \frac{\mathbf{v}^\top \mathbf{e}}{v^2} \mathbf{v}^\top \right) \left( \gamma \left( \mathbf{b} + \frac{1}{c} \mathbf{v} \times \mathbf{e} \right) \right. \\ &\quad \left. + (1 - \gamma) \frac{\mathbf{v}^\top \mathbf{b}}{v^2} \mathbf{v} \right) \\ &= \gamma^2 \mathbf{e}^\top \mathbf{b} + [2\gamma(1 - \gamma) + (1 - \gamma^2)] \frac{\mathbf{e}^\top \mathbf{v} \mathbf{v}^\top \mathbf{b}}{v^2} - \frac{\gamma^2}{c^2} \underbrace{(\mathbf{v} \times \mathbf{b})^\top (\mathbf{v} \times \mathbf{e})}_{\mathbf{b}^\top \mathbf{V}_\times^\top \mathbf{V}_\times \mathbf{e}}. \end{aligned}$$

With  $\mathbf{V}_\times^\top = -\mathbf{V}_\times$  one obtains due to (1.145)

$$\mathbf{V}_\times^\top \mathbf{V}_\times = -\mathbf{V}_\times^2 = v^2 \mathbf{I} - \mathbf{v} \mathbf{v}^\top.$$

This gives

$$\underline{\underline{\mathbf{e}'^\top \mathbf{b}'}} = \gamma^2 \mathbf{e}^\top \mathbf{b} - \frac{\gamma^2}{c^2} (\mathbf{v}^\top \mathbf{e} \mathbf{v}^\top \mathbf{b} + v^2 \mathbf{b}^\top \mathbf{e} - \mathbf{b}^\top \mathbf{v} \mathbf{v}^\top \mathbf{e}) = \left( \gamma^2 - \frac{\gamma^2 v^2}{c^2} \right) \mathbf{e}^\top \mathbf{b} = \underline{\underline{\mathbf{e}^\top \mathbf{b}}},$$

so the scalar product of the electric field intensity  $\mathbf{e}$  and the magnetic induction  $\mathbf{b}$  is invariant with respect to a Lorentz transformation!

Using a slightly long calculation, one can show that the difference of the squares of the field intensities is also invariant:

$$\underline{\underline{b'^2 - e'^2}} = \underline{\underline{b^2 - e^2}}. \quad (1.147)$$

You can also arrive at electromagnetic invariants in another way, namely with the help of the invariant Faraday's matrix

$$\mathbf{F}_{B,e} = \begin{pmatrix} 0 & \mathbf{e}^\top \\ -\mathbf{e} & \mathbf{B}_\times \end{pmatrix}$$

and the invariant Maxwell's matrix

$$\mathbf{F}_{E,b} = \begin{pmatrix} 0 & -\mathbf{b}^\top \\ \mathbf{b} & \mathbf{E}_\times \end{pmatrix}.$$

With

$$\mathbf{F}_{B,e}^* \stackrel{\text{def}}{=} \mathbf{M} \mathbf{F}_{B,e} \mathbf{M} = \begin{pmatrix} 0 & -\mathbf{e}^\top \\ \mathbf{e} & \mathbf{B}_\times \end{pmatrix},$$

one gets, e.g.

$$\mathbf{F}_{B,e}^* \mathbf{F}_{B,e} = \begin{pmatrix} 0 & -\mathbf{e}^\top \\ \mathbf{e} & \mathbf{B}_\times \end{pmatrix} \begin{pmatrix} 0 & \mathbf{e}^\top \\ -\mathbf{e} & \mathbf{B}_\times \end{pmatrix} = \begin{pmatrix} \mathbf{e}^\top \mathbf{e} & -\mathbf{e}^\top \mathbf{B}_\times \\ -\mathbf{B}_\times \mathbf{e} & \mathbf{e} \mathbf{e}^\top + \mathbf{B}_\times \mathbf{B}_\times \end{pmatrix}$$

$$= \begin{pmatrix} e^2 & & & & & \\ & \mathbf{s}^\top & & & & \\ & & e_x^2 - b_z^2 - b_y^2 & & & \\ & & \dots & & \dots & \\ \mathbf{s} & & \dots & e_y^2 - b_z^2 - b_x^2 & & \\ & & \dots & \dots & & \\ & & & & & e_z^2 - b_y^2 - b_x^2 \end{pmatrix},$$

with the Poynting vector  $\mathbf{s} \stackrel{\text{def}}{=} \mathbf{e} \times \mathbf{b}$ . By forming the trace of this matrix product, i.e. taking the sum of the matrix elements on the main diagonal, one obtains

$$\text{trace}(\mathbf{F}_{B,e}^* \mathbf{F}_{B,e}) = 2e^2 - 2b^2,$$

i.e. the above invariant in the form

$$\boxed{-\frac{1}{2} \text{trace}(\mathbf{F}_{B,e}^* \mathbf{F}_{B,e}) = b^2 - e^2.} \quad (1.148)$$

One obtains the second invariant  $\mathbf{e}^\top \mathbf{b}$  above by taking the trace of the product of the modified Faraday's matrix  $\mathbf{F}_{B,e}^*$  and the Maxwell's matrix  $\mathbf{F}_{E,b}$ :

$$\boxed{-\frac{1}{4} \text{trace}(\mathbf{F}_{B,e}^* \mathbf{F}_{E,e}) = \mathbf{e}^\top \mathbf{b}.} \quad (1.149)$$

Making invariants of this type later plays a role in the consideration of the singularities of the Schwarzkopf's solution of General Relativity equations.

### 1.8.4 Electromagnetic Forces

We want to determine the force acting on a charged particle with the charge  $q$ , which is in an electromagnetic field moving with the velocity  $\mathbf{u}$  relative to an inertial frame  $\mathcal{X}$ . Let  $\mathcal{X}'$  be the inertial frame in which the particle rests at the moment. In this system, due to  $\mathbf{u}' = \mathbf{o}$  and  $\mathbf{u}' \times \mathbf{b}' = \mathbf{o}$ ,

$$m_0 \frac{d\mathbf{u}'}{dt'} = q\mathbf{e}' \in \mathbb{R}^3. \quad (1.150)$$

Generally, due to the relativity principle, the Lorentz force in  $\mathcal{X}$  is

$$\mathbf{f} = q \left( \mathbf{e} + \frac{1}{c} \mathbf{u} \times \mathbf{b} \right) = \frac{q}{c} [\mathbf{e} \mid -\mathbf{B}_\times] \begin{pmatrix} c \\ \mathbf{u} \end{pmatrix} \in \mathbb{R}^3. \quad (1.151)$$

The following law expresses how an electromagnetic field acts on a stationary charge  $q$  and current  $\mathbf{j}$ :

$$\mathbf{f} = q\mathbf{e} + \frac{1}{c} \mathbf{j} \times \mathbf{b} = \frac{1}{c} [\mathbf{e} \mid -\mathbf{B}_\times] \begin{pmatrix} cq \\ \mathbf{j} \end{pmatrix} \in \mathbb{R}^3. \quad (1.152)$$

$\mathbf{f}$  is completed to a four-vector  $\vec{\mathbf{f}}$  as in (1.108)

$$\vec{\mathbf{f}} \stackrel{\text{def}}{=} \gamma_u \begin{pmatrix} \mathbf{u}^\top \mathbf{f}/c \\ \mathbf{f} \end{pmatrix} \in \mathbb{R}^4. \quad (1.153)$$

Equation (1.151) applied to (1.153) yields (with  $\mathbf{u}^\top(\mathbf{u} \times \mathbf{b}) = 0$ )

$$\begin{aligned} \vec{\mathbf{f}} &= q\gamma_u \begin{pmatrix} \mathbf{u}^\top \mathbf{e}/c \\ \mathbf{e} + \frac{1}{c} \mathbf{u} \times \mathbf{b} \end{pmatrix} = q \frac{\gamma_u}{c} \underbrace{\begin{pmatrix} 0 & \mathbf{e}^\top \\ \mathbf{e} & -\mathbf{B}_\times \end{pmatrix}}_{\mathbf{MF}_{B,e}} \begin{pmatrix} c \\ \mathbf{u} \end{pmatrix} \\ &= \frac{q}{c} \mathbf{MF}_{B,e} \underbrace{\gamma_u \begin{pmatrix} c \\ \mathbf{u} \end{pmatrix}}_{\vec{\mathbf{u}}}, \end{aligned} \quad (1.154)$$

so

$$\boxed{\vec{\mathbf{f}} = \frac{q}{c} \mathbf{MF}_{B,e} \vec{\mathbf{u}}.} \quad (1.155)$$

With (1.152) one can write instead of (1.155)

$$\boxed{\vec{\mathbf{f}} = \frac{1}{c} \mathbf{F}_{B,e}^* \vec{\mathbf{j}},} \quad (1.156)$$

in which again

$$\vec{\mathbf{j}} = \gamma \begin{pmatrix} cq \\ -\mathbf{j} \end{pmatrix} \in \mathbb{R}^4.$$

Subjecting (1.155) to a linear transformation using the Lorentz matrix  $\mathbf{L}(\mathbf{v})$ , one obtains

$$\vec{\mathbf{f}}' = \mathbf{L}(\mathbf{v}) \vec{\mathbf{f}} = \frac{q}{c} \underbrace{\mathbf{L}(\mathbf{v}) \mathbf{MF}_{B,e} \mathbf{L}^{-1}(\mathbf{v})}_{\mathbf{MF}'_{B',e'}} \underbrace{\mathbf{L}(\mathbf{v}) \vec{\mathbf{u}}}_{\vec{\mathbf{u}}'}. \quad (1.157)$$

i.e.

$$\boxed{\vec{\mathbf{f}}' = \frac{q}{c} \mathbf{MF}'_{B',e'} \vec{\mathbf{u}}'.} \quad (1.158)$$

The power equation is also invariant with respect to a Lorentz transformation!

## 1.9 The Energy–Momentum Matrix

### 1.9.1 The Electromagnetic Energy–Momentum Matrix

We want to derive *one* equation which contains the fundamental dynamic equations of the theory of electricity. This equation should also include the energy theorem and the momentum theorems of electrodynamics. We will find again the therein contained energy–momentum matrix in the main equation of the theory of general relativity.

Due to (1.128), one has

$$\mathbf{F}_{B,e} \vec{\nabla} = \frac{1}{c} \vec{\mathbf{j}}$$

and because of (1.156)

$$\vec{\mathbf{f}} = \frac{1}{c} \mathbf{F}_{B,e}^* \vec{\mathbf{j}}.$$

Is there a matrix such that

$$\vec{\mathbf{f}} = \mathbf{T}_{b,e} \vec{\nabla} \quad (1.159)$$

We try  $\mathbf{T}_{b,e} = \mathbf{F}_{B,e}^* \mathbf{F}_{B,e}$ . Then

$$\begin{aligned} \mathbf{T}_{b,e} \vec{\nabla} &= (\mathbf{F}_{B,e}^* \mathbf{F}_{B,e}) \vec{\nabla} = \begin{pmatrix} \mathbf{e}^\top \mathbf{e} & -\mathbf{e}^\top \mathbf{B}_\times \\ -\mathbf{B}_\times \mathbf{e} & \mathbf{e} \mathbf{e}^\top + \mathbf{B}_\times^2 \end{pmatrix} \vec{\nabla} \\ &= \begin{pmatrix} e^2 & s^\top \\ s & \mathbf{e} \mathbf{e}^\top + \mathbf{b} \mathbf{b}^\top - b^2 \mathbf{I}_3 \end{pmatrix} \vec{\nabla} \stackrel{!}{=} \vec{\mathbf{f}}, \end{aligned} \quad (1.160)$$

by considering the Poynting vector

$$\mathbf{s} \stackrel{\text{def}}{=} \mathbf{e} \times \mathbf{b} \quad (1.161)$$

and the relation  $\mathbf{B}_\times^2 = \mathbf{b} \mathbf{b}^\top - b^2 \mathbf{I}$  in conformance with (1.145). The Poynting vector gives the magnitude and direction of the energy transport in electromagnetic fields. What is the result when differentiating the first line of (1.160)? It is

$$-\frac{1}{c} \frac{\partial(e^2)}{\partial t} + \nabla^\top \mathbf{s} \stackrel{!}{=} \frac{\rho \mathbf{u}^\top \mathbf{f}}{c}.$$

$\frac{1}{c} \frac{\partial(e^2)}{\partial t}$  is proportional to the temporal change of the energy density of the electromagnetic field when no magnetic field is present. But then also  $\mathbf{b} = \mathbf{o}$ , i.e.  $\mathbf{s} = \mathbf{o}$ , and thus the whole equation is without a statement. The matter would be different if instead of  $e^2$  in top left corner of the matrix  $\mathbf{T}_{b,e}$ , one would have the expression  $(e^2 + b^2)/2$ , as then

$$\frac{1}{2c} \frac{\partial(e^2 + b^2)}{\partial t} = \frac{1}{c} \frac{\partial w}{\partial t},$$

namely, the temporal change of the energy density  $w = (e^2 + b^2)/2$  of the electromagnetic field. By adding  $(b^2 - e^2)/2$  to the upper left, there arises, in fact,  $(e^2 + b^2)/2$ ! Therefore, now the

**Definition:** The *electromagnetic energy–momentum matrix*  $\mathbf{T}_{b,e}$  has the form

$$\mathbf{T}_{b,e} \stackrel{\text{def}}{=} \mathbf{F}_{B,e}^* \mathbf{F}_{B,e} + \frac{1}{2}(b^2 - e^2)\mathbf{I}_4 = \begin{pmatrix} w & \mathbf{s}^\top \\ \mathbf{s} & (\mathbf{e}\mathbf{e}^\top + \mathbf{b}\mathbf{b}^\top - w\mathbf{I}_3) \end{pmatrix}, \quad (1.162)$$

with

$$w \stackrel{\text{def}}{=} \frac{1}{2}(e^2 + b^2) \quad (1.163)$$

and

$$\mathbf{T}_{b,e} \cdot \vec{\nabla} = \frac{1}{\gamma_v} \vec{\mathbf{f}}. \quad (1.164)$$

While doing this,  $\vec{\nabla}$  again has the form

$$\vec{\nabla} = \gamma_u \left( -\frac{1}{c} \frac{\partial}{\partial t} \right)$$

and

$$\vec{\mathbf{f}} = \gamma_u \left( \frac{\rho}{c} \mathbf{u}^\top \mathbf{f} \right).$$

For the first row of (1.164) now one obtains

$$-\frac{1}{c} \frac{\partial w}{\partial t} + \mathbf{s}^\top \nabla = \frac{\rho}{c} \mathbf{u}^\top \mathbf{f},$$

i.e.

$$c \operatorname{div} \mathbf{s} = \frac{\partial w}{\partial t} + \rho \mathbf{u}^\top \mathbf{f}. \quad (1.165)$$

On the left is the infinitely small volume unit of the entering or exiting energy flow, and it consists of the temporal change of the energy density  $\frac{\partial w}{\partial t}$  and the conversion of the electromagnetic energy into mechanical energy per unit time and volume  $\rho \mathbf{u}^\top \mathbf{f}$ . So the whole is the *Energy Theorem of Electrodynamics*.

Next we get for the second to fourth components of (1.164):

$$-\frac{1}{c} \frac{\partial \mathbf{s}}{\partial t} + (\mathbf{e}\mathbf{e}^\top) \nabla + (\mathbf{b}\mathbf{b}^\top) \nabla - (w\mathbf{I}_3) \nabla = \mathbf{f}. \quad (1.166)$$

With the help of Maxwell's equations, we obtain for the first term on the left-hand side:

$$-\frac{1}{c} \frac{\partial \mathbf{s}}{\partial t} = \frac{1}{c} \frac{\partial \mathbf{e} \times \mathbf{b}}{\partial t} = \frac{1}{c} \frac{\partial \mathbf{e}}{\partial t} \times \mathbf{b} + \frac{1}{c} \mathbf{e} \times \frac{\partial \mathbf{b}}{\partial t}$$

$$\begin{aligned}
&= \left( \mathbf{curl} \mathbf{b} - \frac{1}{c} \mathbf{j} \right) \times \mathbf{b} + \mathbf{e} \times (-\mathbf{curl} \mathbf{e}) \\
&= -\frac{1}{c} \mathbf{j} \times \mathbf{b} - \mathbf{b} \times \mathbf{curl} \mathbf{b} - \mathbf{e} \times \mathbf{curl} \mathbf{e}. \tag{1.167}
\end{aligned}$$

For the first component, the  $x$ -component of the three-dimensional vector  $(\mathbf{e}\mathbf{e}^\top)\nabla - \frac{1}{c}(e^2\mathbf{I}_3)\nabla$  on the left-hand side, one obtains

$$\begin{aligned}
&\left[ e_x^2 - \frac{1}{2}e^2|e_x e_y|e_x e_z \right] \nabla \\
&= 2e_x \frac{\partial e_x}{\partial x} - \left( e_x \frac{\partial e_x}{\partial x} + e_y \frac{\partial e_y}{\partial y} + e_z \frac{\partial e_z}{\partial z} \right) + \frac{\partial e_x}{\partial y} e_y + e_x \frac{\partial e_y}{\partial y} + \frac{\partial e_x}{\partial z} e_z + e_x \frac{\partial e_z}{\partial z} \\
&= e_x \left( \frac{\partial e_x}{\partial x} + \frac{\partial e_y}{\partial y} + \frac{\partial e_z}{\partial z} \right) + e_z \frac{\partial e_x}{\partial z} + e_y \frac{\partial e_x}{\partial y} - e_z \frac{\partial e_z}{\partial x} - e_y \frac{\partial e_y}{\partial x} \\
&= e_x \operatorname{div} \mathbf{e} - (\mathbf{e} \times \mathbf{curl} \mathbf{e})_x = e_x \cdot \rho - (\mathbf{e} \times \mathbf{curl} \mathbf{e})_x. \tag{1.168}
\end{aligned}$$

Accordingly, one receives for the  $y$ - and  $z$ -component together

$$(\mathbf{e}\mathbf{e}^\top)\nabla - \frac{1}{2}(e^2\mathbf{I}_3)\nabla = \mathbf{e} \cdot \rho - \mathbf{e} \times \mathbf{curl} \mathbf{e}. \tag{1.169}$$

Also we obtain with  $\operatorname{div} \mathbf{b} = 0$ :

$$(\mathbf{b}\mathbf{b}^\top)\nabla - \frac{1}{2}(b^2\mathbf{I}_3)\nabla = -\mathbf{b} \times \mathbf{curl} \mathbf{b}. \tag{1.170}$$

Overall, (1.165), (1.168) and (1.170) result for the second to fourth rows of (1.164) in

$$\mathbf{e} \cdot \rho + \frac{1}{c} \mathbf{j} \times \mathbf{b} = \mathbf{f}. \tag{1.171}$$

But this is the *Momentum Theorem* of Lorentz! Thus the above assertion (1.164) is completely proved.

## 1.9.2 The Mechanical Energy–Momentum Matrix

A particle at rest at the point  $\mathbf{x}_0$  has the speed

$$\vec{\mathbf{u}} = \frac{d\vec{\mathbf{x}}}{d\tau} = \frac{d}{d\tau} \begin{pmatrix} c\tau \\ \mathbf{x}_0 \end{pmatrix} = \begin{pmatrix} c \\ \mathbf{o} \end{pmatrix}$$

and the momentum  $\vec{\mathbf{p}} = m_0 \begin{pmatrix} c \\ \mathbf{o} \end{pmatrix}$ . The zeroth component of  $\vec{\mathbf{p}}$  is the rest energy of the particle divided by  $c$ . For a moving particle, with  $E = \gamma_u m_0 c^2$ , one gets

$$\vec{\mathbf{p}} = \gamma_v m_0 \begin{pmatrix} c \\ \mathbf{v} \end{pmatrix} = \begin{pmatrix} E/c \\ \gamma_v \mathbf{p} \end{pmatrix}.$$



This equation expresses the fact that in the theory of relativity energy and momentum are the temporal and spatial components of the four-vector  $\vec{p}$ . They preserve this distinction even after a Lorentz transformation just like the four-vector  $\vec{x} = \begin{pmatrix} ct \\ \mathbf{x} \end{pmatrix}$ : its zeroth component is always the time component and the rest represents the space components.

We now go over to a distributed matter, such as that in a perfect fluid, i.e. a liquid without internal friction, but of quite variable density. It is described by the two scalar fields, density  $\rho$  and pressure  $p$ , and the velocity vector field  $\vec{u}$ . The aim of this derivation of the energy–momentum matrix is that this matrix somehow represents the energy content of the liquid and the transition to the curved world of the theory of general relativity, which can serve as a source of the field of gravity. The *continuity equation* describes the conservation of mass. The conservation of mass requires that the change of mass  $dV\delta\rho/\delta t$  in the unit of time, associated with the local compression  $\delta\rho/\delta t$ , must be equal to the difference between the entering and exiting masses per unit of time. In the  $x$ -direction, for this difference the following is valid:

$$\rho u_x(x) \cdot dy \cdot dz - \left( \rho u_x + \frac{\partial \rho u_x}{\partial x} dx \right) dy dz = - \frac{\partial \rho u_x}{\partial x} dV.$$

One gets similar expressions for the  $y$ - and  $z$ -direction, making a total for the conservation of mass

$$\frac{\partial \rho}{\partial t} dV = - \left( \frac{\partial \rho u_x}{\partial x} + \frac{\partial \rho u_y}{\partial y} + \frac{\partial \rho u_z}{\partial z} \right) dV.$$

So this is the final continuity equation in differential form:

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho u_x}{\partial x} + \frac{\partial \rho u_y}{\partial y} + \frac{\partial \rho u_z}{\partial z} = \frac{\partial \rho}{\partial t} + \text{div}(\rho \mathbf{u}) = 0. \quad (1.172)$$

The *dynamic* behaviour is described by the Euler's equation. By using Newton's law on the mass contained in a volume element of a perfect fluid, we obtain the Euler equation of motion, initially only in the  $x$ -direction:

$$dm \frac{du_x}{dt} = \rho dx dy dz \frac{du_x}{dt} = dx dy dz f_{D,x} - \left( \frac{\partial p}{\partial x} dx \right) dy dz,$$

which implies

$$\rho \frac{du_x}{dt} = f_{D,x} - \frac{\partial p}{\partial x}, \quad (1.173)$$

where  $f_{D,x}$  is the  $x$ -component of the force per unit volume (power density),  $f_D$ , e.g. the gravitational force. Using the total differential of  $\Delta u_x$ ,

$$\Delta u_x = \frac{\partial u_x}{\partial t} \Delta t + \frac{\partial u_x}{\partial x} \Delta x + \frac{\partial u_x}{\partial y} \Delta y + \frac{\partial u_x}{\partial z} \Delta z,$$

dividing by  $\Delta t$  and passing to the limit  $\Delta t \rightarrow 0$ , we obtain

$$\frac{du_x}{dt} = \frac{\partial u_x}{\partial t} + \frac{\partial u_x}{\partial x}u_x + \frac{\partial u_x}{\partial y}u_y + \frac{\partial u_x}{\partial z}u_z. \quad (1.174)$$

With the corresponding equations for the  $y$ - and  $z$ -direction, we obtain in total

$$\begin{aligned} \rho \left( \frac{\partial u_x}{\partial t} + \frac{\partial u_x}{\partial x}u_x + \frac{\partial u_x}{\partial y}u_y + \frac{\partial u_x}{\partial z}u_z \right) &= f_{D,x} - \frac{\partial p}{\partial x}, \\ \rho \left( \frac{\partial u_y}{\partial t} + \frac{\partial u_y}{\partial x}u_x + \frac{\partial u_y}{\partial y}u_y + \frac{\partial u_y}{\partial z}u_z \right) &= f_{D,y} - \frac{\partial p}{\partial y}, \\ \rho \left( \frac{\partial u_z}{\partial t} + \frac{\partial u_z}{\partial x}u_x + \frac{\partial u_z}{\partial y}u_y + \frac{\partial u_z}{\partial z}u_z \right) &= f_{D,z} - \frac{\partial p}{\partial z}, \end{aligned}$$

summarized in

$$\underline{\underline{\rho \left( \frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{u}}{\partial \mathbf{x}^\top} \mathbf{u} \right) + \mathbf{grad} p = \mathbf{f}_D.}} \quad (1.175)$$

This is the Euler's equation in modern form. The first term in the parentheses is called the local and the second the convective change.

The relativistic generalization of the hydrodynamic equations (1.172) and (1.175) will now be established.  $\rho_0$  is the *rest density*, defined as the rest mass per rest volume. With  $\vec{\mathbf{x}}^\top = [ct|\mathbf{x}^\top]$  and  $\vec{\mathbf{u}}^\top = \gamma_u[c|\mathbf{u}^\top]$ , the Euler's equation (1.175) can also be written as

$$\rho \frac{\partial \mathbf{u}}{\partial \vec{\mathbf{x}}^\top} \vec{\mathbf{u}} = \mathbf{f}_D - \mathbf{grad} p. \quad (1.176)$$

With regard to the subsequent application of the operator  $\vec{\nabla}$ , we now start for the energy–momentum matrix with

$$\mathbf{T}_{\text{mech},1} \stackrel{\text{def}}{=} \rho_0 \vec{\mathbf{u}} \vec{\mathbf{u}}^\top. \quad (1.177)$$

This matrix is symmetric and built up from the two values  $\rho_0$  and  $\vec{\mathbf{u}}$ , which describe completely the dynamics of a perfect fluid with the pressure  $p$  and the acting external forces  $\mathbf{f}_D$ . For the use of the operator  $\vec{\nabla}$  and the further investigation of the result, it is advantageous to first divide the matrix  $\mathbf{T}_{\text{mech},1}$  similarly as the matrix  $\mathbf{T}_{b,e}$ :

$$\mathbf{T}_{\text{mech},1} = \rho_0 \gamma_u^2 \begin{pmatrix} c^2 & c\mathbf{u}^\top \\ c\mathbf{u} & \mathbf{u}\mathbf{u}^\top \end{pmatrix}. \quad (1.178)$$

In a liquid moving with the velocity  $\mathbf{u}$ , the volume decreases with  $\gamma_u$  while the mass increases with  $\gamma_u$ , thus one obtains the total density  $\rho = \gamma_u^2 \rho_0$ . With this we now

define

$$\mathbf{T}_{\text{mech},1} \stackrel{\text{def}}{=} \begin{pmatrix} \rho c^2 & \rho \mathbf{c} \mathbf{u}^\top \\ \rho \mathbf{c} \mathbf{u} & \rho \mathbf{u} \mathbf{u}^\top \end{pmatrix}. \quad (1.179)$$

Multiplying the first line in (1.179) from the right with the operator  $\vec{\nabla}$  yields

$$-c \frac{\partial \rho}{\partial t} + c(\rho \mathbf{u}^\top) \nabla.$$

By setting this expression equal to zero, we obtain with  $(\rho \mathbf{u}^\top) \nabla = \text{div}(\rho \mathbf{u})$  the classical continuity equation (1.172)!

Multiplying now the second row of the matrix in (1.179) from the right with the operator  $\vec{\nabla}$ , one obtains

$$\begin{aligned} -\frac{\partial \rho \mathbf{u}}{\partial t} + (\rho \mathbf{u} \mathbf{u}^\top) \nabla &= -\rho \frac{\partial \mathbf{u}}{\partial t} - \frac{\partial \rho}{\partial t} \mathbf{u} + \rho \frac{\partial \mathbf{u}}{\partial \mathbf{x}^\top} \mathbf{u} + \mathbf{u} \text{div}(\rho \mathbf{u}) \\ &= \left( -\frac{\partial \rho}{\partial t} \mathbf{u} + \text{div}(\rho \mathbf{u}) \right) \mathbf{u} + \rho \left( -\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{u}}{\partial \mathbf{x}^\top} \mathbf{u} \right). \end{aligned} \quad (1.180)$$

The term inside the first parentheses is equal to zero, due to the continuity equation in the non-relativistic case, and the term in the second parentheses contains the force- and pressure-free Euler's equation!

The pressure  $p$  must now be incorporated. Presuppose an isotropic liquid, then the pressure  $p$  is direction-independent. In the Euler's equation, the pressure appears in the form of **grad**  $p$  which can be written as

$$\mathbf{grad} p = \begin{pmatrix} p & 0 & 0 \\ 0 & p & 0 \\ 0 & 0 & p \end{pmatrix} \nabla. \quad (1.181)$$

If this is taken into account in the matrix  $\mathbf{T}_{\text{mech}}$ , it must be remembered that (1.181) holds for a reference system  $\mathcal{X}'$  moving with the liquid; therefore, this approach makes sense:

$$\mathbf{T}'_{\text{mech},2} \stackrel{\text{def}}{=} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix}. \quad (1.182)$$

This matrix can now with the help of the Lorentz matrix

$$\mathbf{L}(-\mathbf{u}) = \begin{pmatrix} \gamma_u & \frac{\gamma_u}{c} \mathbf{u}^\top \\ \frac{\gamma_u}{c} \mathbf{u} & \mathbf{I}_3 + (\gamma_u - 1) \frac{\mathbf{u} \mathbf{u}^\top}{u^2} \end{pmatrix}$$

be transformed back into the resting inertial system  $\mathcal{X}$ :

$$\mathbf{T}_{\text{mech},2} = \mathbf{L}(-\mathbf{u}) \mathbf{T}'_{\text{mech},2} \mathbf{L}^\top(-\mathbf{u}) = p \begin{pmatrix} \frac{\gamma_u^2 u^2}{c^2} & \frac{\gamma_u^2}{c} \mathbf{u}^\top \\ \frac{\gamma_u^2}{c} \mathbf{u} & \mathbf{I}_3 + \frac{\gamma_u^2}{c^2} \mathbf{u} \mathbf{u}^\top \end{pmatrix}.$$

Considering that  $p\gamma^2 u^2/c^2 = p(\gamma^2 - 1)$ , this can also be written as

$$\mathbf{T}_{\text{mech},2} = \frac{p}{c^2} \bar{\mathbf{u}}\bar{\mathbf{u}}^\top + p \begin{pmatrix} -1 & \mathbf{o}^\top \\ \mathbf{o} & \mathbf{I}_3 \end{pmatrix}.$$

For the sum of the two matrices  $\mathbf{T}_{\text{mech},1}$  and  $\mathbf{T}_{\text{mech},2}$ , we finally obtain with the Minkowski matrix  $\mathbf{M}$

$$\underline{\underline{\mathbf{T}_{\text{mech}}}} \stackrel{\text{def}}{=} \left( \rho_0 + \frac{p}{c^2} \right) \bar{\mathbf{u}}\bar{\mathbf{u}}^\top - p\mathbf{M}. \quad (1.183)$$

We now summarize everything together to the relativistic generalization of the hydrodynamic equations

$$\boxed{\mathbf{T}_{\text{mech}} \vec{\nabla} = \vec{f}_D}. \quad (1.184)$$

By the way, one obtains, when the energy–momentum matrix  $\mathbf{T}_{\text{mech}}$  from the right is multiplied with  $\mathbf{M}\bar{\mathbf{u}}$ , the four-vector momentum density  $\rho_0\bar{\mathbf{u}}$  multiplied by  $c^2$ :

$$\mathbf{T}_{\text{mech}}\mathbf{M}\bar{\mathbf{u}} = \left( \rho_0 + \frac{p}{c^2} \right) \bar{\mathbf{u}} \underbrace{\gamma^2 (c^2 - v^2)}_{c^2} - p\bar{\mathbf{u}} = c^2 \rho_0 \bar{\mathbf{u}}. \quad (1.185)$$

### 1.9.3 The Total Energy–Momentum Matrix

The derived energy–momentum matrices include the conservation laws of energy and momentum of a *closed* system. For example, if a force density  $\mathbf{f}_D$  acts from the outside on the fluid, so that an electromagnetic field acts on the electrically charged liquid, then

$$\mathbf{f}_D = -\mathbf{T}_{b,e} \vec{\nabla} = \mathbf{T}_{\text{mech}} \vec{\nabla}, \quad (1.186)$$

or combined,

$$\mathbf{T}_{\text{total}} \stackrel{\text{def}}{=} (\mathbf{T}_{\text{mech}} + \mathbf{T}_{b,e}), \quad (1.187)$$

so

$$\boxed{\mathbf{T}_{\text{total}} \vec{\nabla} = \mathbf{f}_{\text{total}}}. \quad (1.188)$$

The conservation laws now apply to the whole system, which is fluid plus the electromagnetic field. Since the individual matrices are symmetric, the total energy–momentum matrix  $\mathbf{T}_{\text{total}}$  is symmetric. If there are other components in the considered system, they can be incorporated in the symmetric total energy–momentum matrix  $\mathbf{T}_{\text{total}}$  in a similar way as described above, and one obtains (1.188). This form of mathematical representation of the dynamic behaviour of physical systems will play a major role in the main equations of Einstein’s theory of General Relativity!

## 1.10 The Most Important Definitions and Formulas in Special Relativity

For inertial systems, i.e. reference systems that are uniformly moving against each other, the fundamental physical laws are linked with the Lorentz transformation and are invariant. In the Theory of Special Relativity, we derived (see (1.20))

$$\gamma \stackrel{\text{def}}{=} \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

and the Lorentz transformation matrix (see (1.38))

$$\mathbf{L}(v) \stackrel{\text{def}}{=} \left( \begin{array}{c|c} \gamma & -\frac{\gamma}{c} \mathbf{v}^\top \\ \hline -\frac{\gamma}{c} \mathbf{v} & \mathbf{I} + (\gamma - 1) \frac{\mathbf{v} \mathbf{v}^\top}{v^2} \end{array} \right)$$

with the following (see (1.39)) equations:

$$ct' = \gamma ct - \frac{\gamma}{c} \mathbf{v}^\top \mathbf{x}, \quad \mathbf{x}' = \mathbf{x} + (\gamma - 1) \frac{\mathbf{v}^\top \mathbf{x}}{v^2} \mathbf{v} - \gamma \mathbf{v} t.$$

In (1.65), we derived the relativistic velocity addition:

$$\mathbf{w} = \frac{\mathbf{v} + \mathbf{u} + \left(\frac{1}{\gamma_v} - 1\right)(\mathbf{u} - \frac{\mathbf{v}^\top \mathbf{u}}{v^2} \mathbf{v})}{1 + \frac{\mathbf{v}^\top \mathbf{u}}{c^2}}.$$

If the two velocities  $\mathbf{v}$  and  $\mathbf{u}$  are parallel, then (see (1.68))

$$\mathbf{w} = \frac{\mathbf{v} + \mathbf{u}}{1 + \frac{\mathbf{v}^\top \mathbf{u}}{c^2}}.$$

The modified velocity vector in (1.79)

$$\mathbf{u} \stackrel{\text{def}}{=} \frac{d\mathbf{x}}{dt} \in \mathbb{R}^3, \quad \gamma_u \stackrel{\text{def}}{=} \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}}, \quad \vec{\mathbf{u}} \stackrel{\text{def}}{=} \gamma_u \begin{pmatrix} c \\ \mathbf{u} \end{pmatrix} \in \mathbb{R}^4$$

is transformed by a Lorentz transformation  $\mathbf{L}$  into the velocity vector  $\vec{\mathbf{u}}'$  (see (1.80))

$$\vec{\mathbf{u}}' = \mathbf{L} \vec{\mathbf{u}}.$$

With (1.95),

$$\vec{\mathbf{a}} \stackrel{\text{def}}{=} \gamma_u \cdot \frac{d}{dt} \vec{\mathbf{u}} \in \mathbb{R}^4$$

is the Lorentz transformation of the acceleration vector  $\vec{a}$  (see (1.96)):

$$\vec{a}' = L\vec{a}.$$

Einstein's famous formula (1.111) reads

$$E_0 = m_0 c^2.$$

The invariance of the fundamental equation of mechanics ( $m_0$  is the rest mass) is documented in

$$m_0 \vec{a} = \vec{f} \quad L \Rightarrow \Leftarrow L^{-1} \quad m_0 \vec{a}' = \vec{f}'$$

and in electrodynamics

$$\begin{aligned} F_{B,e} \vec{\nabla} &= \frac{1}{c} \vec{j} \\ F_{E,b} \vec{\nabla} &= \mathbf{o} \end{aligned} \quad L \Rightarrow \Leftarrow L^{-1} \quad \begin{aligned} F'_{B',e'} \vec{\nabla}' &= \frac{1}{c} \vec{j}' \\ F'_{E',b'} \vec{\nabla}' &= \mathbf{o} \end{aligned}$$

and

$$\vec{f} = \frac{q}{c} \mathbf{M} F_{B,e} \vec{u} \quad L \Rightarrow \Leftarrow L^{-1} \quad \vec{f}' = \frac{q}{c} \mathbf{M} F'_{B',e'} \vec{u}'$$

with

$$F_{B,e} \stackrel{\text{def}}{=} \begin{pmatrix} 0 & \mathbf{e}^\top \\ -\mathbf{e} & \mathbf{B}_\times \end{pmatrix}, F_{E,b} \stackrel{\text{def}}{=} \begin{pmatrix} 0 & -\mathbf{b}^\top \\ \mathbf{b} & \mathbf{E}_\times \end{pmatrix} \text{ and } \vec{\nabla} \stackrel{\text{def}}{=} \begin{pmatrix} -\frac{1}{c} \frac{\partial}{\partial t} \\ \nabla \end{pmatrix}.$$

With the electromagnetic energy–momentum matrix

$$T_{b,e} \stackrel{\text{def}}{=} \begin{pmatrix} w & \mathbf{s}^\top \\ \mathbf{s} & (\mathbf{e}\mathbf{e}^\top + \mathbf{b}\mathbf{b}^\top - w\mathbf{I}_3) \end{pmatrix}$$

where

$$w \stackrel{\text{def}}{=} 1/2(e^2 + b^2),$$

one gets

$$T_{b,e} \cdot \vec{\nabla} = \frac{1}{\gamma_v} \vec{f},$$

and with the mechanical energy–momentum matrix

$$T_{\text{mech}} \stackrel{\text{def}}{=} \left( \rho_0 + \frac{p}{c^2} \right) \vec{u} \vec{u}^\top - p \mathbf{M},$$

we obtain the relativistic generalization of the hydrodynamic equations

$$\mathbf{T}_{\text{mech}} \vec{\nabla} = \vec{f}_D.$$

*Remark* The operator  $\vec{\nabla}$  was somewhat unusual in the above formulas—it was written to the right to the approaching object, e.g. in  $\mathbf{T}_{\text{mech}} \vec{\nabla}$ . This was done so that *column* vectors appear on both sides of the equation. The usual order would be if one transposes the eligible equations. Then there would be *row* vectors on the left- and on the right-hand side, and the operators, though provided with a transpose sign, would have their usual place as, e.g. in

$$\vec{\nabla}^T \mathbf{T}_{\text{mech}} = \vec{f}_D^T.$$

(Since the matrix  $\mathbf{T}_{\text{mech}}$  is symmetric, it need not be transposed.)

# Chapter 2

## Theory of General Relativity

The chapter begins with the introduction of the metric matrix  $G$  and the effect of a homogeneous field of gravitation on a mass particle. Then the motion on geodesic lines in a gravitational field is considered. The general transformation of coordinates leads to the Christoffel matrix and the Riemannian curvature matrix. With the help of the Ricci matrix, the Theory of General Relativity of Einstein can then be formulated.

### 2.1 General Relativity and Riemannian Geometry

In the theory of general relativity, the invariance of the equations with respect to any coordinate transformation is required. Especially the theorems of physics are to remain unchanged if one transforms them from one coordinate system in another coordinate system by the general transformation equations

$$x_i = x_i(x'_0, x'_1, x'_2, x'_3) \quad \text{for } i = 0, 1, 2 \text{ and } 3. \quad (2.1)$$

Consider an arbitrary coordinate system  $\mathcal{K}$  on the infinitely small neighbourhood of the point  $P$ , in which also a field of gravity may be present. In an infinitely small space and an infinitely small time interval, or in other words, in an infinitely small spacetime interval, the coordinate system  $\mathcal{K}$  can be replaced by a coordinate system  $\mathcal{K}'$  which is accelerated to it and in which no field of gravitation is present.  $\mathcal{K}'$  is the *local* spacetime coordinate system in the neighbourhood of a point, and  $\mathcal{K}$  the *general* coordinate system. It is now assumed that

**For all local coordinate systems  $\mathcal{K}'$  the Special Relativity is valid in any infinitely small four-dimensional neighbourhood.**

The point  $P'$  may be infinitely close to the point  $P$ , and have the real coordinates  $dx_0, dx_1, dx_2$  and  $dx_3$  in the Cartesian (rectangular) coordinate system. For a line element, one has

$$ds^2 = dx_0^2 - (dx_1^2 + dx_2^2 + dx_3^2) \quad (2.2)$$



where  $x_0$  is the time coordinate  $ct$ . If  $ds^2$  is positive, then  $P'$  is reached from  $P$  by a movement with a velocity smaller than the velocity of light. Equation (2.2) can be written using the matrix

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

as a quadratic form

$$ds^2 = d\vec{x}^\top \mathbf{M} d\vec{x}. \quad (2.3)$$

Going on to the coordinate system  $\mathcal{K}'$ , one gets with

$$dx_i = \frac{\partial x_i}{\partial x'_0} dx'_0 + \frac{\partial x_i}{\partial x'_1} dx'_1 + \frac{\partial x_i}{\partial x'_2} dx'_2 + \frac{\partial x_i}{\partial x'_3} dx'_3$$

and the Jacobi matrix

$$\mathbf{J} \stackrel{\text{def}}{=} \begin{pmatrix} \frac{\partial x_0}{\partial x'_0} & \cdots & \frac{\partial x_0}{\partial x'_3} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_3}{\partial x'_0} & \cdots & \frac{\partial x_3}{\partial x'_3} \end{pmatrix} = \frac{\partial \vec{x}}{\partial \vec{x}'^\top} \quad (2.4)$$

the connection

$$d\vec{x} = \mathbf{J} d\vec{x}'. \quad (2.5)$$

This is used in (2.3), yielding

$$ds^2 = d\vec{x}'^\top \mathbf{J}^\top \mathbf{M} \mathbf{J} d\vec{x}' = d\vec{x}'^\top \mathbf{G} d\vec{x}'. \quad (2.6)$$

So the *metric matrix*  $\mathbf{G}$  is defined as

$$\mathbf{G} \stackrel{\text{def}}{=} \mathbf{J}^\top \mathbf{M} \mathbf{J} \in \mathbb{R}^{4 \times 4}. \quad (2.7)$$

The matrix elements  $g_{ik}$  are functions of the parameters  $x'_j$ . They may change their values from point to point. In Special Relativity,  $\mathbf{G} = \mathbf{M}$  in any finite region. An unforced particle moves in such a region straightforwardly and uniformly.

But if the particle is in a gravitational field, the motion is curvilinear and nonuniform. Depending on the nature of the gravitational field, the  $g_{ik}$  are functions of other parameters. At most ten different elements  $g_{ik}$  of the symmetric  $4 \times 4$ -matrix  $\mathbf{G}$  describe the field of gravitation at every point in the coordinate system. In every *local* coordinate system the  $g_{ik}$  are constant and can be transformed by a similarity transformation into the form  $\mathbf{M}$ . Because if  $\mathbf{G} = \mathbf{J}^\top \mathbf{M} \mathbf{J}$ , then one immediately obtains with the transformation  $d\vec{x}' = \mathbf{J}^{-1} d\vec{\xi}$

$$d\vec{x}'^\top \mathbf{G} d\vec{x}' = d\vec{\xi}^\top \mathbf{J}^{-1 \top} \mathbf{J}^\top \mathbf{M} \mathbf{J} \mathbf{J}^{-1} d\vec{\xi} = d\vec{\xi}^\top \mathbf{M} d\vec{\xi}.$$

But this is true only for a point because the matrices  $\mathbf{J}$  are different from point to point. There is *no* transformation matrix  $\mathbf{J}$  which is valid globally.

## 2.2 Motion in a Gravitational Field

What influence does a gravitational field have on a mass particle? Following the relativity principle, in the local inertial system, which means in a coordinate system moving with the particle, the laws of special relativity are valid. For the movement of a particle with no external forces acting on it,

$$\frac{d^2 \vec{\xi}}{d\tau^2} = \mathbf{0}. \quad (2.8)$$

The proper time  $\tau$  follows from

$$ds^2 = c^2 d\tau^2 = d\vec{\xi}^T \mathbf{M} d\vec{\xi}. \quad (2.9)$$

By integrating (2.8), one gets with the initial position  $\vec{\xi}(0)$  and the initial velocity  $\dot{\vec{\xi}}(0)$

$$\vec{\xi}(\tau) = \vec{\xi}(0) + \dot{\vec{\xi}}(0)\tau.$$

A photon also moves in a straight line in the local inertial system. But  $\tau$  is then not the proper time of the photon. A photon has no proper time because for light  $ds = 0 = c d\tau$ . Therefore, we introduce the parameter  $\lambda$  such that

$$\frac{d^2 \vec{\xi}}{d\lambda^2} = \mathbf{0}$$

is the equation of motion of the photon in the local inertial system. Now we move from the local inertial system with the spacetime vector  $\vec{\xi}$  to the global inertial system with the spacetime vector  $\vec{x}$ . With

$$\mathbf{J} \stackrel{\text{def}}{=} \frac{\partial \vec{\xi}}{\partial \vec{x}^T}$$

one gets for (2.9)

$$ds^2 = c^2 d\tau^2 = d\vec{x}^T \mathbf{J}^T \mathbf{M} \mathbf{J} d\vec{x} = d\vec{x}^T \mathbf{G} d\vec{x}. \quad (2.10)$$

For light this becomes

$$d\vec{x}^T \mathbf{G} d\vec{x} = 0. \quad (2.11)$$

### 2.2.1 First Solution

From (2.8) one obtains for the motion of a particle with

$$\dot{\vec{x}} \stackrel{\text{def}}{=} \frac{d\vec{x}}{d\tau},$$

$$\frac{d}{d\tau} \left( \frac{\partial \vec{\xi}}{\partial \vec{x}^\top} \frac{d\vec{x}}{d\tau} \right) = \frac{d}{d\tau} (J \dot{\vec{x}}) = \frac{d}{d\tau} (J) \dot{\vec{x}} + J \ddot{\vec{x}} = \mathbf{0}, \quad (2.12)$$

and with (A.90)

$$\frac{d}{d\tau} (J(\vec{x}(\tau))) = (\dot{\vec{x}}^\top \otimes I_4) \frac{\partial J}{\partial \vec{x}}$$

from (2.12)

$$\ddot{\vec{x}} = -J^{-1} (\dot{\vec{x}}^\top \otimes I_4) \frac{\partial J}{\partial \vec{x}} \dot{\vec{x}}, \quad (2.13)$$

or

$$\ddot{\vec{x}} = -(\dot{\vec{x}}^\top \otimes J^{-1}) \frac{\partial J}{\partial \vec{x}} \dot{\vec{x}} = -(\dot{\vec{x}}^\top \otimes I_4) (I_4 \otimes J^{-1}) \frac{\partial J}{\partial \vec{x}} \dot{\vec{x}},$$

that is,

$$\ddot{\vec{x}} = -(I_4 \otimes \dot{\vec{x}}^\top) U_{4 \times 4} (I_4 \otimes J^{-1}) \frac{\partial J}{\partial \vec{x}} \dot{\vec{x}}. \quad (2.14)$$

With  $J_k \stackrel{\text{def}}{=} \frac{\partial J}{\partial x_k} \in \mathbb{R}^{4 \times 4}$  and

$$\hat{\Gamma} = \begin{pmatrix} \hat{\Gamma}_0 \\ \vdots \\ \hat{\Gamma}_3 \end{pmatrix} \stackrel{\text{def}}{=} U_{4 \times 4} (I_4 \otimes J^{-1}) \frac{\partial J}{\partial \vec{x}} = U_{4 \times 4} \begin{pmatrix} J^{-1} J_0 \\ \vdots \\ J^{-1} J_3 \end{pmatrix} \in \mathbb{R}^{16 \times 4} \quad (2.15)$$

one can write in place of (2.14) the compact equation

$$\ddot{\vec{x}} = -(I_4 \otimes \dot{\vec{x}}^\top) \hat{\Gamma} \dot{\vec{x}}. \quad (2.16)$$

With (2.15) and (2.16) one gets for the individual vector components  $\ddot{x}_k$  according to the form of the permutation matrix  $U_{4 \times 4}$  in the appendix ( $j_k^{-\top} \in \mathbb{R}^4$  is row  $k$  of  $J^{-1}$ )

$$\ddot{x}_k = -\dot{\vec{x}}^\top (I_4 \otimes j_k^{-\top}) \frac{\partial J}{\partial \vec{x}} \dot{\vec{x}}. \quad (2.17)$$

From (2.15) and (2.17) we can read directly

$$\hat{\Gamma}_k = (I_4 \otimes j_k^{-\top}) \frac{\partial J}{\partial \vec{x}}. \quad (2.18)$$

The so-called Christoffel matrices  $\hat{\Gamma}$  can be calculated directly from the Jacobi matrix  $\mathbf{J}$ , that is, from the transformation matrix for the transition from the local inertial system to the accelerated non-inertial system (the coordinate system with a gravitational field). For the motion of a photon, we obtain in the same manner

$$\frac{d^2 \vec{x}}{d\lambda^2} = - \left( \mathbf{I}_4 \otimes \frac{d\vec{x}^\top}{d\lambda} \right) \hat{\Gamma} \frac{d\vec{x}}{d\lambda}. \quad (2.19)$$

### 2.2.2 Second Solution

An alternative solution is obtained from the second form in (A.90), namely

$$\frac{d}{d\tau}(\mathbf{J}) = \frac{\partial \mathbf{J}}{\partial \vec{x}^\top} (\dot{\vec{x}} \otimes \mathbf{I}_4). \quad (2.20)$$

Thus we obtain from (2.12)

$$\ddot{\vec{x}} = -\mathbf{J}^{-1} \frac{\partial \mathbf{J}}{\partial \vec{x}^\top} (\dot{\vec{x}} \otimes \mathbf{I}_4) \dot{\vec{x}} = -\mathbf{J}^{-1} \frac{\partial \mathbf{J}}{\partial \vec{x}^\top} (\dot{\vec{x}} \otimes \dot{\vec{x}}), \quad (2.21)$$

that is, with

$$\tilde{\Gamma} \stackrel{\text{def}}{=} \mathbf{J}^{-1} \frac{\partial \mathbf{J}}{\partial \vec{x}^\top} \quad (2.22)$$

written completely as

$$\tilde{\Gamma} = \mathbf{J}^{-1} \left[ \frac{\partial \mathbf{J}}{\partial x_0} \mid \frac{\partial \mathbf{J}}{\partial x_1} \mid \frac{\partial \mathbf{J}}{\partial x_2} \mid \frac{\partial \mathbf{J}}{\partial x_3} \right] \in \mathbb{R}^{4 \times 16}, \quad (2.23)$$

so with  $\dot{\vec{x}} \otimes \dot{\vec{x}} \in \mathbb{R}^{16}$ ,

$$\ddot{\vec{x}} = -\tilde{\Gamma} (\dot{\vec{x}} \otimes \dot{\vec{x}}). \quad (2.24)$$

This is an alternative representation of the relation (2.16)!

If one defines

$$\tilde{\gamma}_k^\top \stackrel{\text{def}}{=} \mathbf{j}_k^{-\top} \frac{\partial \mathbf{J}}{\partial \vec{x}^\top} \in \mathbb{R}^{16}, \quad (2.25)$$

then for each vector component the following scalar vector product is obtained:

$$\ddot{x}_k = -\tilde{\gamma}_k^\top \cdot (\dot{\vec{x}} \otimes \dot{\vec{x}}). \quad (2.26)$$

### 2.2.3 Relation Between $\tilde{\Gamma}$ and $G$

Since  $G = J^T M J$  and  $\tilde{\Gamma} = J^{-1} \frac{\partial J}{\partial x^T}$ , the matrix  $\tilde{\Gamma}$  must depend on  $\frac{\partial G}{\partial x}$ . This indeed is the case.

On the one hand,

$$G \tilde{\Gamma} = J^T M \frac{\partial J}{\partial x^T}, \quad (2.27)$$

and, on the other hand,  $g_{\mu\nu} = j_\mu^T M j_\nu$ , so

$$\frac{\partial g_{\mu\nu}}{\partial x_\lambda} = \frac{\partial j_\mu^T}{\partial x_\lambda} M j_\nu + j_\mu^T M \frac{\partial j_\nu}{\partial x_\lambda}. \quad (2.28)$$

Furthermore,

$$\frac{\partial g_{\lambda\nu}}{\partial x_\mu} = \frac{\partial j_\lambda^T}{\partial x_\mu} M j_\nu + j_\lambda^T M \frac{\partial j_\nu}{\partial x_\mu} \quad (2.29)$$

and

$$\frac{\partial g_{\mu\lambda}}{\partial x_\nu} = \frac{\partial j_\mu^T}{\partial x_\nu} M j_\lambda + j_\mu^T M \frac{\partial j_\lambda}{\partial x_\nu}. \quad (2.30)$$

If we add (2.28) and (2.29) and then subtract (2.30), we obtain

$$\frac{\partial g_{\mu\nu}}{\partial x_\lambda} + \frac{\partial g_{\lambda\nu}}{\partial x_\mu} - \frac{\partial g_{\mu\lambda}}{\partial x_\nu} = 2 \frac{\partial j_\mu^T}{\partial x_\lambda} M j_\nu = 2 j_\nu^T M \frac{\partial j_\mu}{\partial x_\lambda}. \quad (2.31)$$

Calling  $G \tilde{\Gamma} \stackrel{\text{def}}{=} \check{\Gamma}$ , then we obtain with (2.27) and (2.31) as the element  $\check{\Gamma}_{\nu\mu}^{\check{\lambda}}$  in the  $\nu$ th row and the  $\mu$ th column of  $\check{\Gamma}_\lambda$

$$\check{\Gamma}_{\nu\mu}^{\check{\lambda}} = \frac{1}{2} \left( \frac{\partial g_{\mu\nu}}{\partial x_\lambda} + \frac{\partial g_{\lambda\nu}}{\partial x_\mu} - \frac{\partial g_{\mu\lambda}}{\partial x_\nu} \right). \quad (2.32)$$

Since  $\tilde{\Gamma} = G^{-1} G \tilde{\Gamma} = G^{-1} \check{\Gamma}$ ,

$$\tilde{\Gamma}_\lambda = G^{-1} \check{\Gamma}_\lambda = G^{-1} J^T M \frac{\partial J}{\partial x_\lambda},$$

and one finally obtains with the  $\alpha$ th row  $g_\alpha^{[-T]}$  of the matrix  $G^{-1}$  and the  $\nu$ th element  $g_{\alpha\nu}^{[-1]}$  of this row vector the following relation between the elements in the  $\alpha$ th row and the  $\mu$ th column of  $\tilde{\Gamma}_\lambda$  and the elements of  $G$

$$\tilde{\Gamma}_{\alpha\mu}^{\check{\lambda}} = g_\alpha^{[-T]} J^T M \frac{\partial j_\mu}{\partial x_\lambda} = \sum_{\nu=0}^3 \frac{g_{\alpha\nu}^{[-1]}}{2} \left( \frac{\partial g_{\mu\nu}}{\partial x_\lambda} + \frac{\partial g_{\lambda\nu}}{\partial x_\mu} - \frac{\partial g_{\mu\lambda}}{\partial x_\nu} \right). \quad (2.33)$$

This is the desired relationship between the matrix elements of  $\tilde{\Gamma}$  and  $G$ !

## 2.3 Geodesic Lines and Equations of Motion

The motion of photons and particles in a gravitational field will again be considered, but now using the calculus of variations. The same results as in (2.16) are expected. This is to be seen by using the same Christoffel matrix  $\mathbf{F}$  in the results. In special relativity, the motion of a photon is given by  $c^2 t^2 = \mathbf{x}^\top \mathbf{x}$ , that means  $s^2 = c^2 t^2 - \mathbf{x}^\top \mathbf{x} = 0$ , or  $ds^2 = dx_0^2 - d\mathbf{x}^\top d\mathbf{x} = d\bar{\mathbf{x}}^\top \mathbf{M} d\bar{\mathbf{x}} = 0$  for any small distance  $d\mathbf{x}$ . If paths were straight lines, that would mean the shortest possible connection between the points  $P_1$  and  $P_2$ . Also the theory of general relativity demands that light and particles move on straightest possible paths. These paths are the so-called *geodesic curves* for which the length has an extreme value:

$$\delta \int_{P_1}^{P_2} ds = 0. \quad (2.34)$$

This produces a system of four differential equations. For the variation of  $ds^2$  we obtain

$$\begin{aligned} \delta(ds^2) &= \delta(d\bar{\mathbf{x}}^\top \mathbf{G} d\bar{\mathbf{x}}), \\ 2(\delta ds) ds &= (\delta d\bar{\mathbf{x}}^\top) \mathbf{G} d\bar{\mathbf{x}} + d\bar{\mathbf{x}}^\top (\delta \mathbf{G}) d\bar{\mathbf{x}} + d\bar{\mathbf{x}}^\top \mathbf{G} (\delta d\bar{\mathbf{x}}), \end{aligned} \quad (2.35)$$

but since  $\mathbf{G}$  is symmetric,  $\mathbf{G} = \mathbf{G}^\top$ , we get

$$2(\delta ds) ds = 2 d\bar{\mathbf{x}}^\top \mathbf{G} (\delta d\bar{\mathbf{x}}) + d\bar{\mathbf{x}}^\top (\delta \mathbf{G}) d\bar{\mathbf{x}}. \quad (2.36)$$

Equation (2.36) divided by  $2 ds$  results, with  $d(\delta \bar{\mathbf{x}}) = \delta d\bar{\mathbf{x}}$  and  $\frac{d\bar{\mathbf{x}}}{ds} \stackrel{\text{def}}{=} \dot{\bar{\mathbf{x}}}$ , in

$$\delta ds = \dot{\bar{\mathbf{x}}}^\top \mathbf{G} d(\delta \bar{\mathbf{x}}) + \frac{1}{2} \dot{\bar{\mathbf{x}}}^\top (\delta \mathbf{G}) d\bar{\mathbf{x}}. \quad (2.37)$$

Multiplying the right-hand side by  $ds$  provides

$$\delta ds = \left[ \dot{\bar{\mathbf{x}}}^\top \mathbf{G} \frac{d(\delta \bar{\mathbf{x}})}{ds} + \frac{1}{2} \dot{\bar{\mathbf{x}}}^\top (\delta \mathbf{G}) \dot{\bar{\mathbf{x}}} \right] ds. \quad (2.38)$$

For the variation of the matrix  $\mathbf{G}$  we set

$$\delta \mathbf{G} = \frac{\partial \mathbf{G}}{\partial x_0} \delta x_0 + \frac{\partial \mathbf{G}}{\partial x_1} \delta x_1 + \frac{\partial \mathbf{G}}{\partial x_2} \delta x_2 + \frac{\partial \mathbf{G}}{\partial x_3} \delta x_3. \quad (2.39)$$

This is extended to a quadratic form:

$$\dot{\bar{\mathbf{x}}}^\top \delta \mathbf{G} \dot{\bar{\mathbf{x}}} = \dot{\bar{\mathbf{x}}}^\top \frac{\partial \mathbf{G}}{\partial x_0} \dot{\bar{\mathbf{x}}} \delta x_0 + \dot{\bar{\mathbf{x}}}^\top \frac{\partial \mathbf{G}}{\partial x_1} \dot{\bar{\mathbf{x}}} \delta x_1 + \dot{\bar{\mathbf{x}}}^\top \frac{\partial \mathbf{G}}{\partial x_2} \dot{\bar{\mathbf{x}}} \delta x_2 + \dot{\bar{\mathbf{x}}}^\top \frac{\partial \mathbf{G}}{\partial x_3} \dot{\bar{\mathbf{x}}} \delta x_3 \quad (2.40)$$

and the right-hand side is collected to a vector product, leading to the result

$$\dot{\bar{\mathbf{x}}}^\top \delta \mathbf{G} \dot{\bar{\mathbf{x}}} = \left[ \dot{\bar{\mathbf{x}}}^\top \frac{\partial \mathbf{G}}{\partial x_0} \dot{\bar{\mathbf{x}}}, \dot{\bar{\mathbf{x}}}^\top \frac{\partial \mathbf{G}}{\partial x_1} \dot{\bar{\mathbf{x}}}, \dot{\bar{\mathbf{x}}}^\top \frac{\partial \mathbf{G}}{\partial x_2} \dot{\bar{\mathbf{x}}}, \dot{\bar{\mathbf{x}}}^\top \frac{\partial \mathbf{G}}{\partial x_3} \dot{\bar{\mathbf{x}}} \right] \delta \bar{\mathbf{x}}. \quad (2.41)$$

The row vector on the right-hand side with

$$\left( \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}^\top} \right) \stackrel{\text{def}}{=} \left[ \frac{\partial \mathbf{G}}{\partial x_0}, \frac{\partial \mathbf{G}}{\partial x_1}, \frac{\partial \mathbf{G}}{\partial x_2}, \frac{\partial \mathbf{G}}{\partial x_3} \right]$$

and the Kronecker product can be written as

$$\left[ \dot{\mathbf{x}}^\top \frac{\partial \mathbf{G}}{\partial x_0} \dot{\mathbf{x}}, \dot{\mathbf{x}}^\top \frac{\partial \mathbf{G}}{\partial x_1} \dot{\mathbf{x}}, \dot{\mathbf{x}}^\top \frac{\partial \mathbf{G}}{\partial x_2} \dot{\mathbf{x}}, \dot{\mathbf{x}}^\top \frac{\partial \mathbf{G}}{\partial x_3} \dot{\mathbf{x}} \right] = \dot{\mathbf{x}}^\top \left( \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}^\top} \right) (\mathbf{I}_4 \otimes \dot{\mathbf{x}}). \quad (2.42)$$

With this relationship we obtain now for the varied integral (2.34)

$$\delta \int_{P_1}^{P_2} ds = \int_{P_1}^{P_2} \left[ \dot{\mathbf{x}}^\top \mathbf{G} \frac{d(\delta \dot{\mathbf{x}})}{ds} + \frac{1}{2} \dot{\mathbf{x}}^\top \left( \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}^\top} \right) (\mathbf{I}_4 \otimes \dot{\mathbf{x}}) \delta \dot{\mathbf{x}} \right] ds. \quad (2.43)$$

Taking in consideration  $\delta \dot{\mathbf{x}}(P_1) = \delta \dot{\mathbf{x}}(P_2) = \mathbf{o}$ , and performing integration by parts to the left summand in the integral, gives

$$\begin{aligned} \delta \int_{P_1}^{P_2} ds &= \int_{P_1}^{P_2} \left[ -\frac{d}{ds} (\dot{\mathbf{x}}^\top \mathbf{G}) \delta \dot{\mathbf{x}} + \frac{1}{2} \dot{\mathbf{x}}^\top \left( \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}^\top} \right) (\mathbf{I}_4 \otimes \dot{\mathbf{x}}) \delta \dot{\mathbf{x}} \right] ds \\ &= \int_{P_1}^{P_2} \left[ -\frac{d}{ds} (\dot{\mathbf{x}}^\top \mathbf{G}) + \frac{1}{2} \dot{\mathbf{x}}^\top \left( \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}^\top} \right) (\mathbf{I}_4 \otimes \dot{\mathbf{x}}) \right] \delta \dot{\mathbf{x}} ds = 0. \end{aligned} \quad (2.44)$$

To make the variation of the integral for every arbitrary vector function  $\delta \dot{\mathbf{x}}(\cdot)$  disappear, in spite of the fundamental theorem of the calculus of variations, the vector function in the brackets must be identically zero:

$$-\frac{d}{ds} (\dot{\mathbf{x}}^\top \mathbf{G}) + \frac{1}{2} \dot{\mathbf{x}}^\top \left( \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}^\top} \right) (\mathbf{I}_4 \otimes \dot{\mathbf{x}}) = \mathbf{o}^\top, \quad (2.45)$$

or transposed, where  $(\mathbf{A} \otimes \mathbf{B})^\top = (\mathbf{A}^\top \otimes \mathbf{B}^\top)$  is used,

$$\frac{1}{2} (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}} \dot{\mathbf{x}} - \frac{d}{ds} (\mathbf{G} \dot{\mathbf{x}}) = \mathbf{o}. \quad (2.46)$$

For the second term on the left-hand side one gets

$$\frac{d}{ds} (\mathbf{G} \dot{\mathbf{x}}) = \mathbf{G} \ddot{\mathbf{x}} + \frac{d}{ds} (\mathbf{G}) \dot{\mathbf{x}}. \quad (2.47)$$

And this is

$$\frac{d}{ds} (\mathbf{G}) = (\dot{\mathbf{x}}^\top \otimes \mathbf{I}_4) \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}}.$$

Inserted into (2.47) this finally yields

$$\mathbf{G} \ddot{\mathbf{x}} = \frac{1}{2} (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}} \dot{\mathbf{x}} - (\dot{\mathbf{x}}^\top \otimes \mathbf{I}_4) \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}} \dot{\mathbf{x}} \quad (2.48)$$

or, putting  $\ddot{\mathbf{x}}$  to the left-hand side,

$$\ddot{\mathbf{x}} = \mathbf{G}^{-1} \left[ \frac{1}{2} (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) - (\dot{\mathbf{x}}^\top \otimes \mathbf{I}_4) \right] \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}} \dot{\mathbf{x}}. \quad (2.49)$$

Taking into account the lemma in the Appendix “Vectors and Matrices”:

$$\mathbf{B} \otimes \mathbf{A} = \mathbf{U}_{s \times p} (\mathbf{A} \otimes \mathbf{B}) \mathbf{U}_{q \times t}, \quad \mathbf{A} \in \mathbb{R}^{p \times q}, \mathbf{B} \in \mathbb{R}^{s \times t}, \quad (2.50)$$

(2.49) can be converted, with the help of  $\dot{\mathbf{x}}^\top \otimes \mathbf{I}_4 = (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \mathbf{U}_{4 \times 4}$ , into

$$\ddot{\mathbf{x}} = \mathbf{G}^{-1} (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \left[ \frac{1}{2} \mathbf{I}_{16} - \mathbf{U}_{4 \times 4} \right] \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}} \dot{\mathbf{x}}. \quad (2.51)$$

With

$$\mathbf{G}^{-1} (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) = (\mathbf{G}^{-1} \otimes \mathbf{1}) (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) = (\mathbf{G}^{-1} \otimes \dot{\mathbf{x}}^\top) = (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) (\mathbf{G}^{-1} \otimes \mathbf{I}_4)$$

we finally obtain a form in which  $\dot{\mathbf{x}}$  is pulled out to the left and to the right:

$$\boxed{\ddot{\mathbf{x}} = (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) (\mathbf{G}^{-1} \otimes \mathbf{I}_4) \left[ \frac{1}{2} \mathbf{I}_{16} - \mathbf{U}_{4 \times 4} \right] \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}} \dot{\mathbf{x}}.} \quad (2.52)$$

Summarizing

$$\hat{\mathbf{T}} \stackrel{\text{def}}{=} (\mathbf{G}^{-1} \otimes \mathbf{I}_4) \left[ \mathbf{U}_{4 \times 4} - \frac{1}{2} \mathbf{I}_{16} \right] \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}} \quad (2.53)$$

$$= \mathbf{U}_{4 \times 4} (\mathbf{I}_4 \otimes \mathbf{G}^{-1}) \left[ \frac{1}{2} \mathbf{I}_{16} - \mathbf{U}_{4 \times 4} \right] \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}}, \quad (2.54)$$

one gets the compact equation

$$\boxed{\ddot{\mathbf{x}} = -(\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \hat{\mathbf{T}} \dot{\mathbf{x}}} \quad (2.55)$$

which agrees with the equation of motion (2.16), that is, in the language of the calculus of variations, this equation yields an extremal.

In (2.53) with the  $k$ th row  $\mathbf{g}_k^\top$  of the matrix  $\mathbf{G}$  is

$$\mathbf{U}_{4 \times 4} \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}} = \begin{pmatrix} \frac{\partial \mathbf{g}_0^\top}{\partial \dot{\mathbf{x}}} \\ \vdots \\ \frac{\partial \mathbf{g}_3^\top}{\partial \dot{\mathbf{x}}} \end{pmatrix}.$$



For the four components  $\ddot{x}_k$ ,  $k = 0, 1, 2$  and  $3$ , one obtains with the  $k$ th row  $\mathbf{g}_k^{-T}$  of the matrix  $\mathbf{G}^{-1}$ ,

$$\ddot{x}_k = \dot{\mathbf{x}}^\top (\mathbf{g}_k^{-T} \otimes \mathbf{I}_4) \left[ \begin{pmatrix} \frac{\partial \mathbf{g}_0^\top}{\partial \vec{x}} \\ \vdots \\ \frac{\partial \mathbf{g}_3^\top}{\partial \vec{x}} \end{pmatrix} - \frac{1}{2} \frac{\partial \mathbf{G}}{\partial \vec{x}} \right] \dot{\mathbf{x}}. \quad (2.56)$$

With

$$\hat{\Gamma}_k \stackrel{\text{def}}{=} (\mathbf{g}_k^{-T} \otimes \mathbf{I}_4) \left[ \begin{pmatrix} \frac{\partial \mathbf{g}_0^\top}{\partial \vec{x}} \\ \vdots \\ \frac{\partial \mathbf{g}_3^\top}{\partial \vec{x}} \end{pmatrix} - \frac{1}{2} \frac{\partial \mathbf{G}}{\partial \vec{x}} \right] \quad (2.57)$$

one can also write for (2.56)

$$\ddot{x}_k = -\dot{\mathbf{x}}^\top \hat{\Gamma}_k \dot{\mathbf{x}}. \quad (2.58)$$

The so-obtained matrix  $\hat{\Gamma}_k$  need not be symmetric. But the value of the quadratic form (2.58) is unchanged, if the matrix  $\hat{\Gamma}_k$  in (2.58) is replaced by the symmetric  $4 \times 4$ -matrix

$$\Gamma_k \stackrel{\text{def}}{=} \frac{1}{2} (\hat{\Gamma}_k + \hat{\Gamma}_k^\top). \quad (2.59)$$

In this way, the matrix is said to be *symmetrised*. Expanding (2.57) yields

$$\hat{\Gamma}_k = \sum_{i=0}^3 g_{k,i}^{[-1]} \left( \frac{\partial \mathbf{g}_i^\top}{\partial \vec{x}} - \frac{1}{2} \frac{\partial \mathbf{G}}{\partial x_i} \right),$$

and transposed

$$\hat{\Gamma}_k^\top = \sum_{i=0}^3 g_{k,i}^{[-1]} \left( \frac{\partial \mathbf{g}_i}{\partial \vec{x}^\top} - \frac{1}{2} \frac{\partial \mathbf{G}}{\partial x_i} \right).$$

For this one can also write

$$\hat{\Gamma}_k^\top = (\mathbf{g}_k^{-T} \otimes \mathbf{I}_4) \left[ \begin{pmatrix} \frac{\partial \mathbf{g}_0}{\partial \vec{x}^\top} \\ \vdots \\ \frac{\partial \mathbf{g}_3}{\partial \vec{x}^\top} \end{pmatrix} - \frac{1}{2} \frac{\partial \mathbf{G}}{\partial \vec{x}} \right]. \quad (2.60)$$

Inserting (2.60) into (2.59), we obtain

$$\mathbf{\Gamma}_k = \frac{1}{2}(\hat{\mathbf{\Gamma}}_k + \hat{\mathbf{\Gamma}}_k^\top) = \frac{1}{2}(\mathbf{g}_k^{-T} \otimes \mathbf{I}_4) \left[ \begin{pmatrix} \frac{\partial \mathbf{g}_0^\top}{\partial \dot{\mathbf{x}}} \\ \vdots \\ \frac{\partial \mathbf{g}_3^\top}{\partial \dot{\mathbf{x}}} \end{pmatrix} + \begin{pmatrix} \frac{\partial \mathbf{g}_0}{\partial \dot{\mathbf{x}}^\top} \\ \vdots \\ \frac{\partial \mathbf{g}_3}{\partial \dot{\mathbf{x}}^\top} \end{pmatrix} - \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}} \right]. \quad (2.61)$$

Multiplying out yields the results for the components of the Christoffel matrix  $\mathbf{\Gamma}_k$ , namely the above already derived relationship

$$\Gamma_{\alpha\beta}^k = \sum_{i=0}^3 \frac{g_{ki}^{[-1]}}{2} \left( \frac{\partial g_{\beta i}}{\partial x_\alpha} + \frac{\partial g_{\alpha i}}{\partial x_\beta} - \frac{\partial g_{\alpha\beta}}{\partial x_i} \right). \quad (2.62)$$

This can with

$$\check{\Gamma}_{\alpha\beta}^i \stackrel{\text{def}}{=} \frac{1}{2} \left( \frac{\partial g_{\beta i}}{\partial x_\alpha} + \frac{\partial g_{\alpha i}}{\partial x_\beta} - \frac{\partial g_{\alpha\beta}}{\partial x_i} \right) \quad (2.63)$$

also be written as

$$\Gamma_{\alpha\beta}^k = \sum_{i=0}^3 g_{ki}^{[-1]} \check{\Gamma}_{\alpha\beta}^i. \quad (2.64)$$

In addition, from (2.63) the interesting connection follows:

$$\frac{\partial g_{\alpha i}}{\partial x_\beta} = \check{\Gamma}_{\alpha\beta}^i + \check{\Gamma}_{i\beta}^\alpha. \quad (2.65)$$

Assembling the four components  $\ddot{x}_k$  of the vector  $\ddot{\mathbf{x}}$  with the matrix  $\mathbf{\Gamma}_k$  into one vector yields

$$\ddot{\mathbf{x}} = - \begin{pmatrix} \dot{\mathbf{x}}^\top \mathbf{\Gamma}_0 \dot{\mathbf{x}} \\ \vdots \\ \dot{\mathbf{x}}^\top \mathbf{\Gamma}_3 \dot{\mathbf{x}} \end{pmatrix}, \quad (2.66)$$

respectively,

$$\ddot{\mathbf{x}} = -(\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \mathbf{\Gamma} \dot{\mathbf{x}}, \quad (2.67)$$

with

$$\mathbf{\Gamma} \stackrel{\text{def}}{=} \begin{pmatrix} \mathbf{\Gamma}_0 \\ \vdots \\ \mathbf{\Gamma}_3 \end{pmatrix} = \frac{1}{2}(\mathbf{G}^{-1} \otimes \mathbf{I}_4) \left[ \begin{pmatrix} \frac{\partial \mathbf{g}_0^\top}{\partial \dot{\mathbf{x}}} \\ \vdots \\ \frac{\partial \mathbf{g}_3^\top}{\partial \dot{\mathbf{x}}} \end{pmatrix} + \begin{pmatrix} \frac{\partial \mathbf{g}_0}{\partial \dot{\mathbf{x}}^\top} \\ \vdots \\ \frac{\partial \mathbf{g}_3}{\partial \dot{\mathbf{x}}^\top} \end{pmatrix} - \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}} \right]. \quad (2.68)$$

This can also be written as

$$\Gamma = \frac{1}{2}(\mathbf{G}^{-1} \otimes \mathbf{I}_4) \left[ (\mathbf{U}_{4 \times 4} - \mathbf{I}_{16}) \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}} + \begin{pmatrix} \frac{\partial g_0}{\partial \dot{\mathbf{x}}^\top} \\ \vdots \\ \frac{\partial g_3}{\partial \dot{\mathbf{x}}^\top} \end{pmatrix} \right]. \quad (2.69)$$

Introducing the matrix

$$\check{\Gamma} \stackrel{\text{def}}{=} \frac{1}{2} \left[ (\mathbf{U}_{4 \times 4} - \mathbf{I}_{16}) \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}} + \begin{pmatrix} \frac{\partial g_0}{\partial \dot{\mathbf{x}}^\top} \\ \vdots \\ \frac{\partial g_3}{\partial \dot{\mathbf{x}}^\top} \end{pmatrix} \right], \quad (2.70)$$

it can also be written as

$$\Gamma = (\mathbf{G}^{-1} \otimes \mathbf{I}_4) \check{\Gamma}. \quad (2.71)$$

The matrix difference  $\mathbf{U}_{4 \times 4} - \mathbf{I}_{16}$ , appearing in the matrix  $\check{\Gamma}$ , has the remarkable property that the first,  $(4 + 2)$ th,  $(8 + 3)$ th and the 16th row (resp., column) are equal to a zero-row (resp., zero-column)! For the matrix  $\check{\Gamma}$  this has the consequence that the corresponding rows consist of  $\frac{\partial g_{00}}{\partial \dot{\mathbf{x}}^\top}$ ,  $\frac{\partial g_{11}}{\partial \dot{\mathbf{x}}^\top}$ ,  $\frac{\partial g_{22}}{\partial \dot{\mathbf{x}}^\top}$  and  $\frac{\partial g_{33}}{\partial \dot{\mathbf{x}}^\top}$ . Furthermore, from (2.71) it follows that

$$\check{\Gamma} = (\mathbf{G} \otimes \mathbf{I}_4) \Gamma, \quad (2.72)$$

that is,

$$\check{\Gamma}_k = (\mathbf{g}_k^\top \otimes \mathbf{I}_4) \Gamma = g_{k0} \Gamma_0 + \cdots + g_{k3} \Gamma_3,$$

therefore,

$$\check{\Gamma}_{\alpha\beta}^k = \sum_{i=0}^3 g_{ki} \Gamma_{\alpha\beta}^i. \quad (2.73)$$

### 2.3.1 Alternative Geodesic Equation of Motion

Again the equations of motion can be modified as follows: Firstly,

$$(\dot{\mathbf{x}}^\top \otimes \mathbf{I}_4) \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}} = \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}^\top} (\dot{\mathbf{x}} \otimes \mathbf{I}_4), \quad (2.74)$$

and secondly,

$$(\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}} \dot{\mathbf{x}} = \begin{pmatrix} \dot{\mathbf{x}}^\top \mathbf{G}_0 \dot{\mathbf{x}} \\ \vdots \\ \dot{\mathbf{x}}^\top \mathbf{G}_3 \dot{\mathbf{x}} \end{pmatrix}.$$

Using the  $\mathit{vec}$ -operator from the Appendix (A.51) to the scalar component  $\dot{\mathbf{x}}^\top \mathbf{G}_k \dot{\mathbf{x}}$  yields

$$\mathit{vec}(\dot{\mathbf{x}}^\top \mathbf{G}_k \dot{\mathbf{x}}) = (\dot{\mathbf{x}}^\top \otimes \dot{\mathbf{x}}^\top) \mathit{vec}(\mathbf{G}_k) = (\mathit{vec}(\mathbf{G}_k))^\top (\dot{\mathbf{x}} \otimes \dot{\mathbf{x}}),$$

so

$$\underline{\underline{(\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}} \dot{\mathbf{x}} = \frac{\overline{\partial \mathbf{G}}}{\partial \dot{\mathbf{x}}^\top} (\dot{\mathbf{x}} \otimes \dot{\mathbf{x}})}}, \quad (2.75)$$

with

$$\frac{\overline{\partial \mathbf{G}}}{\partial \dot{\mathbf{x}}^\top} \stackrel{\text{def}}{=} \begin{pmatrix} (\mathit{vec}(\mathbf{G}_0))^\top \\ \vdots \\ (\mathit{vec}(\mathbf{G}_3))^\top \end{pmatrix} = \begin{pmatrix} \mathbf{g}_{0,0}^\top & & \mathbf{g}_{0,3}^\top \\ \mathbf{g}_{1,0}^\top & \cdots & \mathbf{g}_{1,3}^\top \\ \mathbf{g}_{2,0}^\top & & \mathbf{g}_{2,3}^\top \\ \mathbf{g}_{3,0}^\top & & \mathbf{g}_{3,3}^\top \end{pmatrix} \in \mathbb{R}^{4 \times 16}, \quad (2.76)$$

where  $\mathbf{g}_{i,j}^\top$  is the  $j$ th row of  $\mathbf{G}_i$ .

By the “method of careful examination”, one can write

$$\frac{\overline{\partial \mathbf{G}}}{\partial \dot{\mathbf{x}}^\top} = \begin{pmatrix} \mathbf{g}_{0,0}^\top \\ \mathbf{g}_{1,0}^\top \\ \mathbf{g}_{2,0}^\top \\ \mathbf{g}_{3,0}^\top \\ \vdots \\ \mathbf{g}_{0,3}^\top \\ \mathbf{g}_{1,3}^\top \\ \mathbf{g}_{2,3}^\top \\ \mathbf{g}_{3,3}^\top \end{pmatrix}^B = \left( \mathbf{U}_{4 \times 4} \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}} \right)^B, \quad (2.77)$$

where the superscript “ $B$ ” means the *block-transposition* of the corresponding matrix. The block-transposition of a block matrix is defined as

$$\mathbf{A}^B \stackrel{\text{def}}{=} \begin{pmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_n \end{pmatrix}^B = (\mathbf{A}_1 \quad \dots \quad \mathbf{A}_n).$$

Equation (2.75) used in (2.48) yields

$$\ddot{\mathbf{x}} = -\mathbf{G}^{-1} \left[ \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}^\top} - \frac{1}{2} \frac{\overline{\partial \mathbf{G}}}{\partial \dot{\mathbf{x}}^\top} \right] (\dot{\mathbf{x}} \otimes \dot{\mathbf{x}}). \quad (2.78)$$

With

$$\boxed{\tilde{\mathbf{T}} \stackrel{\text{def}}{=} \mathbf{G}^{-1} \left[ \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{x}}^\top} - \frac{1}{2} \frac{\overline{\partial \mathbf{G}}}{\partial \dot{\mathbf{x}}^\top} \right] \in \mathbb{R}^{4 \times 16},} \quad (2.79)$$

one finally obtains

$$\boxed{\ddot{\mathbf{x}} = -\tilde{\Gamma}(\dot{\mathbf{x}} \otimes \dot{\mathbf{x}})}. \quad (2.80)$$

One can also reach the form (2.80) as follows: It is  $\ddot{x}_k = \dot{\mathbf{x}}^\top \Gamma_k \dot{\mathbf{x}}$ . Applying to this the *vec*-operator (see Appendix), one obtains

$$-\ddot{x}_k = \mathit{vec}(\dot{\mathbf{x}}^\top \Gamma_k \dot{\mathbf{x}}) = (\dot{\mathbf{x}}^\top \otimes \dot{\mathbf{x}}^\top) \mathit{vec}(\Gamma_k) = (\mathit{vec}(\Gamma_k))^\top (\dot{\mathbf{x}} \otimes \dot{\mathbf{x}}).$$

With

$$\tilde{\Gamma} \stackrel{\text{def}}{=} \begin{pmatrix} (\mathit{vec}(\Gamma_0))^\top \\ \vdots \\ (\mathit{vec}(\Gamma_3))^\top \end{pmatrix} \quad (2.81)$$

(2.80) is obtained again. Once again, we can write

$$\tilde{\Gamma} = \left( U_{4 \times 4} \frac{\partial \Gamma}{\partial \vec{x}} \right)^B. \quad (2.82)$$

Just a word regarding the derivatives with respect to  $s$ . If  $s^2$  and  $ds^2$  are positive, then it is a so-called time-like event. Then

$$\begin{aligned} (ds)^2 &= c^2 dt^2 - d\mathbf{x}^\top d\mathbf{x} = c^2 dt^2 - \dot{\mathbf{x}}^\top \dot{\mathbf{x}} dt^2, \\ ds &= \sqrt{c^2 - v^2} dt, \end{aligned}$$

so

$$\gamma ds = c dt.$$

Substituting  $ds = c d\tau$ , one obtains

$$\gamma d\tau = dt.$$

A comparison with the results of special relativity theory provides  $d\tau = dt'$ , i.e. the time that elapses in the moving coordinate system  $\mathcal{X}'$ . One calls in this context  $\tau$  as *proper time*.

## 2.4 Example: Uniformly Rotating Systems

We will consider a fixed inertial frame  $\mathcal{X}$  with the coordinates  $t, x, y$  and  $z$  and a uniformly around the  $z$ -axis *rotating* coordinate system  $\mathcal{K}$  with the coordinates  $\tau, r, \varphi$  and  $z$ . Then the transformation equations are

$$\begin{aligned} t &= \tau, \\ x &= r \cos(\varphi + \omega t), \\ y &= r \sin(\varphi + \omega t), \\ z &= z. \end{aligned} \quad (2.83)$$

As Jacobi matrix  $\mathbf{J}$  one obtains

$$\mathbf{J} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -r\frac{\omega}{c}\sin(\varphi + \omega t) & \cos(\varphi + \omega t) & -r\sin(\varphi + \omega t) & 0 \\ r\frac{\omega}{c}\cos(\varphi + \omega t) & \sin(\varphi + \omega t) & r\cos(\varphi + \omega t) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.84)$$

and as the metric matrix

$$\mathbf{G} = \mathbf{J}^T \mathbf{M} \mathbf{J} = \begin{pmatrix} 1 - r^2\frac{\omega^2}{c^2} & 0 & -r^2\frac{\omega}{c} & 0 \\ 0 & -1 & 0 & 0 \\ -r^2\frac{\omega}{c} & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad (2.85)$$

and from this

$$\mathbf{G}^{-1} = \begin{pmatrix} 1 & 0 & -\frac{\omega}{c} & 0 \\ 0 & -1 & 0 & 0 \\ -\frac{\omega}{c} & 0 & \frac{\omega^2}{c^2} - \frac{1}{r^2} & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (2.86)$$

In this case,  $ds^2$  has the value

$$ds^2 = \left(1 - r^2\frac{\omega^2}{c^2}\right) d\tau^2 - dr^2 - r d\varphi^2 - 2r^2\frac{\omega}{c} d\varphi d\tau - dz^2. \quad (2.87)$$

If a clock is in the rotating system at the position  $(r, \theta, z)$  and one considers two temporally directly adjacent events with  $dr = d\varphi = dz = 0$ , then one obtains for the proper time  $ds$  in this case the relationship (with  $v = r\omega$ )

$$ds = d\tau \sqrt{1 - r^2\omega^2/c^2} = d\tau \sqrt{1 - v^2/c^2} = d\tau/\gamma.$$

This is the relationship known from the theory of special relativity! For the calculation of the acceleration, the derivatives of the metric matrix are needed. In this case,  $\mathbf{G}_0 = \mathbf{G}_2 = \mathbf{G}_3 = \mathbf{0}$ , but

$$\mathbf{G}_1 = \frac{\partial \mathbf{G}}{\partial r} = \begin{pmatrix} -2r\frac{\omega^2}{c^2} & 0 & -2r\frac{\omega}{c} & 0 \\ 0 & 0 & 0 & 0 \\ -2r\frac{\omega}{c} & 0 & -2r & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Since only the matrix  $\mathbf{G}_1 \neq \mathbf{0}$ , in this case (2.56) is simplified to

$$\ddot{\mathbf{x}}_k = \left[ \frac{1}{2} (g^{k1} \dot{\mathbf{x}}^\top) - (\dot{x}_1 \mathbf{g}_k^{-\top}) \right] \mathbf{G}_1 \dot{\mathbf{x}}.$$

In particular, with  $\dot{\vec{x}}^\top = [c|\dot{r}|\dot{\varphi}|\dot{z}]$  we obtain

$$\begin{aligned}\ddot{r} &= -\frac{1}{2}\dot{\vec{x}}^\top \mathbf{G}_1 \dot{\vec{x}} \\ &= -\frac{1}{2}[-2r\omega^2/c - \dot{\varphi}2r\omega/c | 0 | -2r\omega - 2r\dot{\varphi} | 0]\dot{\vec{x}} = r(\omega + \dot{\varphi})^2\end{aligned}\quad (2.88)$$

and

$$\ddot{\varphi} = -\dot{r}[-\omega/c | 0 | \omega^2/c^2 - 1/r^2 | 0]\mathbf{G}_1 \dot{\vec{x}} = -2\dot{r}\omega/r - 2\dot{r}\dot{\varphi}/r, \quad (2.89)$$

or

$$r\ddot{\varphi} = -2\dot{r}(\dot{\varphi} + \omega). \quad (2.90)$$

Equation (2.88) multiplied by the mass  $m$  represents the centrifugal force, and (2.90) multiplied by the mass  $m$  is the so-called Coriolis force! The accelerations occurring in this rotating system are determined by the elements  $g_{ij} = g_{ij}(\vec{x})$  of the coordinate-dependent metric matrix  $\mathbf{G}(\vec{x})$ . For a local reference system, one can always specify a coordinate transformation (with  $\mathbf{J}^{-1}(\vec{x})$ ), so that the transformed system is obviously an inertial frame. In general, for an accelerated or non-uniformly moving (e.g. rotating) system no *globally valid* transformation matrix  $\mathbf{J}$  can be specified. The given space is *curved*!

It should now be shown that the Christoffel matrices are the same for the rotating system by applying the formula (2.18). According to (2.84),

$$\mathbf{J} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -r\frac{\omega}{c}\sin(\varphi + \omega t) & \cos(\varphi + \omega t) & -r\sin(\varphi + \omega t) & 0 \\ r\frac{\omega}{c}\cos(\varphi + \omega t) & \sin(\varphi + \omega t) & r\cos(\varphi + \omega t) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (2.91)$$

i.e.

$$\mathbf{J}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\varphi + \omega t) & \sin(\varphi + \omega t) & 0 \\ -\frac{\omega}{c} & \frac{1}{r}\sin(\varphi + \omega t) & \frac{1}{r}\cos(\varphi + \omega t) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (2.92)$$

Equations (2.88) and (2.89) can be used for the Christoffel matrices. The non-zero matrix elements are:

$$\begin{aligned}\Gamma_{00}^1 &= -r\frac{\omega^2}{c^2}, & \Gamma_{02}^1 &= -r\frac{\omega}{c}, & \Gamma_{03}^1 &= -r\frac{\omega}{c}, & \Gamma_{22}^1 &= -1, \\ \Gamma_{01}^2 &= \frac{4\omega}{rc}, & \text{and} & & \Gamma_{12}^2 &= \frac{4}{r}.\end{aligned}$$

Accordingly, we obtain with (2.18), e.g.

$$\begin{aligned} \Gamma_{00}^1 &= [0 \mid \cos(\varphi + \omega t) \mid \sin(\varphi + \omega t) \mid 0] \begin{pmatrix} 0 \\ -r \frac{\omega^2}{c^2} \cos(\varphi + \omega t) \\ -r \frac{\omega^2}{c^2} \sin(\varphi + \omega t) \\ 0 \end{pmatrix} \\ &= -r \frac{\omega^2}{c^2}. \end{aligned}$$

## 2.5 General Coordinate Transformations

In the theory of general relativity, the invariance of the mathematical descriptions is demanded for the general laws of nature with respect to each other arbitrarily moving coordinate systems. Even more generally:

The invariance of the mathematical descriptions with respect to arbitrary coordinate transformations is required.

### 2.5.1 Absolute Derivatives

First, we must clarify how the derivatives eventually have to be modified so that the derived expressions are invariant under coordinate transformations.

Suppose a vector field  $\mathbf{a}(\lambda)$  is defined along a curve whose parametric representation is given by  $\vec{\mathbf{x}}(\lambda)$ . Going on to another coordinate system  $\mathcal{K}'$  with  $\mathbf{a}'$ , for the mathematical description of dynamic processes one is especially interested in how the derivative  $d\mathbf{a}/d\lambda$  is transformed into  $d\mathbf{a}'/d\lambda$ . Because  $\mathbf{T} = \mathbf{T}(\vec{\mathbf{x}}(\lambda))$ , it follows that

$$\frac{d\mathbf{a}'}{d\lambda} = \frac{d(\mathbf{T}\mathbf{a})}{d\lambda} = \mathbf{T} \frac{d\mathbf{a}}{d\lambda} + \left( \frac{d\vec{\mathbf{x}}^\top}{d\lambda} \otimes \mathbf{I}_4 \right) \frac{\partial \mathbf{T}}{\partial \vec{\mathbf{x}}} \mathbf{a}, \quad (2.93)$$

i.e.  $d\mathbf{a}/d\lambda$  is not transformed into  $d\mathbf{a}'/d\lambda$  as was  $\mathbf{a}$  by a simple multiplication with the transformation matrix  $\mathbf{T}$ . For this the reason, the definition of the derivative is

$$\frac{d\mathbf{a}}{d\lambda} = \lim_{\delta\lambda \rightarrow 0} \frac{\mathbf{a}(\lambda + \delta\lambda) - \mathbf{a}(\lambda)}{\delta\lambda},$$

where the difference of the vectors is formed at different places on the curve  $\gamma$ , to which transformation matrices  $\mathbf{T}(\lambda) \neq \mathbf{T}(\lambda + \delta\lambda)$  generally belong.

Thus in order to be always able to take the same transformation matrix, the difference of two vectors must be taken at the *same* place of the curve. It is true that

$$\delta\mathbf{a} \approx \frac{d\mathbf{a}}{d\lambda} \delta\lambda \quad (2.94)$$



and, if the shift of  $\mathbf{a}$  is taken along a geodesic line,

$$\frac{d\mathbf{a}}{d\lambda} + (\mathbf{I}_4 \otimes \mathbf{a}^\top) \boldsymbol{\Gamma} \frac{d\vec{\mathbf{x}}}{d\lambda} = \mathbf{0},$$

so

$$\frac{d\mathbf{a}}{d\lambda} = -(\mathbf{I}_4 \otimes \mathbf{a}^\top) \boldsymbol{\Gamma} \frac{d\vec{\mathbf{x}}}{d\lambda}. \quad (2.95)$$

Here  $\frac{d\vec{\mathbf{x}}}{d\lambda}$  is the tangent vector to the geodesic curve  $\gamma$ , and  $\vec{\mathbf{x}}(\lambda)$  is the parametric representation of  $\gamma$ . Multiplying (2.95) with  $\delta\lambda$ , we obtain

$$\boxed{\delta\mathbf{a} = -(\mathbf{I}_4 \otimes \mathbf{a}^\top) \boldsymbol{\Gamma} \delta\vec{\mathbf{x}}.} \quad (2.96)$$

Moving the vector  $\mathbf{a}(\lambda)$  from the position  $\vec{\mathbf{x}}(\lambda)$  parallel to the position  $\vec{\mathbf{x}}(\lambda + \delta\lambda)$ , the vector

$$\bar{\mathbf{a}} \stackrel{\text{def}}{=} \mathbf{a}(\lambda) + \delta\mathbf{a}$$

is obtained, or by (2.96),

$$\bar{\mathbf{a}} \approx \mathbf{a}(\lambda) - (\mathbf{I}_4 \otimes \mathbf{a}^\top) \boldsymbol{\Gamma} \delta\vec{\mathbf{x}}. \quad (2.97)$$

On the other hand,  $\mathbf{a}(\lambda + \delta\lambda) - \bar{\mathbf{a}}$  is a vector at the position  $\gamma(\lambda + \delta\lambda)$ , as is  $(\mathbf{a}(\lambda + \delta\lambda) - \bar{\mathbf{a}})/\delta\lambda$ . As  $\delta\lambda \rightarrow 0$  the quotient is always a vector at the same location that varies, however.

The limit of this ratio is called *absolute derivative*  $\frac{D\mathbf{a}}{d\lambda}$  of  $\mathbf{a}(\lambda)$  along the curve  $\gamma$ :

$$\lim_{\delta\lambda \rightarrow 0} \frac{\mathbf{a}(\lambda + \delta\lambda) - \bar{\mathbf{a}}}{\delta\lambda} \approx \frac{d\mathbf{a}}{d\lambda} + \lim_{\delta\lambda \rightarrow 0} (\mathbf{I}_4 \otimes \mathbf{a}^\top) \boldsymbol{\Gamma} \frac{\delta\vec{\mathbf{x}}}{\delta\lambda}$$

which is obtained with (2.97), and therefore one defines the absolute derivative as

$$\boxed{\frac{D\mathbf{a}}{d\lambda} \stackrel{\text{def}}{=} \frac{d\mathbf{a}}{d\lambda} + (\mathbf{I}_4 \otimes \mathbf{a}^\top) \boldsymbol{\Gamma} \frac{d\vec{\mathbf{x}}}{d\lambda} = \dot{\mathbf{a}} + (\mathbf{I}_4 \otimes \mathbf{a}^\top) \boldsymbol{\Gamma} \dot{\vec{\mathbf{x}}}.} \quad (2.98)$$

The differentiated vector can be decomposed into

$$\dot{\mathbf{a}} = \frac{\partial \mathbf{a}}{\partial \vec{\mathbf{x}}^\top} \dot{\vec{\mathbf{x}}}, \quad (2.99)$$

so that in (2.98) one can extract  $\dot{\vec{\mathbf{x}}}$  to the right:

$$\boxed{\frac{D\mathbf{a}}{d\lambda} = \left[ \frac{\partial \mathbf{a}}{\partial \vec{\mathbf{x}}^\top} + (\mathbf{I}_4 \otimes \mathbf{a}^\top) \boldsymbol{\Gamma} \right] \dot{\vec{\mathbf{x}}}.} \quad (2.100)$$

The expression appearing in brackets is called the *covariant derivative* of  $\mathbf{a}$  and written  $\mathbf{a}_{\parallel\bar{\mathbf{x}}^\top} (\in \mathbb{R}^{4 \times 4})$ :

$$\mathbf{a}_{\parallel\bar{\mathbf{x}}^\top} \stackrel{\text{def}}{=} \frac{\partial \mathbf{a}}{\partial \bar{\mathbf{x}}^\top} + (\mathbf{I}_4 \otimes \mathbf{a}^\top) \boldsymbol{\Gamma}. \quad (2.101)$$

This covariant derivative  $\mathbf{a}_{\parallel\bar{\mathbf{x}}^\top}$  becomes the normal partial derivative  $\frac{\partial \mathbf{a}}{\partial \bar{\mathbf{x}}^\top}$ , if  $\boldsymbol{\Gamma} = \mathbf{0}$ , that means, there is no gravitational field.

### 2.5.2 Transformation of the Christoffel Matrix $\tilde{\boldsymbol{\Gamma}}$

In (2.22), the Christoffel matrix is defined:

$$\tilde{\boldsymbol{\Gamma}} \stackrel{\text{def}}{=} \mathbf{J}^{-1} \frac{\partial \mathbf{J}}{\partial \bar{\mathbf{x}}^\top} = \frac{\partial \bar{\mathbf{x}}}{\partial \bar{\xi}^\top} \cdot \frac{\partial^2 \bar{\xi}}{\partial \bar{\mathbf{x}}^\top \partial \bar{\mathbf{x}}^\top}. \quad (2.102)$$

If one changes over from the coordinate system with  $\bar{\mathbf{x}}$  to the coordinate system with  $\bar{\mathbf{x}}'$ , one obtains with the transformation matrices

$$\mathbf{T} \stackrel{\text{def}}{=} \frac{\partial \bar{\mathbf{x}}'}{\partial \bar{\mathbf{x}}^\top} \quad (2.103)$$

and

$$\bar{\mathbf{T}} \stackrel{\text{def}}{=} \frac{\partial \bar{\mathbf{x}}}{\partial \bar{\mathbf{x}}'^\top}, \quad (2.104)$$

the Christoffel matrix  $\tilde{\boldsymbol{\Gamma}}'$  in the coordinate system with  $\bar{\mathbf{x}}'$

$$\begin{aligned} \tilde{\boldsymbol{\Gamma}}' &\stackrel{\text{def}}{=} \frac{\partial \bar{\mathbf{x}}'}{\partial \bar{\xi}'^\top} \cdot \frac{\partial^2 \bar{\xi}}{\partial \bar{\mathbf{x}}'^\top \partial \bar{\mathbf{x}}'^\top} = \frac{\partial \bar{\mathbf{x}}'}{\partial \bar{\mathbf{x}}^\top} \frac{\partial \bar{\mathbf{x}}}{\partial \bar{\xi}'^\top} \frac{\partial}{\partial \bar{\mathbf{x}}'^\top} \left( \frac{\partial \bar{\xi}}{\partial \bar{\mathbf{x}}'^\top} \right) \\ &= \mathbf{T} \cdot \underbrace{\frac{\partial \bar{\mathbf{x}}}{\partial \bar{\xi}'^\top}}_{\mathbf{J}^{-1}} \frac{\partial}{\partial \bar{\mathbf{x}}'^\top} \left( \underbrace{\frac{\partial \bar{\xi}}{\partial \bar{\mathbf{x}}^\top}}_{\mathbf{J}} \cdot \underbrace{\frac{\partial \bar{\mathbf{x}}}{\partial \bar{\mathbf{x}}'^\top}}_{\bar{\mathbf{T}}} \right). \end{aligned} \quad (2.105)$$

According to the product and chain rules, one obtains

$$\begin{aligned} \frac{\partial}{\partial \bar{\mathbf{x}}'^\top} (\mathbf{J} \cdot \bar{\mathbf{T}}) &= \frac{\partial \mathbf{J}}{\partial \bar{\mathbf{x}}'^\top} (\mathbf{I}_4 \otimes \bar{\mathbf{T}}) + \mathbf{J} \frac{\partial \bar{\mathbf{T}}}{\partial \bar{\mathbf{x}}'^\top} \\ &= \frac{\partial \mathbf{J}}{\partial \bar{\mathbf{x}}^\top} (\bar{\mathbf{T}} \otimes \mathbf{I}_4) (\mathbf{I}_4 \otimes \bar{\mathbf{T}}) + \mathbf{J} \frac{\partial \bar{\mathbf{T}}}{\partial \bar{\mathbf{x}}^\top} (\bar{\mathbf{T}} \otimes \mathbf{I}_4) \\ &= \frac{\partial \mathbf{J}}{\partial \bar{\mathbf{x}}^\top} (\bar{\mathbf{T}} \otimes \bar{\mathbf{T}}) + \mathbf{J} \frac{\partial \bar{\mathbf{T}}}{\partial \bar{\mathbf{x}}^\top} (\bar{\mathbf{T}} \otimes \mathbf{I}_4). \end{aligned} \quad (2.106)$$

Inserting (2.106) in (2.105) reveals

$$\tilde{\Gamma}' = T \tilde{\Gamma} (\bar{T} \otimes \bar{T}) + T \frac{\partial \tilde{T}}{\partial \bar{x}^\top} (\bar{T} \otimes I_4). \quad (2.107)$$

The second term on the right-hand side shows the coordinate dependence of the transformation matrix  $T$ .

Another important characteristic is obtained as follows. Differentiating  $I_4 = T \bar{T}$  with respect to  $\bar{x}^\top$  yields

$$O = \frac{\partial T}{\partial \bar{x}^\top} (I_4 \otimes \bar{T}) + T \frac{\partial \bar{T}}{\partial \bar{x}^\top},$$

i.e.

$$T \frac{\partial \bar{T}}{\partial \bar{x}^\top} = - \frac{\partial T}{\partial \bar{x}^\top} (I_4 \otimes \bar{T}). \quad (2.108)$$

Putting this in (2.107), one gets another form of the transformed Christoffel matrix, namely

$$\tilde{\Gamma}' = T \tilde{\Gamma} (\bar{T} \otimes \bar{T}) - \frac{\partial T}{\partial \bar{x}^\top} (\bar{T} \otimes \bar{T}). \quad (2.109)$$

Moreover,

$$\frac{d^2 \bar{x}'}{d\tau^2} = \frac{d}{d\tau} \left( \underbrace{\frac{\partial \bar{x}'}{\partial \bar{x}^\top}}_T \cdot \frac{d\bar{x}}{d\tau} \right) = T \frac{d^2 \bar{x}}{d\tau^2} + \frac{\partial T}{\partial \bar{x}^\top} \underbrace{\left( \frac{d\bar{x}}{d\tau} \otimes I_4 \right)}_{\left( \frac{d\bar{x}}{d\tau} \otimes \frac{d\bar{x}}{d\tau} \right)} \frac{d\bar{x}}{d\tau}. \quad (2.110)$$

Multiplying (2.109) by the vector  $\left( \frac{d\bar{x}'}{d\tau} \otimes \frac{d\bar{x}'}{d\tau} \right)$  from the right yields

$$\begin{aligned} \tilde{\Gamma}' \left( \frac{d\bar{x}'}{d\tau} \otimes \frac{d\bar{x}'}{d\tau} \right) &= T \tilde{\Gamma} (\bar{T} \otimes \bar{T}) \underbrace{\left( \frac{d\bar{x}'}{d\tau} \otimes \frac{d\bar{x}'}{d\tau} \right)}_{\left( \frac{d\bar{x}}{d\tau} \otimes \frac{d\bar{x}}{d\tau} \right)} \\ &\quad - \frac{\partial T}{\partial \bar{x}^\top} (\bar{T} \otimes \bar{T}) \underbrace{\left( \frac{d\bar{x}'}{d\tau} \otimes \frac{d\bar{x}'}{d\tau} \right)}_{\left( \frac{d\bar{x}}{d\tau} \otimes \frac{d\bar{x}}{d\tau} \right)}. \end{aligned} \quad (2.111)$$

Adding the two equations (2.110) and (2.111), we finally obtain

$$\frac{d^2 \bar{x}'}{d\tau^2} + \tilde{\Gamma}' \left( \frac{d\bar{x}'}{d\tau} \otimes \frac{d\bar{x}'}{d\tau} \right) = T \left[ \frac{d^2 \bar{x}}{d\tau^2} + \tilde{\Gamma} \left( \frac{d\bar{x}}{d\tau} \otimes \frac{d\bar{x}}{d\tau} \right) \right]. \quad (2.112)$$

The vector in the brackets in (2.112) is transformed as a vector in general! The equation of motion is invariant.

### 2.5.3 Transformation of the Christoffel Matrix $\hat{\Gamma}$

By (2.15),

$$\hat{\Gamma} = U_{4 \times 4} \begin{pmatrix} J^{-1} J_0 \\ \vdots \\ J^{-1} J_3 \end{pmatrix} = U_{4 \times 4} (I_4 \otimes J^{-1}) \frac{\partial J}{\partial \bar{\mathbf{x}}}, \quad (2.113)$$

where

$$J = \frac{\partial \vec{\xi}}{\partial \bar{\mathbf{x}}^\top}.$$

Defining furthermore

$$J' = \frac{\partial \vec{\xi}}{\partial \bar{\mathbf{x}}'^\top}, \quad (2.114)$$

we obtain with the transformation matrices  $T$  and  $\bar{T}$  the relation

$$J = \frac{\partial \vec{\xi}}{\partial \bar{\mathbf{x}}^\top} = \frac{\partial \vec{\xi}}{\partial \bar{\mathbf{x}}'^\top} \frac{\partial \bar{\mathbf{x}}'}{\partial \bar{\mathbf{x}}^\top} = \frac{\partial \vec{\xi}}{\partial \bar{\mathbf{x}}'^\top} T, \quad (2.115)$$

i.e. with (2.114)

$$J = J' T, \quad (2.116)$$

or

$$J' = J \bar{T}. \quad (2.117)$$

Hence

$$\hat{\Gamma}' = U_{4 \times 4} (I_4 \otimes J'^{-1}) \frac{\partial J'}{\partial \bar{\mathbf{x}}'}. \quad (2.118)$$

With

$$J'^{-1} = T J^{-1}$$

and

$$\frac{\partial J'}{\partial \bar{\mathbf{x}}'} = \frac{\partial J \bar{T}}{\partial \bar{\mathbf{x}}'} = \frac{\partial J}{\partial \bar{\mathbf{x}}'} \bar{T} + (I_4 \otimes J) \frac{\partial \bar{T}}{\partial \bar{\mathbf{x}}'},$$

we then obtain

$$\begin{aligned} \hat{\Gamma}' &= U_{4 \times 4} (I_4 \otimes T) (I_4 \otimes J^{-1}) \left( \frac{\partial J}{\partial \bar{\mathbf{x}}'} \bar{T} + (I_4 \otimes J) \frac{\partial \bar{T}}{\partial \bar{\mathbf{x}}'} \right) \\ &= U_{4 \times 4} (I_4 \otimes T) \left( (I_4 \otimes J^{-1}) \frac{\partial J}{\partial \bar{\mathbf{x}}'} \bar{T} + \frac{\partial \bar{T}}{\partial \bar{\mathbf{x}}'} \right). \end{aligned} \quad (2.119)$$

Furthermore,

$$\frac{\partial \mathbf{J}}{\partial \vec{x}'} = \left( \frac{\partial \vec{x}^\top}{\partial \vec{x}'} \otimes \mathbf{I}_4 \right) \frac{\partial \mathbf{J}}{\partial \vec{x}} = (\bar{\mathbf{T}}^\top \otimes \mathbf{I}_4) \frac{\partial \mathbf{J}}{\partial \vec{x}}, \quad (2.120)$$

so

$$\begin{aligned} \hat{\mathbf{T}}' &= U_{4 \times 4} \left[ (\bar{\mathbf{T}}^\top \otimes \mathbf{T} \mathbf{J}^{-1}) \frac{\partial \mathbf{J}}{\partial \vec{x}} \bar{\mathbf{T}} + (\mathbf{I}_4 \otimes \mathbf{T}) \frac{\partial \bar{\mathbf{T}}}{\partial \vec{x}'} \right] \\ &= \underbrace{U_{4 \times 4} (\bar{\mathbf{T}}^\top \otimes \mathbf{T}) U_{4 \times 4}}_{\mathbf{T} \otimes \bar{\mathbf{T}}^\top} \underbrace{U_{4 \times 4} (\mathbf{I}_4 \otimes \mathbf{J}^{-1})}_{\hat{\mathbf{T}}} \frac{\partial \mathbf{J}}{\partial \vec{x}} \bar{\mathbf{T}} + U_{4 \times 4} (\mathbf{I}_4 \otimes \mathbf{T}) \frac{\partial \bar{\mathbf{T}}}{\partial \vec{x}'}, \end{aligned}$$

and with (2.113)

$$\boxed{\hat{\mathbf{T}}' = (\mathbf{T} \otimes \bar{\mathbf{T}}^\top) \hat{\mathbf{T}} \bar{\mathbf{T}} + U_{4 \times 4} (\mathbf{I}_4 \otimes \mathbf{T}) \frac{\partial \bar{\mathbf{T}}}{\partial \vec{x}'}. \quad (2.121)}$$

The second term on the right-hand side expresses again the dependence on the transformation matrix  $\mathbf{T}$ .

### 2.5.4 Coordinate Transformation and Covariant Derivative

It is true that

$$\mathbf{T} \bar{\mathbf{T}} = \mathbf{I},$$

so

$$\frac{\partial}{\partial \vec{x}'} (\mathbf{T} \bar{\mathbf{T}}) = \mathbf{0} = \frac{\partial \mathbf{T}}{\partial \vec{x}'} \bar{\mathbf{T}} + (\mathbf{I}_4 \otimes \mathbf{T}) \frac{\partial \bar{\mathbf{T}}}{\partial \vec{x}'},$$

or

$$(\mathbf{I}_4 \otimes \mathbf{T}) \frac{\partial \bar{\mathbf{T}}}{\partial \vec{x}'} = - \frac{\partial \mathbf{T}}{\partial \vec{x}'} \bar{\mathbf{T}}. \quad (2.122)$$

Look now at the coordinate transformation

$$\mathbf{a}' = \mathbf{T} \mathbf{a} \quad (2.123)$$

and its partial derivative with respect to  $\vec{x}'$ :

$$\frac{\partial \mathbf{a}'}{\partial \vec{x}'} = \frac{\partial}{\partial \vec{x}'} (\mathbf{T} \mathbf{a}) = \frac{\partial \mathbf{T}}{\partial \vec{x}'} \mathbf{a} + (\mathbf{I}_4 \otimes \mathbf{T}) \frac{\partial \mathbf{a}}{\partial \vec{x}'}. \quad (2.124)$$

On the other hand,

$$\frac{\partial \mathbf{a}}{\partial \vec{x}'} = \left( \frac{\partial \vec{x}^\top}{\partial \vec{x}'} \otimes \mathbf{I}_4 \right) \frac{\partial \mathbf{a}}{\partial \vec{x}} = (\bar{\mathbf{T}}^\top \otimes \mathbf{I}_4) \frac{\partial \mathbf{a}}{\partial \vec{x}}.$$

Inserting this into (2.124) yields

$$\frac{\partial \mathbf{a}'}{\partial \bar{\mathbf{x}}'} = (\bar{\mathbf{T}}^\top \otimes \mathbf{T}) \frac{\partial \mathbf{a}}{\partial \bar{\mathbf{x}}} + \frac{\partial \mathbf{T}}{\partial \bar{\mathbf{x}}'} \mathbf{a}. \quad (2.125)$$

How does the product of the matrix  $\mathbf{\Gamma}$  and the vector  $\mathbf{a}$  transform? We have

$$\mathbf{\Gamma}' \mathbf{a}' = (\mathbf{T} \otimes \bar{\mathbf{T}}^\top) \mathbf{\Gamma} \bar{\mathbf{T}} \mathbf{T} \mathbf{a} + U_{4 \times 4} (\mathbf{I}_4 \otimes \mathbf{T}) \frac{\partial \mathbf{T}^{-1}}{\partial \bar{\mathbf{x}}'} \mathbf{T} \mathbf{a}, \quad (2.126)$$

and hence obtain

$$\frac{\partial \mathbf{T}^{-1}}{\partial \bar{\mathbf{x}}'} = -(\mathbf{I}_4 \otimes \mathbf{T}^{-1}) \frac{\partial \mathbf{T}}{\partial \bar{\mathbf{x}}'} \mathbf{T}^{-1}.$$

This in (2.126) yields

$$\mathbf{\Gamma}' \mathbf{a}' = (\mathbf{T} \otimes \bar{\mathbf{T}}^\top) \mathbf{\Gamma} \mathbf{a} - U_{4 \times 4} \frac{\partial \mathbf{T}}{\partial \bar{\mathbf{x}}'} \mathbf{a}. \quad (2.127)$$

After adding (2.125) and (2.127), multiplied from the left by  $U_{4 \times 4}$ , we get

$$\frac{\partial \mathbf{a}'}{\partial \bar{\mathbf{x}}'} + U_{4 \times 4} \mathbf{\Gamma}' \mathbf{a}' = (\bar{\mathbf{T}}^\top \otimes \mathbf{T}) \frac{\partial \mathbf{a}}{\partial \bar{\mathbf{x}}} + U_{4 \times 4} (\mathbf{T} \otimes \bar{\mathbf{T}}^\top) U_{4 \times 4} U_{4 \times 4} \mathbf{\Gamma} \mathbf{a},$$

so

$$\frac{\partial \mathbf{a}'}{\partial \bar{\mathbf{x}}'} + U_{4 \times 4} \mathbf{\Gamma}' \mathbf{a}' = (\bar{\mathbf{T}}^\top \otimes \mathbf{T}) \left[ \frac{\partial \mathbf{a}}{\partial \bar{\mathbf{x}}} + U_{4 \times 4} \mathbf{\Gamma} \mathbf{a} \right]. \quad (2.128)$$

With the definition

$$\mathbf{\Gamma}^* \stackrel{\text{def}}{=} U_{4 \times 4} \mathbf{\Gamma} = \frac{1}{2} (\mathbf{I} \otimes \mathbf{G}^{-1}) \left[ (\mathbf{I}_{16} - U_{4 \times 4}) \frac{\partial \mathbf{G}}{\partial \bar{\mathbf{x}}} + U_{4 \times 4} \begin{pmatrix} \frac{\partial \mathbf{g}_0}{\partial \bar{\mathbf{x}}^\top} \\ \vdots \\ \frac{\partial \mathbf{g}_3}{\partial \bar{\mathbf{x}}^\top} \end{pmatrix} \right] \quad (2.129)$$

the following compact relation is obtained:

$$\boxed{\frac{\partial \mathbf{a}'}{\partial \bar{\mathbf{x}}'} + \mathbf{\Gamma}^* \mathbf{a}' = (\bar{\mathbf{T}}^\top \otimes \mathbf{T}) \left[ \frac{\partial \mathbf{a}}{\partial \bar{\mathbf{x}}} + \mathbf{\Gamma}^* \mathbf{a} \right]}. \quad (2.130)$$

On the right-hand side of this equation, a vector from  $\mathbb{R}^{16}$  is multiplied by a  $16 \times 16$ -matrix! We introduce the abbreviations

$$\mathbf{a}_{|\bar{\mathbf{x}}} \stackrel{\text{def}}{=} \frac{\partial \mathbf{a}}{\partial \bar{\mathbf{x}}} \quad (2.131)$$

and

$$\underline{\underline{\mathbf{a}_{\|\bar{\mathbf{x}}}}} \stackrel{\text{def}}{=} \mathbf{a}_{|\bar{\mathbf{x}}} + \mathbf{\Gamma}^* \mathbf{a}, \quad (2.132)$$

and call  $\mathbf{a}_{\parallel\bar{\mathbf{x}}}$  the *covariant derivative of  $\mathbf{a}$  with respect to  $\bar{\mathbf{x}}$* . Now (2.130) is written as

$$\boxed{\mathbf{a}'_{\parallel\bar{\mathbf{x}'}} = (\bar{\mathbf{T}}^\top \otimes \mathbf{T})\mathbf{a}_{\parallel\bar{\mathbf{x}}}.} \quad (2.133)$$

The form on the right-hand side of (2.130) reminds us of the right-hand side of the expression used in the *vec*-operator (see Appendix), that is,

$$\mathbf{vec}(\mathbf{ABC}) = (\mathbf{C}^\top \otimes \mathbf{A})\mathbf{vec}(\mathbf{B}). \quad (2.134)$$

To use this lemma, (2.130) is first written in some detail. With the new matrix

$$\mathbf{\Gamma}_k^* \stackrel{\text{def}}{=} (\mathbf{i}_k^\top \otimes \mathbf{G}^{-1}) \left[ \frac{1}{2} \mathbf{I}_{16} - \mathbf{U}_{4 \times 4} \right] \frac{\partial \mathbf{G}}{\partial \bar{\mathbf{x}}}, \quad (2.135)$$

where  $\mathbf{i}_k^\top$  is the  $k$ th row of the unit matrix  $\mathbf{I}_4$ , one obtains for (2.130)

$$\begin{aligned} \begin{pmatrix} \frac{\partial \mathbf{a}'}{\partial x'_0} \\ \vdots \\ \frac{\partial \mathbf{a}'}{\partial x'_3} \end{pmatrix} + \begin{pmatrix} \mathbf{\Gamma}_0^{*'} \mathbf{a}' \\ \vdots \\ \mathbf{\Gamma}_3^{*'} \mathbf{a}' \end{pmatrix} &= (\mathbf{I}_4^\top \otimes \mathbf{I}_4) \left[ \begin{pmatrix} \frac{\partial \mathbf{a}'}{\partial x'_0} \\ \vdots \\ \frac{\partial \mathbf{a}'}{\partial x'_3} \end{pmatrix} + \begin{pmatrix} \mathbf{\Gamma}_0^{*'} \mathbf{a}' \\ \vdots \\ \mathbf{\Gamma}_3^{*'} \mathbf{a}' \end{pmatrix} \right] \\ &= (\bar{\mathbf{T}}^\top \otimes \mathbf{T}) \left[ \begin{pmatrix} \frac{\partial \mathbf{a}}{\partial x_0} \\ \vdots \\ \frac{\partial \mathbf{a}}{\partial x_3} \end{pmatrix} + \begin{pmatrix} \mathbf{\Gamma}_0^* \mathbf{a} \\ \vdots \\ \mathbf{\Gamma}_3^* \mathbf{a} \end{pmatrix} \right]. \end{aligned} \quad (2.136)$$

Now Lemma (2.134) applied to this equation shows that

$$\begin{aligned} \left[ \frac{\partial \mathbf{a}'}{\partial x'_0} \middle| \dots \middle| \frac{\partial \mathbf{a}'}{\partial x'_3} \right] + [\mathbf{\Gamma}_0^{*'} \mathbf{a}' \middle| \dots \middle| \mathbf{\Gamma}_3^{*'} \mathbf{a}'] \\ = \mathbf{T} \left\{ \left[ \frac{\partial \mathbf{a}}{\partial x_0} \middle| \dots \middle| \frac{\partial \mathbf{a}}{\partial x_3} \right] + [\mathbf{\Gamma}_0^* \mathbf{a} \middle| \dots \middle| \mathbf{\Gamma}_3^* \mathbf{a}] \right\} \mathbf{T}^{-1}. \end{aligned} \quad (2.137)$$

With

$$\frac{\partial \mathbf{a}}{\partial \bar{\mathbf{x}}^\top} \stackrel{\text{def}}{=} \left[ \frac{\partial \mathbf{a}}{\partial x_0} \middle| \dots \middle| \frac{\partial \mathbf{a}}{\partial x_3} \right] \in \mathbb{R}^{4 \times 4} \quad \text{and} \quad \bar{\mathbf{\Gamma}} \stackrel{\text{def}}{=} [\mathbf{\Gamma}_0^* \middle| \dots \middle| \mathbf{\Gamma}_3^*] \in \mathbb{R}^{4 \times 16}$$

one obtains

$$\boxed{\frac{\partial \mathbf{a}'}{\partial \bar{\mathbf{x}}'^\top} + \bar{\mathbf{\Gamma}}'(\mathbf{I}_4 \otimes \mathbf{a}') = \mathbf{T} \left[ \frac{\partial \mathbf{a}}{\partial \bar{\mathbf{x}}^\top} + \bar{\mathbf{\Gamma}}(\mathbf{I}_4 \otimes \mathbf{a}) \right] \mathbf{T}^{-1}.} \quad (2.138)$$

The sum of matrices

$$\frac{\partial \mathbf{a}}{\partial \bar{\mathbf{x}}^\top} + \bar{\mathbf{\Gamma}}(\mathbf{I}_4 \otimes \mathbf{a})$$

is therefore transformed by a normal similarity transformation in to the following sum of matrices

$$\frac{\partial \mathbf{a}'}{\partial \bar{\mathbf{x}}'^{\top}} + \bar{\Gamma}'(\mathbf{I}_4 \otimes \mathbf{a}').$$

And again some abbreviations are introduced (now the vector  $\bar{\mathbf{x}}$  is transposed!):

$$\mathbf{a}_{|\bar{\mathbf{x}}^{\top}} \stackrel{\text{def}}{=} \frac{\partial \mathbf{a}}{\partial \bar{\mathbf{x}}^{\top}} \tag{2.139}$$

and

$$\mathbf{a}_{\|\bar{\mathbf{x}}^{\top}} \stackrel{\text{def}}{=} \mathbf{a}_{|\bar{\mathbf{x}}^{\top}} + \bar{\Gamma}(\mathbf{I}_4 \otimes \mathbf{a}) \in \mathbb{R}^{4 \times 4}. \tag{2.140}$$

$\mathbf{a}_{\|\bar{\mathbf{x}}^{\top}}$  is called again the *covariant derivative* of  $\mathbf{a}$  and it is

$$\mathbf{a}'_{\|\bar{\mathbf{x}}'^{\top}} = \mathbf{T} \mathbf{a}_{\|\bar{\mathbf{x}}^{\top}} \mathbf{T}^{-1}. \tag{2.141}$$

Important conclusion:

**The formulas in the theory of general relativity are invariant with respect to coordinate transformations if in formulas of special relativity ordinary derivatives  $\frac{\partial \mathbf{a}}{\partial \mathbf{x}^{\top}}$  are replaced by covariant derivatives  $\mathbf{a}_{\|\mathbf{x}^{\top}}$ !**

The covariant derivative defined in (2.140) differs from the derivative defined in (2.100) by the summand  $\bar{\Gamma}(\mathbf{I}_4 \otimes \mathbf{a})$ ; there one has instead  $(\mathbf{I}_4 \otimes \mathbf{a}^{\top})\Gamma$ . But it is indeed true that

$$\bar{\Gamma}(\mathbf{I}_4 \otimes \mathbf{a}) = (\mathbf{I}_4 \otimes \mathbf{a}^{\top})\Gamma. \tag{2.142}$$

In fact, if we rename the  $j$ th row of the sub-matrix  $\Gamma_i$  by  $\boldsymbol{\gamma}_j^{i\top}$ , then the matrix  $\Gamma^*$  is composed as

$$\Gamma^* = U_{4 \times 4} \Gamma = \begin{pmatrix} \boldsymbol{\gamma}_0^{0\top} \\ \boldsymbol{\gamma}_0^{1\top} \\ \boldsymbol{\gamma}_0^{2\top} \\ \boldsymbol{\gamma}_0^{3\top} \\ \hline \vdots \\ \boldsymbol{\gamma}_3^{0\top} \\ \boldsymbol{\gamma}_3^{1\top} \\ \boldsymbol{\gamma}_3^{2\top} \\ \boldsymbol{\gamma}_3^{3\top} \end{pmatrix},$$



i.e.

$$\bar{\Gamma} = \begin{pmatrix} \boldsymbol{\gamma}_0^{0\top} & & \boldsymbol{\gamma}_3^{0\top} \\ \boldsymbol{\gamma}_0^{1\top} & & \boldsymbol{\gamma}_3^{1\top} \\ \boldsymbol{\gamma}_0^{2\top} & \dots & \boldsymbol{\gamma}_3^{2\top} \\ \boldsymbol{\gamma}_0^{3\top} & & \boldsymbol{\gamma}_3^{3\top} \end{pmatrix}. \quad (2.143)$$

With  $\Gamma_i = \Gamma_i^\top$  one obtains:

$$\begin{aligned} \underline{\underline{\bar{\Gamma}(I_4 \otimes \mathbf{a})}} &= \begin{pmatrix} \boldsymbol{\gamma}_0^{0\top} \mathbf{a} & & \boldsymbol{\gamma}_3^{0\top} \mathbf{a} \\ \boldsymbol{\gamma}_0^{1\top} \mathbf{a} & & \boldsymbol{\gamma}_3^{1\top} \mathbf{a} \\ \boldsymbol{\gamma}_0^{2\top} \mathbf{a} & \dots & \boldsymbol{\gamma}_3^{2\top} \mathbf{a} \\ \boldsymbol{\gamma}_0^{3\top} \mathbf{a} & & \boldsymbol{\gamma}_3^{3\top} \mathbf{a} \end{pmatrix} = \begin{pmatrix} \mathbf{a}^\top \boldsymbol{\gamma}_0^0 & & \mathbf{a}^\top \boldsymbol{\gamma}_3^0 \\ \mathbf{a}^\top \boldsymbol{\gamma}_0^1 & & \mathbf{a}^\top \boldsymbol{\gamma}_3^1 \\ \mathbf{a}^\top \boldsymbol{\gamma}_0^2 & \dots & \mathbf{a}^\top \boldsymbol{\gamma}_3^2 \\ \mathbf{a}^\top \boldsymbol{\gamma}_0^3 & & \mathbf{a}^\top \boldsymbol{\gamma}_3^3 \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{a}^\top \Gamma_0^\top \\ \vdots \\ \mathbf{a}^\top \Gamma_3^\top \end{pmatrix} = \begin{pmatrix} \mathbf{a}^\top \Gamma_0 \\ \vdots \\ \mathbf{a}^\top \Gamma_3 \end{pmatrix} = \underline{\underline{(I_4 \otimes \mathbf{a}^\top) \Gamma}}, \end{aligned} \quad (2.144)$$

so we can instead of (2.140) finally write the same as in (2.100)

$$\mathbf{a}_{\|\bar{\mathbf{x}}^\top} = \mathbf{a}_{\|\bar{\mathbf{x}}^\top} + (I_4 \otimes \mathbf{a}^\top) \Gamma \in \mathbb{R}^{4 \times 4}. \quad (2.145)$$

## 2.6 Incidental Remark

If one starts with an equation which is valid in the presence of gravity in general relativity, then this equation must for  $v^2 \ll c^2$  pass over to Newton's equation. The force of interaction of two discrete masses  $m$  and  $m_1$  is proportional to the product of the two known masses and inversely proportional to the square of the distance of the two centres of gravity:

$$\mathbf{f} = G \frac{mm_1}{|\mathbf{x} - \mathbf{x}_1|^2} \frac{\mathbf{x} - \mathbf{x}_1}{|\mathbf{x} - \mathbf{x}_1|},$$

or

$$\mathbf{f} = m \cdot G \frac{m_1}{|\mathbf{x} - \mathbf{x}_1|^3} (\mathbf{x} - \mathbf{x}_1).$$

For several discrete masses  $m_i$ , this attraction is obtained as

$$\mathbf{f} = m \cdot G \sum_i \frac{m_i}{|\mathbf{x} - \mathbf{x}_i|^3} (\mathbf{x} - \mathbf{x}_i),$$

and for a distributed mass with the mass density  $\rho$

$$\mathbf{f} = m \cdot G \int_V \rho(\mathbf{x}_i) \frac{\mathbf{x} - \mathbf{x}_i}{|\mathbf{x} - \mathbf{x}_i|^3} dV.$$

In an electric field of strength  $\mathbf{e}$ , which together with a charge  $q$  produces the force  $\mathbf{f} = q\mathbf{e}$ , we define the gravitational field strength  $\mathbf{e}_G$ , producing the force  $\mathbf{f} = m\mathbf{e}_G$  acting on the mass  $m$ . For several discrete masses, the gravitational field strength is

$$\mathbf{e}_G = G \sum_i \frac{m_i}{|\mathbf{x} - \mathbf{x}_i|^3} (\mathbf{x} - \mathbf{x}_i).$$

So we can divide the analysis of the problem into two steps. In the first step, the gravitational field generated by several masses  $m_i$  at the point  $\mathbf{x}$  is determined, and in the second step, the force acting on the mass  $m$  at the point  $\mathbf{x}$  is determined.

The potential energy is the integral of force times distance, so

$$U = - \int \mathbf{f}^\top ds = -m \int \mathbf{e}_G^\top ds \stackrel{\text{def}}{=} m\phi.$$

If a mass  $m$  is displaced by a small distance  $\Delta x$ , the work done is equal to the change in potential energy

$$\Delta W = -\Delta U = f_x \Delta x.$$

Dividing this equation by  $\Delta x$ , this force in the  $x$ -direction is given as

$$f_x = - \frac{\Delta U}{\Delta x}.$$

Dividing by the mass  $m$ , one gets the  $x$ -component of the strength of the gravitational field

$$e_x = - \frac{\Delta \phi}{\Delta x}.$$

Whence it follows generally with  $\Delta \mathbf{x} \rightarrow \mathbf{0}$ , that

$$\mathbf{e} = -\nabla \phi,$$

so

$$\mathbf{f} = -m\nabla \phi,$$

or

$$\frac{d^2 \mathbf{x}}{dt^2} = -\nabla \phi(\mathbf{x}), \quad (2.146)$$

where the gravitational potential  $\phi$ , which is a scalar function of the position  $\mathbf{x}$ , is obtained from the second-order linear partial differential equation, the Poisson equation

$$\Delta \phi(\mathbf{x}) = 4\pi G \rho(\mathbf{x}), \quad (2.147)$$

with the gravitational constant  $G$  and the mass density  $\rho(\mathbf{x})$ . This equation shows the relationship between gravitational potential and matter in the Newtonian Physics. The above two steps are thus:

**Step 1.** Find the solution  $\phi(\mathbf{x})$  of the Poisson equation (2.147).

**Step 2.** Use the solution of (2.147) to find  $\mathbf{x}(t)$ .

This is the approach in the classical Newtonian Physics. How must we do or modify the two steps in the theory of general relativity, i.e. how does one generally get the  $g_{ik}$ 's and how does one establish the dynamic equations? In the case when only the element  $g_{00}$  depends on  $\mathbf{x}$ , i.e. only the 00-element of the sub-matrix  $\mathbf{F}_i$  is different from zero, then

$$\mathbf{F}_i = \begin{pmatrix} \frac{\partial g_{00}}{\partial x_i} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

According to (2.146), the acceleration is proportional to the partial derivative of the gravitational potential  $\phi$  with respect to the coordinates  $x_i$ . Looking at the equation

$$\ddot{\mathbf{x}} = -(\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \mathbf{F} \dot{\mathbf{x}}, \quad (2.148)$$

on the left-hand side of the equation one has acceleration and on the right-hand side there is the matrix  $\mathbf{F}$  comprising partial derivatives of the  $g_{ij}$ 's with respect to the coordinates  $x_\ell$ . The  $g_{ij}$ 's apparently play in General Relativity the same role as the gravitational potential  $\phi$  does in classical Physics. There the gravitational potential  $\phi$  was determined by the Poisson equation whose form is, for the most part, given by the Laplace operator  $\Delta$ :

$$\Delta\phi = \frac{\partial^2\phi}{\partial x_1^2} + \frac{\partial^2\phi}{\partial x_2^2} + \frac{\partial^2\phi}{\partial x_3^2} = 4\pi G\rho(\mathbf{x}).$$

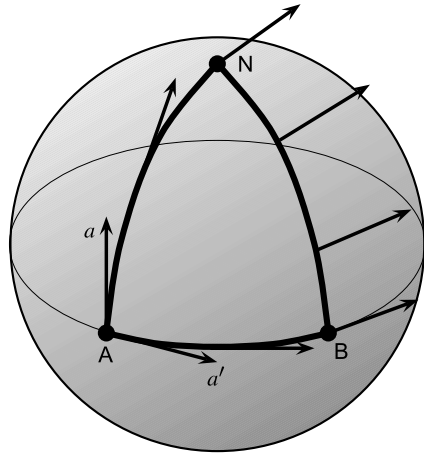
So one is now looking for a mathematical expression involving the second derivatives of the  $g_{ij}$ 's with respect to the four spacetime coordinates  $x_i$ . Such a term appears, in fact, in the differential geometry of Gauss and Riemann, to be precise, in the investigation of the curvature of surfaces or hyper-surfaces in three- or  $n$ -dimensional spaces, where the surfaces are described by quadratic forms with the  $g_{ij}$ 's as the corresponding matrix elements. Therefore, the appendix "Some Differential Geometry" deals with the theory of curvature of surfaces in three- and  $n$ -dimensional spaces.

## 2.7 Parallel Transport

For further considerations, we need the definition of *parallel transport*:

1. The parallel displacement of a vector  $\mathbf{a}$ , which is tangential to the curved surface and runs along a geodesic of this surface, is defined as follows: The origin point of the vector moves along the geodesic and the vector itself moves continuously so that its angle with the geodesic and its length remains constant. It then changes by a parallel transport along  $\delta\mathbf{x}$  according to (2.96) with  $\delta\mathbf{a} = -(\mathbf{I}_4 \otimes \mathbf{a}^\top) \mathbf{F} \delta\mathbf{x}$ .

**Fig. 2.1** Parallel transport



2. The parallel displacement of a vector on a surface along a broken line, consisting of some geodesic pieces, is done so that from the first corner to the second corner the vector is moved along the first geodesic arc, then along the second arc, etc.
3. Finally, the parallel translation of a vector along a smooth curve is described by the limiting process, in which the curve is approximated by broken lines consisting of geodesic pieces.

If a vector  $\mathbf{a}$  in a *flat* space, where  $\mathbf{\Gamma} = \mathbf{0}$ , moves in parallel along a closed loop, it comes back to the starting point with the same length and direction. But if a modified vector comes back, so it must be true that  $\mathbf{\Gamma} \neq \mathbf{0}$ —there exists curvature.

*Example* If you move a vector  $\mathbf{a}$  on a sphere (Fig. 2.1), beginning at the equator at a point A, along a meridian to the north pole N, then along another meridian back to the equator at a point B, and finally back along the equator to the starting point A, then the incoming vector  $\mathbf{a}'$  has a direction other than the initial vector  $\mathbf{a}$ . Calling the difference between the starting and end vector

$$\Delta \mathbf{a} \stackrel{\text{def}}{=} \mathbf{a}' - \mathbf{a},$$

the question is: What happens to  $\Delta \mathbf{a}$  when the circulated area is smaller and smaller? Of course,  $\Delta \mathbf{a}$  approaches the zero vector, but not the ratio  $\Delta \mathbf{a}/(\text{circulated area})$ .

We now replace the spherical triangle of the example by a differentially small square and consider not the difference vector  $\Delta \mathbf{a}$  of a complete circulation. Instead, the initial vector  $\mathbf{a}$  is moved the half way around the square in one direction. Then the same vector  $\mathbf{a}$  is shifted half way in the other direction. At the meeting point the difference  $\Delta \mathbf{a}$  arises. This difference vector is deduced and considered more precisely in the following section.

## 2.8 Riemannian Curvature Matrix

If one moves a vector  $\mathbf{a}(p_0) \in \mathbb{R}^4$  from a point  $p_0$  by  $\delta\mathbf{x} \in \mathbb{R}^4$  to the point  $p_1$ , then it changes according to (2.96) by

$$\delta\mathbf{a} = -(\mathbf{I}_4 \otimes \mathbf{a}(p_0)^\top) \mathbf{\Gamma}(p_0) \delta\mathbf{x}. \quad (2.149)$$

Therefore,

$$\mathbf{a}(p_1) = \mathbf{a}(p_0) + \delta\mathbf{a} = \mathbf{a}(p_0) - (\mathbf{I}_4 \otimes \mathbf{a}(p_0)^\top) \mathbf{\Gamma}(p_0) \delta\mathbf{x}. \quad (2.150)$$

A further displacement of  $p_1$  in the direction  $\delta\bar{\mathbf{x}}$  to  $p_2$  gives the change

$$\delta\bar{\mathbf{a}} = -(\mathbf{I}_4 \otimes \mathbf{a}(p_1)^\top) \mathbf{\Gamma}(p_1) \delta\bar{\mathbf{x}}. \quad (2.151)$$

For  $\mathbf{\Gamma}(p_1)$  as a first approximation one can write

$$\mathbf{\Gamma}(p_1) \approx \mathbf{\Gamma}(p_0) + \sum_{v=0}^3 \frac{\partial \mathbf{\Gamma}}{\partial x_v} \delta x_v = \mathbf{\Gamma}(p_0) + \frac{\partial \mathbf{\Gamma}}{\partial \mathbf{x}^\top} (\delta\mathbf{x} \otimes \mathbf{I}_4). \quad (2.152)$$

Equations (2.150) and (2.152) used in (2.151) yield

$$\begin{aligned} \delta\bar{\mathbf{a}} &= -(\mathbf{I}_4 \otimes [\mathbf{a}(p_0) - (\mathbf{I}_4 \otimes \mathbf{a}(p_0)^\top) \mathbf{\Gamma}(p_0) \delta\mathbf{x}]^\top) \left( \mathbf{\Gamma}(p_0) + \frac{\partial \mathbf{\Gamma}}{\partial \mathbf{x}^\top} (\delta\mathbf{x} \otimes \mathbf{I}_4) \right) \delta\bar{\mathbf{x}} \\ &= -(\mathbf{I}_4 \otimes \mathbf{a}(p_0)^\top) \mathbf{\Gamma}(p_0) \delta\bar{\mathbf{x}} - (\mathbf{I}_4 \otimes \mathbf{a}(p_0)^\top) \frac{\partial \mathbf{\Gamma}}{\partial \mathbf{x}^\top} (\delta\mathbf{x} \otimes \mathbf{I}_4) \delta\bar{\mathbf{x}} \\ &\quad + (\mathbf{I}_4 \otimes [(\mathbf{I}_4 \otimes \mathbf{a}(p_0)^\top) \mathbf{\Gamma} \delta\mathbf{x}]^\top) \mathbf{\Gamma} \delta\bar{\mathbf{x}} + \mathcal{O}(\delta\bar{\mathbf{x}} \cdot (d\mathbf{x}^2)) \\ &= \left[ -(\mathbf{I}_4 \otimes \mathbf{a}^\top) \mathbf{\Gamma} - (\mathbf{I}_4 \otimes \mathbf{a}^\top) \frac{\partial \mathbf{\Gamma}}{\partial \mathbf{x}^\top} (\delta\mathbf{x} \otimes \mathbf{I}_4) \right. \\ &\quad \left. + (\mathbf{I}_4 \otimes [(\mathbf{I}_4 \otimes \mathbf{a}(p_0)^\top) \mathbf{\Gamma} \delta\mathbf{x}]^\top) \mathbf{\Gamma} \right] \delta\bar{\mathbf{x}}. \end{aligned} \quad (2.153)$$

The third term in the brackets can be transformed into the following form:

$$\begin{aligned} (\mathbf{I}_4 \otimes [(\mathbf{I}_4 \otimes \mathbf{a}^\top) \mathbf{\Gamma} \delta\mathbf{x}]^\top) \mathbf{\Gamma} &= \bar{\mathbf{\Gamma}} (\mathbf{I}_4 \otimes (\mathbf{I}_4 \otimes \mathbf{a}^\top) \mathbf{\Gamma} \delta\mathbf{x}) \\ &= \bar{\mathbf{\Gamma}} (\mathbf{I}_{16} \otimes \mathbf{a}^\top) (\mathbf{I}_4 \otimes \mathbf{\Gamma} \delta\mathbf{x}) \\ &= (\bar{\mathbf{\Gamma}} \otimes \mathbf{a}^\top) (\mathbf{I}_4 \otimes \mathbf{\Gamma}) (\mathbf{I}_4 \otimes \delta\mathbf{x}) \\ &= (\mathbf{I}_4 \otimes \mathbf{a}^\top) (\bar{\mathbf{\Gamma}} \otimes \mathbf{I}_4) (\mathbf{I}_4 \otimes \mathbf{\Gamma}) (\mathbf{I}_4 \otimes \delta\mathbf{x}). \end{aligned} \quad (2.154)$$

Using this in (2.153) results in

$$\delta\bar{\mathbf{a}} = -(\mathbf{I}_4 \otimes \mathbf{a}^\top) \left[ \mathbf{\Gamma} + \frac{\partial \mathbf{\Gamma}}{\partial \mathbf{x}^\top} (\delta\mathbf{x} \otimes \mathbf{I}_4) - (\bar{\mathbf{\Gamma}} \otimes \mathbf{I}_4) (\mathbf{I}_4 \otimes \mathbf{\Gamma}) (\mathbf{I}_4 \otimes \delta\mathbf{x}) \right] \delta\bar{\mathbf{x}}. \quad (2.155)$$

If we now first go in the direction  $\delta\bar{\mathbf{x}}$  and then in the direction  $\delta\mathbf{x}$ , we obtain accordingly

$$\delta\bar{\bar{\mathbf{a}}} = -(\mathbf{I}_4 \otimes \mathbf{a}^\top) \left[ \boldsymbol{\Gamma} + \frac{\partial \boldsymbol{\Gamma}}{\partial \mathbf{x}^\top} (\delta\bar{\mathbf{x}} \otimes \mathbf{I}_4) - (\bar{\boldsymbol{\Gamma}} \otimes \mathbf{I}_4)(\mathbf{I}_4 \otimes \boldsymbol{\Gamma})(\mathbf{I}_4 \otimes \delta\bar{\mathbf{x}}) \right] \delta\mathbf{x}. \quad (2.156)$$

For the final product in the third summand of (2.155) we can also write

$$(\mathbf{I}_4 \otimes \delta\mathbf{x})\delta\bar{\mathbf{x}} = (\mathbf{I}_4 \otimes \delta\mathbf{x})(\delta\bar{\mathbf{x}} \otimes 1) = (\delta\bar{\mathbf{x}} \otimes \delta\mathbf{x}) = \mathbf{U}_{4 \times 4}(\delta\mathbf{x} \otimes \delta\bar{\mathbf{x}}). \quad (2.157)$$

It is true that

$$\Delta \mathbf{a} = (\mathbf{a} + \delta\bar{\mathbf{a}}) - (\mathbf{a} + \delta\bar{\bar{\mathbf{a}}}) = \delta\bar{\mathbf{a}} - \delta\bar{\bar{\mathbf{a}}}, \quad (2.158)$$

so with (2.155), (2.156) and (2.157),

$$\begin{aligned} \Delta \mathbf{a} &= (\mathbf{I}_4 \otimes \mathbf{a}^\top) \left( \boldsymbol{\Gamma}(\delta\mathbf{x} - \delta\bar{\mathbf{x}}) \right. \\ &\quad \left. + \left[ \frac{\partial \boldsymbol{\Gamma}}{\partial \mathbf{x}^\top} + (\bar{\boldsymbol{\Gamma}} \otimes \mathbf{I}_4)(\mathbf{I}_4 \otimes \boldsymbol{\Gamma}) \right] (\mathbf{U}_{4 \times 4} - \mathbf{I}_{16})(\delta\mathbf{x} \otimes \delta\bar{\mathbf{x}}) \right). \end{aligned} \quad (2.159)$$

With the  $16 \times 16$  Riemannian curvature matrix

$$\mathbf{R} \stackrel{\text{def}}{=} \left[ \frac{\partial \boldsymbol{\Gamma}}{\partial \mathbf{x}^\top} + (\bar{\boldsymbol{\Gamma}} \otimes \mathbf{I}_4)(\mathbf{I}_4 \otimes \boldsymbol{\Gamma}) \right] (\mathbf{U}_{4 \times 4} - \mathbf{I}_{16}) \in \mathbb{R}^{16 \times 16}, \quad (2.160)$$

we have

$$\Delta \mathbf{a} = (\mathbf{I}_4 \otimes \mathbf{a}^\top) [\boldsymbol{\Gamma}(\delta\mathbf{x} - \delta\bar{\mathbf{x}}) + \mathbf{R}(\delta\mathbf{x} \otimes \delta\bar{\mathbf{x}})] \in \mathbb{R}^4. \quad (2.161)$$

We also define a slightly modified curvature matrix:

$$\check{\mathbf{R}} \stackrel{\text{def}}{=} (\mathbf{G} \otimes \mathbf{I}_4) \mathbf{R}. \quad (2.162)$$

## 2.9 Properties of the Riemannian Curvature Matrix

### 2.9.1 Composition of $\mathbf{R}$ and $\check{\mathbf{R}}$

Which form do the elements of the Riemannian curvature matrix

$$\mathbf{R} = \left[ \frac{\partial \boldsymbol{\Gamma}}{\partial \mathbf{x}^\top} + (\bar{\boldsymbol{\Gamma}} \otimes \mathbf{I}_4)(\mathbf{I}_4 \otimes \boldsymbol{\Gamma}) \right] (\mathbf{U}_{4 \times 4} - \mathbf{I}_{16}) \in \mathbb{R}^{16 \times 16}$$

have? The property, already mentioned earlier, that in the occurring matrix difference  $\mathbf{U}_{4 \times 4} - \mathbf{I}_{16}$  the first,  $(4 + 2)$ th,  $(8 + 3)$ th and the 16-th row/column are equal

to the zero row/column, has for the Riemannian curvature matrix the consequence that its corresponding columns are zero columns!

Next we obtain with (2.143)

$$\bar{\Gamma} = \begin{pmatrix} \boldsymbol{\gamma}_0^{0\top} & & \boldsymbol{\gamma}_3^{0\top} \\ \boldsymbol{\gamma}_0^{1\top} & & \boldsymbol{\gamma}_3^{1\top} \\ \boldsymbol{\gamma}_0^{2\top} & \dots & \boldsymbol{\gamma}_3^{2\top} \\ \boldsymbol{\gamma}_0^{3\top} & & \boldsymbol{\gamma}_3^{3\top} \end{pmatrix} = [\bar{\Gamma}_0, \dots, \bar{\Gamma}_3]$$

for

$$\begin{aligned} (\bar{\Gamma} \otimes I_4)(I_4 \otimes \Gamma) &= [(\bar{\Gamma}_0 \otimes I_4), \dots, (\bar{\Gamma}_3 \otimes I_4)] \begin{pmatrix} \Gamma & 0 & 0 & 0 \\ 0 & \Gamma & 0 & 0 \\ 0 & 0 & \Gamma & 0 \\ 0 & 0 & 0 & \Gamma \end{pmatrix} \\ &= [(\bar{\Gamma}_0 \otimes I_4)\Gamma, \dots, (\bar{\Gamma}_3 \otimes I_4)\Gamma]. \end{aligned}$$

This matrix product contributes to the matrix element  $R_{\alpha\beta}^{\gamma\delta}$  the sum

$$[(\boldsymbol{\gamma}_\delta^{\gamma\top} \otimes I_4)\Gamma]_{\alpha\beta} = [\Gamma_{\delta 0}^\gamma \Gamma_0 + \dots + \Gamma_{\delta 3}^\gamma \Gamma_3]_{\alpha\beta} = \sum_{\nu} \Gamma_{\delta\nu}^\gamma \Gamma_{\alpha\beta}^\nu, \quad (2.163)$$

where  $\boldsymbol{\gamma}_\delta^{\gamma\top} \in \mathbb{R}^4$  is the  $\gamma$ th row of the sub-matrix  $\bar{\Gamma}_\delta$ , i.e. the  $\delta$ th row of the sub-matrix  $\bar{\Gamma}_\gamma$ .

Furthermore,

$$\begin{aligned} &(\bar{\Gamma} \otimes I_4)(I_4 \otimes \Gamma)U_{4 \times 4} \\ &= (\bar{\Gamma} \otimes I_4) \begin{pmatrix} \Gamma & 0 & 0 & 0 \\ 0 & \Gamma & 0 & 0 \\ 0 & 0 & \Gamma & 0 \\ 0 & 0 & 0 & \Gamma \end{pmatrix} U_{4 \times 4} \\ &= (\bar{\Gamma} \otimes I_4) \begin{pmatrix} \boldsymbol{\gamma}_0 & | \boldsymbol{\gamma}_1 & | \boldsymbol{\gamma}_2 & | \boldsymbol{\gamma}_3 \\ \cdot & \cdot & \cdot & \cdot \\ \boldsymbol{\gamma}_0 & | \boldsymbol{\gamma}_1 & | \boldsymbol{\gamma}_2 & | \boldsymbol{\gamma}_3 \end{pmatrix}. \end{aligned}$$

This matrix product ( $\boldsymbol{\gamma}_\delta \in \mathbb{R}^{16}$  is the  $\delta$ th column of  $\Gamma$ ) contributes to the matrix element  $R_{\alpha\beta}^{\gamma\delta}$  the sum

$$\left[ (\boldsymbol{\gamma}^{\gamma\top} \otimes I_4) \begin{pmatrix} \boldsymbol{\gamma}_\delta & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \boldsymbol{\gamma}_\delta \end{pmatrix} \right]_{\alpha\beta}$$

$$\begin{aligned}
&= \left[ \left[ \boldsymbol{\gamma}_0^{\gamma\top} \otimes \mathbf{I}_4, \dots, \boldsymbol{\gamma}_0^{\gamma\top} \otimes \mathbf{I}_4 \right] \begin{pmatrix} \boldsymbol{\gamma}_\delta & & \\ & \ddots & \\ & & \boldsymbol{\gamma}_\delta \end{pmatrix} \right]_{\alpha\beta} \\
&= [(\boldsymbol{\gamma}_0^{\gamma\top} \otimes \mathbf{I}_4)\boldsymbol{\gamma}_\delta, \dots, (\boldsymbol{\gamma}_3^{\gamma\top} \otimes \mathbf{I}_4)\boldsymbol{\gamma}_\delta]_{\alpha\beta} = [(\boldsymbol{\gamma}_\beta^{\gamma\top} \otimes \mathbf{I}_4)\boldsymbol{\gamma}_\delta]_\alpha \\
&= [\Gamma_{\beta 0}^\gamma \boldsymbol{\gamma}_\delta^0 + \dots + \Gamma_{\beta 3}^\gamma \boldsymbol{\gamma}_\delta^3]_\alpha = \sum_{\underline{\nu}} \Gamma_{\beta\nu}^\gamma \Gamma_{\delta\alpha}^\nu. \tag{2.164}
\end{aligned}$$

In accordance with (2.160), one gets finally with (2.163) and (2.164)

$$\boxed{R_{\alpha\beta}^{\gamma\delta} = \frac{\partial}{\partial x_\beta} \Gamma_{\alpha\delta}^\gamma - \frac{\partial}{\partial x_\delta} \Gamma_{\alpha\beta}^\gamma + \sum_{\nu} \Gamma_{\beta\nu}^\gamma \Gamma_{\delta\alpha}^\nu - \sum_{\nu} \Gamma_{\delta\nu}^\gamma \Gamma_{\alpha\beta}^\nu.} \tag{2.165}$$

From this form, one can immediately read off the property

$$\underline{\underline{R_{\alpha\beta}^{\gamma\delta} = -R_{\alpha\delta}^{\gamma\beta}}}. \tag{2.166}$$

With the help of (2.165) one can also verify the so-called *cyclic identity*:

$$\boxed{R_{\alpha\beta}^{\gamma\delta} + R_{\beta\delta}^{\gamma\alpha} + R_{\delta\alpha}^{\gamma\beta} = 0.} \tag{2.167}$$

From (2.165), using (2.63), (2.73) and (2.65), one can also derive a closed form for  $\check{R}_{\alpha\beta}^{\gamma\delta}$  in (2.160) as follows:

$$\begin{aligned}
\check{R}_{\alpha\beta}^{\gamma\delta} &= \sum_i g_{\gamma i} R_{\alpha\beta}^{i\delta} = \sum_i g_{\gamma i} \left( \frac{\partial}{\partial x_\beta} \Gamma_{\alpha\delta}^i - \frac{\partial}{\partial x_\delta} \Gamma_{\alpha\beta}^i + \sum_{\nu} \Gamma_{\beta\nu}^i \Gamma_{\delta\alpha}^\nu - \sum_{\nu} \Gamma_{\delta\nu}^i \Gamma_{\alpha\beta}^\nu \right) \\
&= \left( \frac{\partial}{\partial x_\beta} \check{\Gamma}_{\alpha\delta}^\gamma - \sum_i \Gamma_{\alpha\delta}^i \frac{\partial g_{\gamma i}}{\partial x_\beta} \right) - \left( \frac{\partial}{\partial x_\delta} \check{\Gamma}_{\alpha\beta}^\gamma - \sum_i \Gamma_{\alpha\beta}^i \frac{\partial g_{\gamma i}}{\partial x_\delta} \right) \\
&\quad + \sum_{\nu} \check{\Gamma}_{\beta\nu}^\gamma \Gamma_{\delta\alpha}^\nu - \sum_{\nu} \check{\Gamma}_{\delta\nu}^\gamma \Gamma_{\alpha\beta}^\nu \\
&= \left( \frac{\partial}{\partial x_\beta} \check{\Gamma}_{\alpha\delta}^\gamma - \sum_i \Gamma_{\alpha\delta}^i (\check{\Gamma}_{\gamma\beta}^i + \check{\Gamma}_{i\beta}^\gamma) \right) - \left( \frac{\partial}{\partial x_\delta} \check{\Gamma}_{\alpha\beta}^\gamma \right. \\
&\quad \left. - \sum_i \Gamma_{\alpha\beta}^i (\check{\Gamma}_{\gamma\delta}^i + \check{\Gamma}_{i\delta}^\gamma) \right) + \sum_{\nu} \check{\Gamma}_{\beta\nu}^\gamma \Gamma_{\delta\alpha}^\nu - \sum_{\nu} \check{\Gamma}_{\delta\nu}^\gamma \Gamma_{\alpha\beta}^\nu \\
&= \frac{1}{2} \left( \frac{\partial}{\partial x_\beta} \left( \frac{\partial g_{\delta\gamma}}{\partial x_\alpha} + \frac{\partial g_{\alpha\gamma}}{\partial x_\delta} - \frac{\partial g_{\alpha\delta}}{\partial x_\gamma} \right) - \frac{\partial}{\partial x_\delta} \left( \frac{\partial g_{\beta\gamma}}{\partial x_\alpha} + \frac{\partial g_{\alpha\gamma}}{\partial x_\beta} - \frac{\partial g_{\alpha\beta}}{\partial x_\gamma} \right) \right)
\end{aligned}$$



$$\begin{aligned}
& - \sum_i \Gamma_{\alpha\delta}^i \check{\Gamma}_{\gamma\beta}^i - \sum_i \Gamma_{\alpha\delta}^i \check{\Gamma}_{i\beta}^\gamma + \sum_i \Gamma_{\alpha\beta}^i \check{\Gamma}_{\gamma\delta}^i + \sum_i \Gamma_{\alpha\beta}^i \check{\Gamma}_{i\delta}^\gamma \\
& + \sum_\nu \check{\Gamma}_{\beta\nu}^\gamma \Gamma_{\delta\alpha}^\nu - \sum_\nu \check{\Gamma}_{\delta\nu}^\gamma \Gamma_{\alpha\beta}^\nu.
\end{aligned}$$

After cancelling out some of the terms, we finally get the closed form

$$\check{R}_{\alpha\beta}^{\gamma\delta} = \frac{\partial}{\partial x_\beta} \check{\Gamma}_{\alpha\delta}^\gamma - \frac{\partial}{\partial x_\delta} \check{\Gamma}_{\alpha\beta}^\gamma + \sum_i \Gamma_{\alpha\beta}^i \check{\Gamma}_{\gamma\delta}^i - \sum_i \Gamma_{\alpha\delta}^i \check{\Gamma}_{\gamma\beta}^i \quad (2.168)$$

$$\begin{aligned}
& = \frac{1}{2} \left( \frac{\partial^2 g_{\delta\gamma}}{\partial x_\alpha \partial x_\beta} - \frac{\partial^2 g_{\alpha\delta}}{\partial x_\gamma \partial x_\beta} - \frac{\partial^2 g_{\beta\gamma}}{\partial x_\alpha \partial x_\delta} + \frac{\partial^2 g_{\alpha\beta}}{\partial x_\gamma \partial x_\delta} \right) \\
& + \sum_i \Gamma_{\alpha\beta}^i \check{\Gamma}_{\gamma\delta}^i - \sum_i \Gamma_{\alpha\delta}^i \check{\Gamma}_{\gamma\beta}^i,
\end{aligned} \quad (2.169)$$

or

$$\begin{aligned}
\check{R}_{\alpha\beta}^{\gamma\delta} & = \frac{1}{2} \left( \frac{\partial^2 g_{\delta\gamma}}{\partial x_\alpha \partial x_\beta} - \frac{\partial^2 g_{\alpha\delta}}{\partial x_\gamma \partial x_\beta} - \frac{\partial^2 g_{\beta\gamma}}{\partial x_\alpha \partial x_\delta} + \frac{\partial^2 g_{\alpha\beta}}{\partial x_\gamma \partial x_\delta} \right) \\
& + \sum_i \check{\Gamma}_{\gamma\delta}^i \sum_\nu g_{i\nu}^{(-1)} \check{\Gamma}_{\alpha\beta}^\nu - \sum_i \check{\Gamma}_{\gamma\beta}^i \sum_\nu g_{i\nu}^{(-1)} \check{\Gamma}_{\alpha\delta}^\nu.
\end{aligned} \quad (2.170)$$

From (2.170) the following identities are obtained directly by comparing the corresponding forms:

$$\underline{\underline{\check{R}_{\alpha\beta}^{\gamma\delta}}} = -\underline{\underline{\check{R}_{\gamma\beta}^{\alpha\delta}}}, \quad (2.171)$$

$$\underline{\underline{\check{R}_{\alpha\beta}^{\gamma\delta}}} = -\underline{\underline{\check{R}_{\alpha\delta}^{\gamma\beta}}}, \quad (2.172)$$

$$\underline{\underline{\check{R}_{\alpha\beta}^{\gamma\delta}}} = \underline{\underline{\check{R}_{\gamma\delta}^{\alpha\beta}}}. \quad (2.173)$$

Also in this case, a *cyclic identity* is valid:

$$\check{R}_{\alpha\beta}^{\gamma\delta} + \check{R}_{\beta\delta}^{\gamma\alpha} + \check{R}_{\delta\alpha}^{\gamma\beta} = 0. \quad (2.174)$$

If the two vectors  $\mathbf{dx}$  and  $\mathbf{d\bar{x}}$  are perpendicular to each other, then the area of the formed rectangle is equal to  $|\mathbf{dx}| \cdot |\mathbf{d\bar{x}}|$ . In the *differential geometry*, the limit of the ratio of  $\Delta a$  to the area is now called the *curvature*  $\kappa$ :

$$\kappa \stackrel{\text{def}}{=} \lim_{|\mathbf{dx}|, |\mathbf{d\bar{x}}| \rightarrow 0} \frac{|\Delta a(\mathbf{dx}, \mathbf{d\bar{x}})|}{|\mathbf{dx}| \cdot |\mathbf{d\bar{x}}|}$$

or with  $\mathbf{n} \stackrel{\text{def}}{=} \frac{d\mathbf{x}}{|d\mathbf{x}|}$  and  $\bar{\mathbf{n}} \stackrel{\text{def}}{=} \frac{d\bar{\mathbf{x}}}{|d\bar{\mathbf{x}}|}$ ,

$$\kappa \stackrel{\text{def}}{=} \lim_{\epsilon \rightarrow 0} \frac{|\Delta \mathbf{a}(\epsilon \mathbf{n}, \epsilon \bar{\mathbf{n}})|}{\epsilon^2}.$$

With the help of (2.161) we then obtain

$$\kappa = |(\mathbf{I}_4 \otimes \mathbf{a}^\top) \mathbf{R}(\mathbf{n} \otimes \bar{\mathbf{n}})|. \quad (2.175)$$

**Riemannian Coordinate System** For the study of the properties of the Riemannian curvature matrix, it is advantageous to first perform a coordinate transformation, so that in the new coordinate system, the Christoffel matrices have the property  $\mathbf{\Gamma} = \mathbf{0}$ . Such a coordinate transformation is in the case of a curved spacetime, where  $\mathbf{G}$  depends on  $\bar{\mathbf{x}}$ , only locally possible, but then for each  $\bar{\mathbf{x}}$ ! By the inverse transformation of the obtained statements, they are again globally valid. So, we look for a local coordinate transformation such that in the new coordinate system  $\mathbf{\Gamma} = \mathbf{0}$  is valid. For geodesic lines the following applies to the four coordinates:

$$\frac{d^2 x_k}{ds^2} + \left( \frac{d\bar{\mathbf{x}}}{ds} \right)^\top \mathbf{\Gamma}_k \frac{d\bar{\mathbf{x}}}{ds} = 0. \quad (2.176)$$

The  $x_k$ 's belong to an arbitrary coordinate system in which the geodesic is described by  $x_k = x_k(s)$  and  $s$  is the arc length along the curve. At a fixed point  $\mathcal{P}$  with the coordinate  $\bar{\mathbf{x}}^{(0)}$ , any coordinate can be developed in a power series:

$$x_k = x_k^{(0)} + \zeta_k s + \frac{1}{2} \left( \frac{d^2 x_k}{ds^2} \right)_{\mathcal{P}} s^2 + \frac{1}{3!} \left( \frac{d^3 x_k}{ds^3} \right)_{\mathcal{P}} s^3 + \dots, \quad (2.177)$$

where  $\zeta_k$  is the  $k$ th component of the tangent vector

$$\boldsymbol{\zeta} \stackrel{\text{def}}{=} \left( \frac{d\bar{\mathbf{x}}}{ds} \right)_{\mathcal{P}}$$

to the geodesic at the point  $\mathcal{P}$ . Then, however, according to (2.176),

$$\frac{d^2 x_k}{ds^2} = -\boldsymbol{\zeta}^\top (\mathbf{\Gamma}_k)_{\mathcal{P}} \boldsymbol{\zeta}. \quad (2.178)$$

Plugging this into (2.177), considering a small neighbourhood of  $\mathcal{P}$ , a small  $x_k - x_k^{(0)}$ , and neglecting powers higher than two gives

$$x_k = x_k^{(0)} + \zeta_k s - \frac{1}{2} \boldsymbol{\zeta}^\top (\mathbf{\Gamma}_k)_{\mathcal{P}} \boldsymbol{\zeta} s^2. \quad (2.179)$$

Now, calling  $\boldsymbol{\zeta} s = \bar{\mathbf{x}}'$ , the following is obtained from (2.179):

$$x_k = x_k^{(0)} + x_k' - \frac{1}{2} \bar{\mathbf{x}}'^\top (\mathbf{\Gamma}_k)_{\mathcal{P}} \bar{\mathbf{x}}'.$$

This relation suggests the following coordinate transformation from  $\vec{x}$  to  $\vec{x}'$ :

$$x'_k = x_k - x_k^{(0)} + \frac{1}{2}(\vec{x} - \vec{x}^{(0)})^\top (\Gamma_k)_{\mathcal{P}} (\vec{x} - \vec{x}^{(0)}). \quad (2.180)$$

What is the corresponding metric matrix  $\mathbf{G}'$  in

$$ds^2 = \vec{x}'^\top \mathbf{G}' \vec{x}'? \quad (2.181)$$

In this coordinate system, the geodesic has the equation

$$\frac{d^2 \vec{x}'}{ds^2} + \left( \mathbf{I}_4 \otimes \left( \frac{d\vec{x}'}{ds} \right)^\top \right) \Gamma'(\vec{x}') \left( \frac{d\vec{x}'}{ds} \right) = \mathbf{0}. \quad (2.182)$$

But since in the new coordinate system the geodesics are straight lines of the form  $\vec{x}' = \zeta s$ , in (2.182) the expression  $(\mathbf{I}_4 \otimes \zeta^\top) \Gamma'(\zeta s) \zeta$  must be equal to the zero vector. Since  $\zeta$  are arbitrary vectors, the following must be valid for  $s = 0$ :

$$\underline{\underline{\Gamma'(0) = \mathbf{0}}}. \quad (2.183)$$

**Implications for the Riemannian Curvature Matrix** If at a point  $\mathcal{P}$  for a special coordinate system one has  $\Gamma_{\mathcal{P}} = \mathbf{0}$ , then naturally any  $\check{\Gamma}_{\alpha\beta}^\gamma$  defined in (2.63) is equal to zero. But since

$$\underline{\underline{\check{\Gamma}_{k\ell}^i + \check{\Gamma}_{ki}^\ell = \frac{1}{2} \left( \frac{\partial g_{\ell i}}{\partial x_k} + \frac{\partial g_{ki}}{\partial x_\ell} - \frac{\partial g_{k\ell}}{\partial x_i} \right) + \frac{1}{2} \left( \frac{\partial g_{i\ell}}{\partial x_k} + \frac{\partial g_{k\ell}}{\partial x_i} - \frac{\partial g_{ki}}{\partial x_\ell} \right) = \frac{\partial g_{i\ell}}{\partial x_k}}}. \quad (2.184)$$

also all first partial derivatives of the elements of the metric matrix are zero, i.e.

$$\underline{\underline{\frac{\partial \mathbf{G}}{\partial \vec{x}} \Big|_{\mathcal{P}} = \mathbf{0}}}. \quad (2.185)$$

In the local coordinate system with  $\Gamma_{\mathcal{P}} = \mathbf{0}$ , the Riemannian curvature matrix has the form

$$\mathbf{R}_{\mathcal{P}} = \frac{\partial \Gamma}{\partial \vec{x}^\top} \Big|_{\mathcal{P}} (\mathbf{U}_{4 \times 4} - \mathbf{I}_{16}). \quad (2.186)$$

The structure is

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}^{00} & \mathbf{R}^{01} & \mathbf{R}^{02} & \mathbf{R}^{03} \\ \mathbf{R}^{10} & \mathbf{R}^{11} & \mathbf{R}^{12} & \mathbf{R}^{13} \\ \mathbf{R}^{20} & \mathbf{R}^{21} & \mathbf{R}^{22} & \mathbf{R}^{23} \\ \mathbf{R}^{30} & \mathbf{R}^{31} & \mathbf{R}^{32} & \mathbf{R}^{33} \end{pmatrix}. \quad (2.187)$$

Each sub-matrix  $\mathbf{R}^{\gamma\delta}$  has the form

$$\mathbf{R}^{\gamma\delta} = \begin{pmatrix} R_{00}^{\gamma\delta} & R_{01}^{\gamma\delta} & R_{02}^{\gamma\delta} & R_{03}^{\gamma\delta} \\ R_{10}^{\gamma\delta} & R_{11}^{\gamma\delta} & R_{12}^{\gamma\delta} & R_{13}^{\gamma\delta} \\ R_{20}^{\gamma\delta} & R_{21}^{\gamma\delta} & R_{22}^{\gamma\delta} & R_{23}^{\gamma\delta} \\ R_{30}^{\gamma\delta} & R_{31}^{\gamma\delta} & R_{32}^{\gamma\delta} & R_{33}^{\gamma\delta} \end{pmatrix}. \quad (2.188)$$

$\mathbf{R}^{\gamma\delta}$  is thus the sub-matrix of  $\mathbf{R}$  at the intersection of the  $\gamma$ th row and  $\delta$ th column. Furthermore,  $R_{\alpha\beta}^{\gamma\delta}$  is the matrix element of the sub-matrix  $\mathbf{R}^{\gamma\delta}$  at the intersection of the  $\alpha$ th row and  $\beta$ th column.

With (2.68) and (2.185) one has

$$\begin{aligned} \left. \frac{\partial \mathbf{\Gamma}}{\partial \bar{\mathbf{x}}^\top} \right|_{\mathcal{P}} &= \frac{1}{2} \frac{\partial}{\partial \bar{\mathbf{x}}^\top} \left( (\mathbf{G}^{-1} \otimes \mathbf{I}_4) \left[ \begin{pmatrix} \frac{\partial \mathbf{g}_0^\top}{\partial \bar{\mathbf{x}}} \\ \vdots \\ \frac{\partial \mathbf{g}_3^\top}{\partial \bar{\mathbf{x}}} \end{pmatrix} + \begin{pmatrix} \frac{\partial \mathbf{g}_0^\top}{\partial \bar{\mathbf{x}}^\top} \\ \vdots \\ \frac{\partial \mathbf{g}_3^\top}{\partial \bar{\mathbf{x}}^\top} \end{pmatrix} - \frac{\partial \mathbf{G}}{\partial \bar{\mathbf{x}}} \right] \right) \\ &= \frac{1}{2} \frac{\partial}{\partial \bar{\mathbf{x}}^\top} (\mathbf{G}^{-1} \otimes \mathbf{I}_4) [\mathbf{0} + \mathbf{0} - \mathbf{0}] \\ &\quad + \frac{1}{2} (\mathbf{G}^{-1} \otimes \mathbf{I}_4) \left[ \begin{pmatrix} \frac{\partial^2 \mathbf{g}_0^\top}{\partial \mathbf{x}^\top \partial \bar{\mathbf{x}}} \\ \vdots \\ \frac{\partial^2 \mathbf{g}_3^\top}{\partial \mathbf{x}^\top \partial \bar{\mathbf{x}}} \end{pmatrix} + \begin{pmatrix} \frac{\partial^2 \mathbf{g}_0^\top}{\partial \bar{\mathbf{x}}^\top \partial \bar{\mathbf{x}}} \\ \vdots \\ \frac{\partial^2 \mathbf{g}_3^\top}{\partial \bar{\mathbf{x}}^\top \partial \bar{\mathbf{x}}} \end{pmatrix} - \frac{\partial^2 \mathbf{G}}{\partial \bar{\mathbf{x}}^\top \partial \bar{\mathbf{x}}} \right]. \end{aligned} \quad (2.189)$$

For

$$\check{\mathbf{R}} = (\mathbf{G} \otimes \mathbf{I}_4) \mathbf{R} \quad (2.190)$$

with

$$\check{\mathbf{\Gamma}} \stackrel{\text{def}}{=} (\mathbf{G} \otimes \mathbf{I}_4) \mathbf{\Gamma} \quad (2.191)$$

one finally obtains

$$\check{\mathbf{R}}_{\mathcal{P}} \stackrel{\text{def}}{=} \frac{\partial \check{\mathbf{\Gamma}}}{\partial \bar{\mathbf{x}}^\top} (\mathbf{U}_{4 \times 4} - \mathbf{I}_{16}). \quad (2.192)$$

Naturally, because of the in  $\mathbf{R}$  occurring matrix difference  $\mathbf{U}_{4 \times 4} - \mathbf{I}_{16}$ , the first,  $(4 + 2)$ th,  $(8 + 3)$ th and the 16th column of  $\check{\mathbf{R}}$  and of  $\check{\mathbf{R}}_{\mathcal{P}}$  are equal to the zero column. In

$$\check{\mathbf{\Gamma}} = \begin{pmatrix} \check{\mathbf{\Gamma}}_0 \\ \check{\mathbf{\Gamma}}_1 \\ \check{\mathbf{\Gamma}}_2 \\ \check{\mathbf{\Gamma}}_3 \end{pmatrix},$$

$\check{\Gamma}_\nu$  has the form

$$\check{\Gamma}_\nu = \frac{1}{2} \left( \frac{\partial \mathbf{g}_\nu^\top}{\partial \mathbf{x}} + \frac{\partial \mathbf{g}_\nu}{\partial \mathbf{x}^\top} - \frac{\partial \mathbf{G}}{\partial x_\nu} \right),$$

i.e. the elements of  $\check{\Gamma}_\nu$  are (see also (2.63))

$$\check{\Gamma}_{\alpha\beta}^\nu = \frac{1}{2} \left( \frac{\partial g_{\beta\nu}}{\partial x_\alpha} + \frac{\partial g_{\alpha\nu}}{\partial x_\beta} - \frac{\partial g_{\alpha\beta}}{\partial x_\nu} \right). \quad (2.193)$$

For  $(\check{R}_{\alpha\beta}^{\gamma\delta})_{\mathcal{P}}$  from (2.165) one obtains

$$\underline{\underline{(\check{R}_{\alpha\beta}^{\gamma\delta})_{\mathcal{P}}}} = \frac{1}{2} \left( \frac{\partial^2 g_{\gamma\delta}}{\partial x_\alpha \partial x_\beta} - \frac{\partial^2 g_{\alpha\delta}}{\partial x_\gamma \partial x_\beta} - \frac{\partial^2 g_{\beta\gamma}}{\partial x_\alpha \partial x_\delta} + \frac{\partial^2 g_{\alpha\beta}}{\partial x_\gamma \partial x_\delta} \right). \quad (2.194)$$

From (2.194), by comparing the corresponding forms, follows

$$\underline{\underline{(\check{R}_{\alpha\beta}^{\gamma\delta})_{\mathcal{P}}}} = \underline{\underline{(\check{R}_{\alpha\beta}^{\delta\gamma})_{\mathcal{P}}}}, \quad (2.195)$$

and the so-called *cyclic identity*

$$\underline{\underline{(\check{R}_{\alpha\beta}^{\gamma\delta})_{\mathcal{P}}}} + \underline{\underline{(\check{R}_{\alpha\gamma}^{\delta\beta})_{\mathcal{P}}}} + \underline{\underline{(\check{R}_{\alpha\delta}^{\beta\gamma})_{\mathcal{P}}}} = 0. \quad (2.196)$$

From (2.165), at a point  $\mathcal{P}$ , we get

$$\underline{\underline{(R_{\alpha\beta}^{\gamma\delta})_{\mathcal{P}}}} = \left( \frac{\partial}{\partial x_\beta} \Gamma_{\alpha\delta}^\gamma - \frac{\partial}{\partial x_\delta} \Gamma_{\alpha\beta}^\gamma \right)_{\mathcal{P}}. \quad (2.197)$$

Partially differentiating (2.197) yields

$$\underline{\underline{\left( \frac{\partial}{\partial x_\kappa} R_{\alpha\beta}^{\gamma\delta} \right)_{\mathcal{P}}}} = \left( \frac{\partial^2}{\partial x_\kappa \partial x_\beta} \Gamma_{\alpha\delta}^\gamma - \frac{\partial^2}{\partial x_\kappa \partial x_\delta} \Gamma_{\alpha\beta}^\gamma \right)_{\mathcal{P}}. \quad (2.198)$$

With the help of (2.198), by substituting the corresponding terms, the so-called *Bianchi identity* is obtained:

$$\underline{\underline{\frac{\partial}{\partial x_\kappa} R_{\alpha\beta}^{\gamma\delta} + \frac{\partial}{\partial x_\beta} R_{\alpha\delta}^{\gamma\kappa} + \frac{\partial}{\partial x_\delta} R_{\alpha\kappa}^{\gamma\beta} = 0.}} \quad (2.199)$$

Since this is valid in *any* event  $\mathcal{P}$ , it is valid *anywhere*.

## 2.10 The Ricci Matrix and Its Properties

The reason for considering the Riemannian curvature theory is to find with its help a way to determine the components of the Christoffel matrix  $\mathbf{\Gamma}$ , which are needed

to calculate the solution of the equations

$$\ddot{\mathbf{x}} = -(\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \mathbf{\Gamma} \dot{\mathbf{x}},$$

describing the dynamic behaviour of particles in a gravitational field. But the Riemannian curvature matrix has, being a  $16 \times 16$ -matrix, 256 components, so, it would provide 256 equations in extreme cases. Considering, however, that there are 4 rows and 4 columns of  $\mathbf{R}$  with all zeros, only  $12 \cdot 12 = 144$  equations remain—still too many. The number of equations can be reduced significantly, however, by reducing, using clever addition of matrix elements, the number of components of the newly formed matrix. Setting up in this way a *symmetric*  $4 \times 4$ -matrix, we would get exactly 10 independent equations for the determination of the 10 independent components of  $\mathbf{G}$ . Such a matrix is called the Ricci matrix, which can be obtained in two ways. One way is through the sum of sub-matrices on the diagonal of  $\mathbf{R}$ ; it is described in the Appendix. The second way is as follows:

The Ricci matrix  $\mathbf{R}_{\text{Ric}}$  consists of the traces of sub-matrices  $\mathbf{R}^{\gamma\delta}$  of  $\mathbf{R}$

$$R_{\text{Ric},\gamma\delta} \stackrel{\text{def}}{=} \text{trace}(\mathbf{R}^{\gamma\delta}) = \sum_{\nu=0}^3 R_{\nu\nu}^{\gamma\delta}. \quad (2.200)$$

Accordingly, the components of the new matrix  $\check{\mathbf{R}}_{\text{Ric}}$  are defined as

$$\check{R}_{\text{Ric},\gamma\delta} \stackrel{\text{def}}{=} \sum_{\nu=0}^3 \check{R}_{\nu\nu}^{\gamma\delta}. \quad (2.201)$$

From (2.194) one may read immediately that the Ricci matrix  $\check{\mathbf{R}}_{\text{Ric}}$  is *symmetric* because  $\check{R}_{\nu\nu}^{\gamma\delta} = \check{R}_{\nu\nu}^{\delta\gamma}$ .

Moreover, because of (2.162),

$$\mathbf{R} = (\mathbf{G}^{-1} \otimes \mathbf{I}_4) \check{\mathbf{R}}, \quad (2.202)$$

so

$$\mathbf{R}^{\gamma\delta} = (\mathbf{g}_\gamma^{-T} \otimes \mathbf{I}_4) \check{\mathbf{R}}^\delta = \sum_{\mu=0}^3 g_{\gamma\mu}^{[-1]} \check{\mathbf{R}}^{\mu\delta}, \quad (2.203)$$

where  $\mathbf{g}_\gamma^{-T}$  is the  $\gamma$ th row of  $\mathbf{G}^{-1}$  and  $\check{\mathbf{R}}^\delta$  the  $16 \times 4$ -matrix consisting of the four sub-matrices in the  $\delta$ th block column of  $\check{\mathbf{R}}$ , i.e. for the matrix elements one has

$$R_{\alpha\beta}^{\gamma\delta} = \sum_{\mu=0}^3 g_{\gamma\mu}^{[-1]} \check{R}_{\alpha\beta}^{\mu\delta}. \quad (2.204)$$

With the help of (2.200), for the Ricci matrix components one gets

$$R_{\text{Ric},\gamma\delta} = \sum_{\nu} R_{\nu\nu}^{\gamma\delta} = \sum_{\nu} \sum_{\mu} g_{\gamma\mu}^{[-1]} \check{R}_{\nu\nu}^{\mu\delta}, \quad (2.205)$$

or, due to (2.173),

$$R_{\text{Ric},\gamma\delta} = \sum_{\nu} \sum_{\mu} g_{\gamma\mu}^{[-1]} \check{R}_{\mu\delta}^{\nu\nu}. \quad (2.206)$$

The *curvature scalar*  $R$  is obtained by taking the trace of the Ricci matrix:

$$\begin{aligned} R &\stackrel{\text{def}}{=} \text{trace}(\mathbf{R}_{\text{Ric}}) = \sum_{\alpha} \sum_{\nu} R_{\nu\nu}^{\alpha\alpha} \\ &= \sum_{\alpha} \sum_{\nu} \sum_{\mu} g_{\alpha\mu}^{[-1]} \check{R}_{\nu\nu}^{\mu\alpha} = \sum_{\alpha} \sum_{\mu} g_{\alpha\mu}^{[-1]} \check{R}_{\text{Ric},\mu\alpha}. \end{aligned} \quad (2.207)$$

Conversely, one obtains

$$\check{R}_{\alpha\beta}^{\gamma\delta} = \sum_{\mu=0}^3 g_{\gamma\mu} R_{\alpha\beta}^{\mu\delta}. \quad (2.208)$$

Because of (2.165) it follows directly that

$$R_{\text{Ric},\gamma\delta} = \sum_{\nu=0}^3 \left( \frac{\partial}{\partial x_{\delta}} \Gamma_{\gamma\nu}^{\nu} - \frac{\partial}{\partial x_{\nu}} \Gamma_{\gamma\delta}^{\nu} + \sum_{\mu=0}^3 \Gamma_{\delta\mu}^{\nu} \Gamma_{\nu\gamma}^{\mu} - \sum_{\mu=0}^3 \Gamma_{\nu\mu}^{\mu} \Gamma_{\gamma\delta}^{\mu} \right) \quad (2.209)$$

and from (2.168)

$$\check{R}_{\text{Ric},\gamma\delta} = \sum_{\nu=0}^3 \left( \frac{\partial}{\partial x_{\delta}} \check{\Gamma}_{\gamma\nu}^{\nu} - \frac{\partial}{\partial x_{\nu}} \check{\Gamma}_{\gamma\delta}^{\nu} + \sum_{\mu=0}^3 \Gamma_{\gamma\delta}^{\mu} \check{\Gamma}_{\nu\nu}^{\mu} - \sum_{\mu=0}^3 \Gamma_{\gamma\nu}^{\mu} \check{\Gamma}_{\nu\delta}^{\mu} \right). \quad (2.210)$$

### 2.10.1 Symmetry of the Ricci Matrix $\mathbf{R}_{\text{Ric}}$

Even if the Riemannian matrix  $\mathbf{R}$  itself is not symmetric, the derived Ricci matrix  $\mathbf{R}_{\text{Ric}}$  is nevertheless symmetric; and that will be shown below. The symmetry will follow from the component equations (2.209) of the Ricci matrix. One sees immediately that the second and fourth summand are symmetric in  $\gamma$  and  $\delta$ . Looking at  $\sum_{\nu=0}^3 \frac{\partial}{\partial x_{\delta}} \Gamma_{\gamma\nu}^{\nu}$ , it is not directly seen that this term is symmetric in  $\gamma$  and  $\delta$ . This can be seen using the Laplace expansion theorem for determinants.<sup>1</sup> Developing the determinant of  $\mathbf{G}$  along the  $\nu$ th row, one gets

$$g \stackrel{\text{def}}{=} \det(\mathbf{G}) = g_{\nu 1} A_{\nu 1} + \cdots + g_{\nu \delta} A_{\nu \delta} + \cdots + g_{\nu n} A_{\nu n},$$

<sup>1</sup>The sum of the products of all elements of a row (or column) with its adjuncts is equal to the determinant's value:  $\det A = \sum_{j=1}^n (-1)^{i+j} \cdot a_{ij} \cdot \det A_{ij}$  (development along the  $i$ th row)

where  $A_{\nu\delta}$  is the element in the  $\nu$ th row and  $\delta$ th column of the adjoint of  $\mathbf{G}$ . If  $g_{\delta\nu}^{[-1]}$  is the  $(\delta\nu)$ -element of the inverse matrix  $\mathbf{G}^{-1}$ , then  $g_{\delta\nu}^{[-1]} = \frac{1}{g} A_{\nu\delta}$ , so  $A_{\nu\delta} = g g_{\delta\nu}^{[-1]}$ . Thus we obtain

$$\frac{\partial g}{\partial g_{\nu\delta}} = A_{\nu\delta} = g g_{\delta\nu}^{[-1]},$$

or

$$\delta g = g g_{\delta\nu}^{[-1]} \delta g_{\nu\delta},$$

and

$$\frac{\partial g}{\partial x_\gamma} = g g_{\delta\nu}^{[-1]} \frac{\partial g_{\nu\delta}}{\partial x_\gamma},$$

i.e.

$$\frac{1}{g} \frac{\partial g}{\partial x_\gamma} = g_{\delta\nu}^{[-1]} \frac{\partial g_{\nu\delta}}{\partial x_\gamma}. \quad (2.211)$$

Due to (2.62), on the other hand, one has

$$\sum_{\nu=0}^3 \Gamma_{\gamma\nu}^\nu = \sum_{\nu=0}^3 \sum_{\delta=0}^3 \frac{g_{\delta\nu}^{[-1]}}{2} \left( \frac{\partial g_{\nu\delta}}{\partial x_\gamma} + \frac{\partial g_{\gamma\delta}}{\partial x_\nu} - \frac{\partial g_{\gamma\nu}}{\partial x_\delta} \right),$$

i.e. the last two summands cancel out, and the following remains:

$$\sum_{\nu=0}^3 \Gamma_{\gamma\nu}^\nu = \sum_{\nu=0}^3 \sum_{\delta=0}^3 \frac{g_{\delta\nu}^{[-1]}}{2} \frac{\partial g_{\nu\delta}}{\partial x_\gamma}.$$

Then it follows by (2.211) that

$$\underline{\underline{\sum_{\nu=0}^3 \frac{\partial}{\partial x_\delta} \Gamma_{\gamma\nu}^\nu = \sum_{\nu=0}^3 \sum_{\delta=0}^3 \frac{1}{\sqrt{|g|}} \frac{\partial^2 \sqrt{|g|}}{\partial x_\gamma \partial x_\delta}}}. \quad (2.212)$$

From this form the symmetry in  $\gamma$  and  $\delta$  can immediately be observed.

Now remains to show that the third term

$$\sum_{\nu=0}^3 \sum_{\mu=0}^3 \Gamma_{\delta\mu}^\nu \Gamma_{\nu\gamma}^\mu$$

in (2.209) is symmetric. But this can be spotted from the following equalities:

$$\sum_{\nu,\mu=0}^3 \Gamma_{\delta\mu}^\nu \Gamma_{\nu\gamma}^\mu = \sum_{\nu,\mu=0}^3 \Gamma_{\mu\delta}^\nu \Gamma_{\gamma\nu}^\mu = \sum_{\nu,\mu=0}^3 \Gamma_{\nu\delta}^\mu \Gamma_{\gamma\mu}^\nu.$$

Combining everything together demonstrates that the Ricci matrix  $\mathbf{R}_{\text{Ric}}$  is, in fact, symmetric!



### 2.10.2 The Divergence of the Ricci Matrix

Multiplying the Bianchi identities (2.199) in the form

$$\frac{\partial}{\partial x_\kappa} R_{\alpha\beta}^{\nu\delta} + \frac{\partial}{\partial x_\beta} R_{\alpha\delta}^{\nu\kappa} + \frac{\partial}{\partial x_\delta} R_{\alpha\kappa}^{\nu\beta} = 0$$

with  $g_{\gamma\nu}$  and summing over  $\nu$ , we obtain at a point  $\mathcal{P}$ , where  $\frac{\partial G}{\partial x} = \mathbf{0}$ ,

$$\frac{\partial}{\partial x_\kappa} \sum_{\nu=0}^3 g_{\gamma\nu} R_{\alpha\beta}^{\nu\delta} + \frac{\partial}{\partial x_\beta} \sum_{\nu=0}^3 g_{\gamma\nu} R_{\alpha\delta}^{\nu\kappa} + \frac{\partial}{\partial x_\delta} \sum_{\nu=0}^3 g_{\gamma\nu} R_{\alpha\kappa}^{\nu\beta} = 0.$$

With (2.208) we obtain the following modified Bianchi identity:

$$\frac{\partial}{\partial x_\kappa} \check{R}_{\alpha\beta}^{\gamma\delta} + \frac{\partial}{\partial x_\beta} \check{R}_{\alpha\delta}^{\gamma\kappa} + \frac{\partial}{\partial x_\delta} \check{R}_{\alpha\kappa}^{\gamma\beta} = 0. \quad (2.213)$$

For the third term we can, with respect to (2.172), also write

$$-\frac{\partial}{\partial x_\delta} \check{R}_{\alpha\beta}^{\gamma\kappa}.$$

Substituting now  $\alpha = \beta$  and summing over  $\alpha$ , we obtain

$$\frac{\partial}{\partial x_\kappa} \check{R}_{\text{Ric},\gamma\delta} + \sum_{\alpha=0}^3 \frac{\partial}{\partial x_\alpha} \check{R}_{\alpha\delta}^{\gamma\kappa} - \frac{\partial}{\partial x_\delta} \check{R}_{\text{Ric},\gamma\kappa} = 0. \quad (2.214)$$

In the second term, one can, based on (2.171), replace  $\check{R}_{\alpha\delta}^{\gamma\kappa}$  by  $-\check{R}_{\gamma\delta}^{\alpha\kappa}$ . If we set  $\gamma = \delta$  and sum over  $\gamma$ , we obtain from (2.214) with the ‘‘curvature scalar’’  $\check{R} \stackrel{\text{def}}{=} \sum_{\gamma=0}^3 \check{R}_{\text{Ric},\gamma\gamma}$ , the trace of the Ricci matrix  $\check{\mathbf{R}}_{\text{Ric}}$ :

$$\frac{\partial}{\partial x_\kappa} \check{R} - \sum_{\alpha=0}^3 \frac{\partial}{\partial x_\alpha} \check{R}_{\text{Ric},\alpha\kappa} - \sum_{\gamma=0}^3 \frac{\partial}{\partial x_\gamma} \check{R}_{\text{Ric},\gamma\kappa} = 0. \quad (2.215)$$

If in the last sum the summation index  $\gamma$  is replaced by  $\alpha$ , we can finally summarize

$$\frac{\partial}{\partial x_\kappa} \check{R} - 2 \sum_{\alpha=0}^3 \frac{\partial}{\partial x_\alpha} \check{R}_{\text{Ric},\alpha\kappa} = 0. \quad (2.216)$$

We would get the same result, when this equation was assumed:

$$\frac{\partial}{\partial x_\kappa} \check{R}_{\gamma\delta}^{\alpha\beta} - 2 \frac{\partial}{\partial x_\beta} \check{R}_{\gamma\delta}^{\alpha\kappa} = 0. \quad (2.217)$$

Indeed, setting  $\delta = \gamma$  and summing over  $\gamma$ , we receive first:

$$\frac{\partial}{\partial x_\kappa} \check{R}_{\text{Ric},\alpha\beta} - 2 \frac{\partial}{\partial x_\beta} \check{R}_{\text{Ric},\alpha\kappa} = 0.$$

Substituting now  $\alpha = \beta$  and summing over  $\alpha$ , we receive again (2.216).

A different result is obtained when starting at first from (2.217) (with  $\nu$  instead of  $\alpha$ ), multiplying this equation by  $g_{\alpha\nu}^{[-1]}$ ,

$$\frac{\partial}{\partial x_\kappa} g_{\alpha\nu}^{[-1]} \check{R}_{\gamma\delta}^{\nu\beta} - 2 \frac{\partial}{\partial x_\beta} g_{\alpha\nu}^{[-1]} \check{R}_{\gamma\delta}^{\nu\kappa} = 0,$$

then setting  $\gamma = \delta$ , summing over  $\gamma$  and  $\nu$ , and using (2.214):

$$\begin{aligned} & \sum_\gamma \sum_\nu \frac{\partial}{\partial x_\kappa} g_{\alpha\nu}^{[-1]} \check{R}_{\gamma\delta}^{\nu\beta} - 2 \sum_\gamma \sum_\nu \frac{\partial}{\partial x_\beta} g_{\alpha\nu}^{[-1]} \check{R}_{\gamma\delta}^{\nu\kappa} \\ &= \frac{\partial}{\partial x_\kappa} R_{\text{Ric},\alpha\beta} - 2 \frac{\partial}{\partial x_\beta} R_{\text{Ric},\alpha\kappa} = 0. \end{aligned}$$

If we now set  $\alpha = \beta$  and sum over  $\alpha$ , we finally obtain the important relationship

$$\frac{\partial}{\partial x_\kappa} R - 2 \sum_\alpha \frac{\partial}{\partial x_\alpha} R_{\text{Ric},\alpha\kappa} = 0. \quad (2.218)$$

These are the four equations for the four spacetime coordinates  $x_0, \dots, x_3$ . Finally, the overall result can be represented as

$$\vec{\nabla}^\top \left( R_{\text{Ric}} - \frac{1}{2} R I_4 \right) = \mathbf{0}^\top. \quad (2.219)$$

## 2.11 General Theory of Gravitation

### 2.11.1 The Einstein's Matrix $\mathfrak{G}$

With the Einstein's matrix

$$\mathfrak{G} \stackrel{\text{def}}{=} R_{\text{Ric}} - \frac{1}{2} R I_4, \quad (2.220)$$

taking into account that the matrix  $\mathfrak{G}$  is symmetric, (2.219) can be restated as

$$\vec{\nabla}^\top \mathfrak{G} = \mathbf{0}^\top. \quad (2.221)$$

The following is a very important property of the Einstein's matrix:

**The divergence of the Einstein's matrix  $\mathfrak{G}$  vanishes!**

### 2.11.2 Newton's Theory of Gravity

According to Newton, for the acceleration the following is valid:

$$\frac{d^2 \mathbf{x}}{dt^2} = -\nabla \phi(\mathbf{x}), \quad (2.222)$$

where  $\phi(\mathbf{x})$  is the gravitational potential and  $\mathbf{x} \in \mathbb{R}^3$ . One can also write

$$\frac{d^2 \mathbf{x}}{dt^2} + \nabla \phi(\mathbf{x}) = \mathbf{0}. \quad (2.223)$$

The Newtonian universal time is a parameter which has two degrees of freedom, namely the time origin  $t_0$  and  $a$ , the unit of time, both can be chosen arbitrarily:  $t = t_0 + a\tau$ . Thus we obtain

$$\frac{d^2 t}{d\tau^2} = 0, \quad \frac{d^2 \mathbf{x}}{d\tau^2} + \frac{\partial \phi}{\partial \mathbf{x}} \left( \frac{dt}{d\tau} \right)^2 = \mathbf{0}. \quad (2.224)$$

This can also be written with the spacetime vector  $\vec{\mathbf{x}} = \begin{pmatrix} ct \\ \mathbf{x} \end{pmatrix} \in \mathbb{R}^4$ :

$$\frac{d^2 \vec{\mathbf{x}}}{d\tau^2} + (\mathbf{I}_4 \otimes \dot{\vec{\mathbf{x}}}^\top) \boldsymbol{\Gamma} \dot{\vec{\mathbf{x}}} = \mathbf{0}. \quad (2.225)$$

Here  $\boldsymbol{\Gamma} \in \mathbb{R}^{16 \times 4}$  has the form

$$\boldsymbol{\Gamma} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline \frac{\partial \phi}{\partial x_1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline \frac{\partial \phi}{\partial x_2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline \frac{\partial \phi}{\partial x_3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Now, how can one get the expression of the Poisson equation

$$\Delta \phi(\mathbf{x}) = \frac{\partial^2 \phi}{\partial x_1^2} + \frac{\partial^2 \phi}{\partial x_2^2} + \frac{\partial^2 \phi}{\partial x_3^2} = 4\pi G \rho(\mathbf{x})? \quad (2.226)$$





The  $4 \times 4$ -sub-matrices on the main diagonal of the curvature matrix  $\mathbf{R}$  contain exactly the components of the left-hand side of the Poisson equation. This subsequently provides the motivation for the introduction of the Ricci matrix. Here it is specifically equal to

$$\mathbf{R}_{\text{Ric}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -\frac{\partial^2 \phi}{\partial x_1^2} & -\frac{\partial^2 \phi}{\partial x_1 \partial x_2} & -\frac{\partial^2 \phi}{\partial x_1 \partial x_3} \\ 0 & -\frac{\partial^2 \phi}{\partial x_2 \partial x_1} & -\frac{\partial^2 \phi}{\partial x_2^2} & -\frac{\partial^2 \phi}{\partial x_2 \partial x_3} \\ 0 & -\frac{\partial^2 \phi}{\partial x_3 \partial x_1} & -\frac{\partial^2 \phi}{\partial x_3 \partial x_2} & -\frac{\partial^2 \phi}{\partial x_3^2} \end{pmatrix}. \quad (2.228)$$

If we form the trace of the Ricci matrix, then

$$R = \text{trace}(\mathbf{R}_{\text{Ric}}) = -\left(\frac{\partial^2 \phi}{\partial x_1^2} + \frac{\partial^2 \phi}{\partial x_2^2} + \frac{\partial^2 \phi}{\partial x_3^2}\right), \quad (2.229)$$

so that finally the Poisson equation can also be succinctly written as

$$\boxed{-R = 4\pi G\rho}. \quad (2.230)$$

The relationship in Newtonian mechanics between the gravitational potential  $\phi$  and matter is the Poisson equation

$$\Delta\phi = 4\pi G\rho,$$

where  $\rho$  is the mass density and  $G$  the Newtonian Gravitational Constant. In the theory of General Relativity, it is now required to set up general invariant equations of gravitation between the  $g_{ik}$ 's and matter. It is useful to characterise the matter by the energy–momentum matrix  $\mathbf{T} = \mathbf{T}_{\text{total}}$ .

For the rotating system in Sect. 2.5, we had  $\ddot{r} = r(\omega + \dot{\phi})^2$  so that there the potential was  $\phi = -\frac{\omega^2 r^2}{2}$  and  $g_{00} = 1 - \frac{r^2 \omega^2}{c^2} = 1 + \frac{2\phi}{c^2}$ . On the other hand, in Newtonian non-relativistic limit  $\frac{v^2}{c^2} \ll 1$ , we obtained for the energy–momentum matrix (see Sect. 1.9.2)  $T_{00} = c^2 \rho_0$  and for the remaining  $T_{ij} \approx 0$ .

With  $\phi = \frac{c^2}{2}(g_{00} - 1)$  we obtain  $\Delta\phi = \frac{c^2}{2}\Delta g_{00}$ . So one can write for the above Poisson equation

$$\Delta g_{00} = \frac{8\pi G}{c^4} T_{00}. \quad (2.231)$$

### 2.11.3 The Einstein's Equation with $\mathfrak{G}$

If one assumes that in the general case, i.e. in the presence of gravitational fields on the right-hand side of (2.231), the symmetric energy–momentum matrix is  $\mathbf{T}$ , a matrix containing the second partial derivatives of the elements of the metric matrix  $\mathbf{G}$

must appear on the left-hand side of the equation. If therefore the Einstein's matrix  $\mathfrak{G}$  is taken, one obtains the Einstein's Field Equation

$$\mathfrak{G} = -\frac{8\pi G}{c^4}T. \quad (2.232)$$

Because the matrix  $T \in \mathbb{R}^{4 \times 4}$  is symmetric, also the Einstein's matrix

$$\mathfrak{G} = R_{\text{Ric}} - \frac{1}{2}RI_4 \quad (2.233)$$

must be symmetric. This is indeed the case because both the Ricci matrix  $R_{\text{Ric}}$  and the diagonal matrix  $RI_4$  are symmetric.  $R_{\text{Ric}}$  is extracted, according to (2.200), from the Riemannian curvature matrix

$$R = \left( \frac{\partial \Gamma}{\partial x^\top} + (\bar{\Gamma} \otimes I_4)(I_4 \otimes \Gamma) \right) (U_{4 \times 4} - I_{16}). \quad (2.234)$$

**Final Form of the Einstein's Equation** The energy–momentum matrix  $T$  on the right-hand side of the Einstein's field equation (2.232) has the property that  $T\vec{\nabla} = \mathbf{0}$ , if one considers a closed system, i.e. if there are no external forces acting. The same must apply on the left:  $\mathfrak{G}\vec{\nabla} = \mathbf{0}$ . But that is, due to (2.221), true for the symmetric matrix  $\mathfrak{G}$ !

Overall, the final Einstein's Field Equation is fixed as an axiom as

$$\mathfrak{G} = R_{\text{Ric}} - \frac{R}{2}I_4 = -\frac{8\pi G}{c^4}T. \quad (2.235)$$

This is a matrix differential equation for determining the metric matrix  $G$ . This is no longer an “action-at-a-distance-equation” as Newton's, but describes the relationships at a spacetime point  $\vec{x}$ !  $R_{\text{Ric}}$  depends linearly on the second order derivatives of  $g_{ik}$ , and depends nonlinearly on the  $g_{ik}$  directly. So the Einstein's equation is a coupled system of second-order nonlinear partial differential equations for the determination of the components  $g_{ik}$  of the metric matrix  $G$  as a function of the given distribution of matter, i.e. the matrix  $T$ , the source of the gravitational field.

By taking the trace, from (2.235) it follows that

$$R - \frac{R}{2} \cdot 4 = -\frac{8\pi G}{c^4}T,$$

so

$$R = \frac{8\pi G}{c^4}T, \quad (2.236)$$

where

$$T \stackrel{\text{def}}{=} \text{trace}(T).$$

Substituting (2.236) into (2.235), one obtains the following form of the Einstein's Field Equation:

$$\mathbf{R}_{\text{Ric}} = \frac{8\pi G}{c^4} \left( \frac{T}{2} \mathbf{I}_4 - \mathbf{T} \right). \tag{2.237}$$

By the way, the constant factor  $\frac{8\pi G}{c^4}$  has the numerical value

$$\frac{8\pi G}{c^4} = 1.86 \cdot 10^{-27} \text{ cm/g}. \tag{2.238}$$

## 2.12 Summary

### 2.12.1 Covariance Principle

Einstein postulated the *Equivalence Principle*:

**Gravitational forces are equivalent to inertial forces.**

Gravitational fields can be eliminated by transitioning to an accelerated coordinate system. In this new local inertial frame, the laws of the theory of special relativity are valid.

Directly from the equivalence principle follows the *Covariance Principle*:

**The equations of physics must be invariant under general coordinate transformations.**

In particular, this means that even in a local inertial frame the equations must be valid, so the transition from a metric matrix  $\mathbf{G}$  to a Minkowski matrix  $\mathbf{M}$  results in the laws of special relativity.

The connection is  $\mathbf{G} = \mathbf{J}^\top \mathbf{M} \mathbf{J}$  and  $\mathbf{M} = \mathbf{J}^{-1\top} \mathbf{G} \mathbf{J}^{-1}$ , i.e. using the special transformation matrix  $\mathbf{T}(\vec{x}) = \mathbf{J}^{-1}(\vec{x})$  leads for a particular event  $\vec{x}$  to a local inertial frame in which the laws of special relativity apply. Conversely, one passes from the special local inertial frame with the event  $\vec{\xi}$  to the same event in the general coordinate system  $\vec{x}$  by the transformation  $\vec{x} = \mathbf{J}^{-1}\vec{\xi}$ .

The physical equations in the theory of general relativity must be formulated so that they are invariant (covariant) with respect to general coordinate transformations.

Above we derived: The formulas in the theory of general relativity are invariant under coordinate transformations if one replaces in the formulas of special relativity the ordinary derivatives  $\frac{\partial a}{\partial x^\tau}$  by the covariant derivatives  $a_{\parallel x^\tau}$ . Then we're done! The law applies generally.



For a particle on which no force acts, for example, in an inertial system with  $\dot{\vec{\xi}} \stackrel{\text{def}}{=} \frac{d\vec{\xi}}{d\tau}$ , applies

$$\frac{d\dot{\vec{\xi}}}{d\tau} = \mathbf{0}. \quad (2.239)$$

Herein the ordinary differential  $d\dot{\vec{\xi}}$  must be replaced by the covariant differential  $D\vec{u}$ , with  $\vec{u} = \dot{\vec{x}} \stackrel{\text{def}}{=} \mathbf{J}^{-1}\dot{\vec{\xi}}$ . According to (2.145), we first obtain

$$D\vec{u} = \vec{u}_{\parallel\vec{x}^\top} d\vec{x} = \frac{\partial\vec{u}}{\partial\vec{x}^\top} d\vec{x} + (\mathbf{I}_4 \otimes \vec{u}^\top) \mathbf{\Gamma} d\vec{x}, \quad (2.240)$$

so

$$\frac{D\vec{u}}{d\tau} = \vec{u}_{\parallel\vec{x}^\top} \frac{d\vec{x}}{d\tau} = \frac{d\vec{u}}{d\tau} + (\mathbf{I}_4 \otimes \vec{u}^\top) \mathbf{\Gamma} \frac{d\vec{x}}{d\tau}.$$

This used in (2.239) instead of  $\frac{d\vec{u}}{d\tau}$  generally yields

$$\boxed{\frac{D\vec{u}}{d\tau} = \mathbf{0}}, \quad (2.241)$$

or

$$\underline{\underline{\frac{d\vec{u}}{d\tau} = -(\mathbf{I}_4 \otimes \vec{u}^\top) \mathbf{\Gamma} \vec{u}}}. \quad (2.242)$$

Through the Christoffel matrix  $\mathbf{\Gamma}$ , the effect of the gravitational field is expressed. If no gravitational field is present, then  $\mathbf{\Gamma} = \mathbf{0}$  and (2.239) is again obtained. A comparison of (2.242) with the equation of a geodesic shows that a material particle moves on a geodesic, so to speak, “on the shortest path in curved space”.

If, in addition to gravitational forces, other forces appear, e.g. exerted by electric fields, then for an inertial frame one has the equation

$$m \frac{d\dot{\vec{\xi}}}{d\tau} = \vec{f}. \quad (2.243)$$

This equation multiplied from the left with  $\mathbf{J}^{-1}$  and again used with the covariant derivative leads with

$$\vec{f}_x \stackrel{\text{def}}{=} \mathbf{J}^{-1} \vec{f}$$

to

$$\boxed{m \frac{D\vec{u}}{d\tau} = \vec{f}_x}, \quad (2.244)$$

or

$$\underline{\underline{m \frac{d\vec{u}}{d\tau} = \vec{f}_x - m(\mathbf{I}_4 \otimes \vec{u}^\top) \mathbf{\Gamma} \vec{u}}}, \quad (2.245)$$

where on the right-hand side, next to the other forces  $\vec{f}_x$ , the gravitational forces are written.

### 2.12.2 Einstein's Field Equation and Momentum

The Einstein's field equation expresses that each form of matter and energy is a source of the gravitational field. The gravitational field is described by the metric matrix  $\mathbf{G}$ , the components of which must be determined by using Einstein's Field Equation

$$\mathbf{R}_{\text{Ric}} - \frac{R}{2} \mathbf{I}_4 = -\frac{8\pi G}{c^4} \mathbf{T}. \quad (2.246)$$

So one basically has to perform the same procedure as in Newton's dynamics:

1. Solve the Poisson equation  $\Delta\phi(\mathbf{x}) = 4\pi G\rho(\mathbf{x})$ , to determine the potential function  $\phi$ .
2. Establish the solution of the equation  $\frac{d^2\mathbf{x}}{dt^2} = -\nabla\phi(\mathbf{x})$ .

In the theory of general relativity, we now arrive at the following procedure:

1. Solve Einstein's Field Equation (2.246)  $\mathbf{R}_{\text{Ric}} - \frac{R}{2} \mathbf{I}_4 = -\frac{8\pi G}{c^4} \mathbf{T}$ , to determine the metric matrix  $\mathbf{G}$ .
2. Establish the solution of (2.245)

$$m \frac{d\vec{u}}{d\tau} = \vec{f}_x - m(\mathbf{I}_4 \otimes \vec{u}^\top) \Gamma \vec{u}.$$

## 2.13 Hilbert's Action Functional

Above, the Einstein's equation was postulated as an axiom. Einstein has found this equation through years of hard work. As a supplement, we will now shown that one can obtain the Einstein's equation with the help of the calculus of variations. Following Hilbert, this is derived from a variational principle, but initially only for the free gravitational field, i.e. when  $\mathbf{T} = \mathbf{0}$ .

The gravitational field is determined primarily by the metric matrix  $\mathbf{G}$ , i.e. the curvature of space caused by this field. All curvature parameters are concentrated in the curvature scalar  $R$  which is formed through the trace of the Ricci matrix  $\mathbf{R}_{\text{Ric}}$ . Now a variational functional, the Hilbert's action functional, is formed so that the space curvature is minimal:

$$W_{\text{Grav}} = \int R \, dV. \quad (2.247)$$

However, this integral is not invariant under coordinate transformations  $\Theta$ . Therefore, we need the following considerations: Let

$$\Theta^\top(\vec{x})\mathbf{G}(\vec{x})\Theta(\vec{x}) = \mathbf{M}, \quad (2.248)$$

i.e.  $\Theta(\vec{x})$  is the matrix which at the point  $\vec{x}$  transforms the metric matrix  $\mathbf{G}$  into the Minkowski matrix  $\mathbf{M}$ . Computing the determinants on both sides of (2.248), we obtain

$$\underbrace{\det(\mathbf{G})}_g \underbrace{\det(\Theta)}_{\Theta^2} = \det(\mathbf{M}) = -1, \quad (2.249)$$

i.e.

$$\sqrt{-g} = \frac{1}{\Theta}.$$

In a Cartesian coordinate system, the integral of a scalar with respect to the scalar  $dV = dx_0 \cdot dx_1 \cdot dx_2 \cdot dx_3$  is also a scalar. When passing to curvilinear coordinates  $\vec{x}'$ , the integration element  $dV$  is

$$\frac{1}{\Theta} dV' = \sqrt{-g'} dV'.$$

In curvilinear coordinates,  $\sqrt{-g} dV$  behaves as an invariant when integrated over any area of the four-dimensional space. If  $f$  is a scalar, then  $f\sqrt{-g}$ , which is invariant when integrated over  $dV$ , is named the *scalar density*. This quantity provides a scalar for its multiplication with the four-dimensional volume element  $dV$ .

For this reason, we now consider only the effect

$$W_{\text{Grav}} = \int R\left(\Gamma(\mathbf{x}), \frac{\partial \Gamma}{\partial \mathbf{x}^\top}\right) \sqrt{-g} d^4 \vec{x}. \quad (2.250)$$

Under  $d^4 \vec{x}$  the four-dimensional volume element  $dx_0 \cdot dx_1 \cdot dx_2 \cdot dx_3$  is understood. The Einstein's equation will follow from (2.250) and the condition  $\delta W_{\text{Grav}} = 0$  for any variation  $\delta g_{ik}$ .  $R\sqrt{-g}$  is a Lagrange density which is integrated over the volume. The elements of  $\Gamma_{\alpha\beta}^k$  of the Christoffel matrix  $\Gamma$  are (2.62)

$$\Gamma_{\alpha\beta}^k = \sum_{i=0}^3 \frac{g_{ki}^{[-1]}}{2} \left( \frac{\partial g_{\beta i}}{\partial x_\alpha} + \frac{\partial g_{\alpha i}}{\partial x_\beta} - \frac{\partial g_{\alpha\beta}}{\partial x_i} \right) \quad (2.251)$$

and the elements of the Riemannian curvature matrix  $\mathbf{R}$  are, according to (2.165),

$$R_{\alpha\beta}^{\gamma\delta} = \frac{\partial}{\partial x_\beta} \Gamma_{\alpha\delta}^\gamma - \frac{\partial}{\partial x_\delta} \Gamma_{\alpha\beta}^\gamma + \sum_\nu \Gamma_{\beta\nu}^\gamma \Gamma_{\delta\alpha}^\nu - \sum_\nu \Gamma_{\delta\nu}^\gamma \Gamma_{\alpha\beta}^\nu. \quad (2.252)$$

The curvature scalar  $R$ , due to (2.207), is obtained from the Ricci matrix by

$$\begin{aligned} R &= \text{trace}(\mathbf{R}_{\text{Ric}}) = \sum_{\alpha} \sum_{\nu} R_{\nu\nu}^{\alpha\alpha} \\ &= \sum_{\alpha} \sum_{\nu} \sum_{\mu} g_{\alpha\mu}^{[-1]} \check{R}_{\nu\nu}^{\mu\alpha} = \sum_{\alpha} \sum_{\mu} g_{\alpha\mu}^{[-1]} \check{R}_{\text{Ric},\mu\alpha}, \end{aligned} \quad (2.253)$$

which, in accordance with (2.210), is

$$\check{R}_{\text{Ric},\gamma\delta} = \sum_{\nu=0}^3 \left( \frac{\partial}{\partial x_{\delta}} \check{r}_{\gamma\nu}^{\nu} - \frac{\partial}{\partial x_{\nu}} \check{r}_{\gamma\delta}^{\nu} + \sum_{\mu=0}^3 \Gamma_{\gamma\delta}^{\mu} \check{r}_{\nu\nu}^{\mu} - \sum_{\mu=0}^3 \Gamma_{\gamma\nu}^{\mu} \check{r}_{\nu\delta}^{\mu} \right). \quad (2.254)$$

The Lagrange/Hamilton theory applied to the action integral (2.250) provides the Euler/Lagrange equations for the variational problem associated to (2.250). We consider the elements of  $g_{ki}^{[-1]}$  and  $\Gamma_{\alpha\beta}^k$  as independent functions  $f_i(\vec{x})$ . The integral therefore is a functional of the form

$$\int L \left( f_i(\vec{x}), \frac{\partial f_i}{\partial x_k}(\vec{x}) \right) d^4 \vec{x}. \quad (2.255)$$

This yields the Euler/Lagrange equations

$$\frac{\partial L}{\partial f_i} = \sum_{k=0}^3 \frac{\partial}{\partial x_k} \frac{\partial L}{\partial \left( \frac{\partial f_i}{\partial x_k} \right)}. \quad (2.256)$$

The Euler/Lagrange equations for (2.250) are

$$\frac{\partial}{\partial g_{\alpha\beta}^{[-1]}} \left( \sqrt{-g} \sum_{\delta} \sum_{\mu} g_{\mu\delta}^{[-1]} \check{R}_{\text{Ric},\mu\delta} \right) = 0, \quad (2.257)$$

$$\frac{\partial}{\partial \Gamma_{\alpha\beta}^{\gamma}} \left( \sqrt{-g} \sum_{\delta} \sum_{\mu} g_{\mu\delta}^{[-1]} \check{R}_{\text{Ric},\mu\delta} \right) = \sum_{\delta=0}^3 \frac{\partial}{\partial x_{\delta}} \left( \sqrt{-g} \sum_{\delta} \sum_{\mu} g_{\mu\delta}^{[-1]} \frac{\partial \check{R}_{\text{Ric},\mu\delta}}{\partial \left( \frac{\partial \Gamma_{\alpha\beta}^{\gamma}}{\partial x_{\delta}} \right)} \right). \quad (2.258)$$

The first equation (2.257) provides the Einstein's field equation. Indeed, one can first write this compactly:

$$\frac{\partial L}{\partial \mathbf{G}^{-1}} = \frac{\partial(\sqrt{-g}R)}{\partial \mathbf{G}^{-1}} = \frac{\partial\sqrt{-g}}{\partial \mathbf{G}^{-1}} R + \sqrt{-g} \frac{\partial R}{\partial \mathbf{G}^{-1}} = 0. \quad (2.259)$$

For  $\frac{\partial\sqrt{-g}}{\partial \mathbf{G}^{-1}}$ , one then obtains

$$\frac{\partial\sqrt{-g}}{\partial \mathbf{G}^{-1}} = \frac{-1}{2\sqrt{-g}} \cdot \frac{\partial g}{\partial \mathbf{G}^{-1}}. \quad (2.260)$$

In addition,

$$\frac{\partial}{\partial \mathbf{G}^{-1}} \left( \frac{1}{g} \cdot g \right) = \mathbf{0} = \frac{1}{g} \frac{\partial g}{\partial \mathbf{G}^{-1}} + g \frac{\partial(1/g)}{\partial \mathbf{G}^{-1}}. \quad (2.261)$$

According to the Laplace expansion theorem for determinants (“The sum of the products of all elements of a row with their cofactors is equal to the value of the determinant”), we obtain by developing the determinant of  $\mathbf{G}^{-1}$  along the  $\gamma$ th row:

$$\det(\mathbf{G}^{-1}) = \frac{1}{g} = g_{\gamma 0}^{[-1]} A_{\gamma 0}^{[-1]} + \dots + g_{\gamma \beta}^{[-1]} A_{\gamma \beta}^{[-1]} + \dots + g_{\gamma 3}^{[-1]} A_{\gamma 3}^{[-1]}, \quad (2.262)$$

where  $A_{\gamma\beta}^{[-1]}$  is the element in the  $\gamma$ th row and  $\beta$ th column of the adjoint of  $\mathbf{G}^{-1}$ ;  $g_{\gamma\beta}^{[-1]}$  is the  $(\gamma\beta)$ -element of  $\mathbf{G}^{-1}$ . It is, of course, true that

$$\mathbf{G} = \frac{\text{adj}(\mathbf{G}^{-1})}{\det(\mathbf{G}^{-1})}$$

so element-wise  $g_{\beta\gamma} = g \cdot A_{\gamma\beta}^{[-1]}$ , or  $A_{\gamma\beta}^{[-1]} = \frac{1}{g} g_{\beta\gamma}$ . Thus, we obtain by partial differentiation of (2.262) with respect to  $g_{\gamma\beta}^{[-1]}$

$$\frac{\partial(1/g)}{\partial g_{\gamma\beta}^{[-1]}} = A_{\gamma\beta}^{[-1]} = \frac{1}{g} g_{\beta\gamma},$$

so

$$\frac{\partial(1/g)}{\partial \mathbf{G}^{-1}} = \frac{1}{g} \mathbf{G}. \quad (2.263)$$

Using this in (2.261) provides

$$\frac{\partial g}{\partial \mathbf{G}^{-1}} = -g \mathbf{G}, \quad (2.264)$$

and, together with (2.260), finally yields

$$\frac{\partial \sqrt{-g}}{\partial \mathbf{G}^{-1}} = \frac{-1}{2} \sqrt{-g} \cdot \mathbf{G}. \quad (2.265)$$

All what is missing in (2.259) is  $\frac{\partial R}{\partial \mathbf{G}^{-1}}$ . Using  $R = \sum_{\alpha} \sum_{\mu} g_{\alpha\mu}^{[-1]} \check{R}_{\text{Ric},\mu\alpha}$ ,

$$\frac{\partial R}{\partial \mathbf{G}^{-1}} = \begin{pmatrix} \frac{\partial R}{\partial g_{00}^{[-1]}} & \cdots & \frac{\partial R}{\partial g_{03}^{[-1]}} \\ \vdots & \ddots & \vdots \\ \frac{\partial R}{\partial g_{30}^{[-1]}} & \cdots & \frac{\partial R}{\partial g_{33}^{[-1]}} \end{pmatrix} = \begin{pmatrix} \check{R}_{\text{Ric},00} & \cdots & \check{R}_{\text{Ric},03} \\ \vdots & \ddots & \vdots \\ \check{R}_{\text{Ric},30} & \cdots & \check{R}_{\text{Ric},33} \end{pmatrix} = \underline{\underline{\check{R}_{\text{Ric}}}}. \quad (2.266)$$

Multiplying this matrix with  $\mathbf{G}^{-1}$ , by the way, yields

$$\mathbf{R}_{\text{Ric}} = \bar{\mathbf{G}}^{-1} \check{\mathbf{R}}_{\text{Ric}}. \quad (2.267)$$

Inserting (2.265) and (2.266) into (2.259) yields first

$$\sqrt{-g} \left( \frac{-1}{2} \cdot \mathbf{G} \mathbf{R} + \check{\mathbf{R}}_{\text{Ric}} \right) = \mathbf{0}, \quad (2.268)$$

i.e. the following particular form of the Einstein's field equation

$$\check{\mathbf{R}}_{\text{Ric}} - \frac{R}{2} \cdot \mathbf{G} = \mathbf{0}. \quad (2.269)$$

Multiplying this equation from the left with the inverse metric matrix  $\mathbf{G}^{-1}$ , we finally obtain the Einstein's field equation for a source-free gravitational field ( $\mathbf{T} = \mathbf{0}$ ) as in (2.235)

$$\boxed{\mathbf{R}_{\text{Ric}} - \frac{R}{2} \mathbf{I}_4 = \mathbf{0}.} \quad (2.270)$$

### 2.13.1 Effects of Matter

So far, only the gravitational field in vacuum has been treated. To account for the sources of the gravitational field, e.g. matter, the action functional must include an additive term  $W_M$  that describes the source:

$$W \stackrel{\text{def}}{=} W_{\text{Grav}} + W_M = \int (kR + \mathcal{L}_M) \sqrt{-g} d^4 \vec{x}. \quad (2.271)$$

The Lagrange equation with respect to  $\mathbf{G}^{-1}$  is

$$\begin{aligned} \frac{\partial(\sqrt{-g}(kR + \mathcal{L}_M))}{\partial \mathbf{G}^{-1}} &= \mathbf{0} \\ &= k \left( \frac{\partial \sqrt{-g}}{\partial \mathbf{G}^{-1}} R + \sqrt{-g} \frac{\partial R}{\partial \mathbf{G}^{-1}} \right) + \frac{\partial(\sqrt{-g} \mathcal{L}_M)}{\partial \mathbf{G}^{-1}}. \end{aligned} \quad (2.272)$$

The term inside the large parentheses gives the left-hand side of (2.268), i.e. together we first get for (2.272)

$$\mathbf{0} = k \sqrt{-g} \left( \frac{-1}{2} \cdot \mathbf{G} \mathbf{R} + \check{\mathbf{R}}_{\text{Ric}} \right) + \sqrt{-g} \left( \frac{\partial \mathcal{L}_M}{\partial \mathbf{G}^{-1}} + \frac{\mathcal{L}_M}{\sqrt{-g}} \underbrace{\frac{\partial \sqrt{-g}}{\partial \mathbf{G}^{-1}}}_{-\frac{1}{2} \sqrt{-g} \mathbf{G}} \right).$$

Now, if we define the energy–momentum matrix as

$$-\frac{1}{2}\check{\mathbf{T}} \stackrel{\text{def}}{=} \frac{\partial \mathcal{L}_M}{\partial \mathbf{G}^{-1}} - \frac{1}{2}\mathcal{L}_M \mathbf{G}, \quad (2.273)$$

then we finally obtain with  $k = c^4/(16\pi G)$  by left multiplication of (2.273) by  $\mathbf{G}^{-1}$  and with  $\mathbf{T} \stackrel{\text{def}}{=} \mathbf{G}^{-1}\check{\mathbf{T}}$  the Einstein's field equation

$$\boxed{\mathbf{R}_{\text{Ric}} - \frac{R}{2}\mathbf{I}_4 = \frac{8\pi G}{c^4}\mathbf{T}.} \quad (2.274)$$

## 2.14 Most Important Definitions and Formulas

**Christoffel Matrix (2.68):**

$$\boxed{\mathbf{\Gamma} \stackrel{\text{def}}{=} \begin{pmatrix} \Gamma_0 \\ \vdots \\ \Gamma_3 \end{pmatrix} = \frac{1}{2}(\mathbf{G}^{-1} \otimes \mathbf{I}_4) \left[ \begin{pmatrix} \frac{\partial g_0^\top}{\partial \mathbf{x}} \\ \vdots \\ \frac{\partial g_3^\top}{\partial \mathbf{x}} \end{pmatrix} + \begin{pmatrix} \frac{\partial g_0}{\partial \mathbf{x}^\top} \\ \vdots \\ \frac{\partial g_3}{\partial \mathbf{x}^\top} \end{pmatrix} - \frac{\partial \mathbf{G}}{\partial \mathbf{x}} \right] \in \mathbb{R}^{16 \times 4}}$$

with the components given, in accordance with (2.63) and (2.64), as

$$\Gamma_{\alpha\beta}^k = \sum_{i=0}^3 g_{ki}^{[-1]} \frac{1}{2} \left( \frac{\partial g_{\beta i}}{\partial x_\alpha} + \frac{\partial g_{\alpha i}}{\partial x_\beta} - \frac{\partial g_{\alpha\beta}}{\partial x_i} \right).$$

**Motion in a Gravitational Field (2.67):**

$$\boxed{\ddot{\mathbf{x}} = -(\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \mathbf{\Gamma} \dot{\mathbf{x}}.}$$

**Riemannian Curvature Matrix (2.160):**

$$\boxed{\mathbf{R} \stackrel{\text{def}}{=} \left[ \frac{\partial \mathbf{\Gamma}}{\partial \mathbf{x}^\top} + (\overline{\mathbf{\Gamma}} \otimes \mathbf{I}_4)(\mathbf{I}_4 \otimes \mathbf{\Gamma}) \right] (\mathbf{U}_{4 \times 4} - \mathbf{I}_{16}) \in \mathbb{R}^{16 \times 16},}$$

with the components (2.165)

$$R_{\alpha\beta}^{\gamma\delta} = \frac{\partial}{\partial x_\beta} \Gamma_{\alpha\delta}^\gamma - \frac{\partial}{\partial x_\delta} \Gamma_{\alpha\beta}^\gamma + \sum_\nu \Gamma_{\beta\nu}^\gamma \Gamma_{\delta\alpha}^\nu - \sum_\nu \Gamma_{\delta\nu}^\gamma \Gamma_{\alpha\beta}^\nu.$$

**Ricci Matrix  $R_{\text{Ric}} \in \mathbb{R}^{4 \times 4}$  (2.200):**

$$R_{\text{Ric}, \gamma\delta} \stackrel{\text{def}}{=} \text{trace}(R^{\gamma\delta}) = \sum_{\nu=0}^3 R_{\nu\nu}^{\gamma\delta},$$

consisting of the traces of the sub-matrices  $R^{\gamma\delta}$  of  $R$  with the components in accordance with (2.209):

$$R_{\text{Ric}, \gamma\delta} = \sum_{\nu=0}^3 \left( \frac{\partial}{\partial x_\delta} \Gamma_{\gamma\nu}^\nu - \frac{\partial}{\partial x_\nu} \Gamma_{\gamma\delta}^\nu + \sum_{\mu=0}^3 \Gamma_{\delta\mu}^\nu \Gamma_{\nu\gamma}^\mu - \sum_{\mu=0}^3 \Gamma_{\nu\mu}^\mu \Gamma_{\gamma\delta}^\mu \right).$$

**Curvature Scalar  $R$  (2.200):**

$$R \stackrel{\text{def}}{=} \text{trace}(R_{\text{Ric}}) = \sum_{\alpha} \sum_{\nu} R_{\nu\nu}^{\alpha\alpha},$$

which is obtained by taking the trace of the Ricci matrix.

**Einstein's Field Equation (2.235) (fixed as an axiom):**

$$R_{\text{Ric}} - \frac{R}{2} I_4 = -\frac{8\pi G}{c^4} T,$$

or with  $T \stackrel{\text{def}}{=} \text{trace } T$ , see (2.237):

$$R_{\text{Ric}} = \frac{8\pi G}{c^4} \left( \frac{T}{2} I_4 - T \right).$$



# Chapter 3

## Gravitation of a Spherical Mass

The chapter begins with the first solution of Einstein’s equations obtained in 1916 by Schwarzschild for a spherical mass. From these results, the influence of a mass in time and space and the redshift of the spectral lines are given. This gives also a first indication of the existence of “black holes”. Finally, a brief look at the effect of rotating masses and on the Lense–Thirring effect is taken.

### 3.1 Schwarzschild’s Solution

Because of its simplicity, the solution of Einstein’s field equation will now be determined for the *outside* of a spherically symmetric, uniform, time-invariant mass distribution. This first exact solution of Einstein’s field equation was given in 1916 by Schwarzschild.

As  $r \rightarrow \infty$ , the desired metric should be the Minkowski metric, i.e.

$$ds^2 = c^2 dt^2 - dr^2 - r^2(d\theta^2 + \sin^2 \theta d\varphi^2). \tag{3.1}$$

where  $r, \theta$  and  $\varphi$  are the spherical coordinates. For a gravitational field, we now start with

$$ds^2 = A(r) dt^2 - B(r) dr^2 - r^2(d\theta^2 + \sin^2 \theta d\varphi^2). \tag{3.2}$$

Since the field must be spherically symmetric, the factors  $A$  and  $B$  may depend only on  $r$  and not on  $\theta$  or  $\varphi$ . Due to (3.1), the factor  $A(r)$  must tend to  $c^2$  as  $r \rightarrow \infty$ , and the factor  $B(r) \rightarrow 1$ . The metric matrix thus has the form

$$\mathbf{G} = \begin{pmatrix} A(r) & 0 & 0 & 0 \\ 0 & -B(r) & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2 \sin^2 \theta \end{pmatrix}. \tag{3.3}$$

### 3.1.1 Christoffel Matrix $\Gamma$

Due to (2.57)

$$\hat{\Gamma}_k \stackrel{\text{def}}{=} (\mathbf{I}_4 \otimes \mathbf{g}_k^{-T}) \left[ \mathbf{I}_{16} - \frac{1}{2} \mathbf{U}_{4 \times 4} \right] \frac{\partial \mathbf{G}}{\partial \bar{\mathbf{x}}}$$

and

$$\mathbf{G}^{-1} = \begin{pmatrix} \frac{1}{A(r)} & 0 & 0 & 0 \\ 0 & \frac{-1}{B(r)} & 0 & 0 \\ 0 & 0 & \frac{-1}{r^2} & 0 \\ 0 & 0 & 0 & \frac{-1}{r^2 \sin^2 \theta} \end{pmatrix}, \quad (3.4)$$

with

$$\frac{\partial \mathbf{G}}{\partial \bar{\mathbf{x}}} = \begin{pmatrix} \mathbf{0}_4 \\ \hline A'(r) & 0 & 0 & 0 \\ 0 & -B'(r) & 0 & 0 \\ 0 & 0 & -2r & 0 \\ 0 & 0 & 0 & -2r \sin^2 \theta \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2r^2 \sin \theta \cos \theta \\ \hline \mathbf{0}_4 \end{pmatrix},$$

in which  $A' \stackrel{\text{def}}{=} \frac{\partial A}{\partial r}$  and  $\mathbf{0}_4$  is a  $4 \times 4$ -zero-matrix, and

$$\left[ \mathbf{I}_{16} - \frac{1}{2} \mathbf{U}_{4 \times 4} \right] \frac{\partial \mathbf{G}}{\partial \bar{\mathbf{x}}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -\frac{A'}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline A' & 0 & 0 & 0 \\ 0 & -\frac{B'}{2} & 0 & 0 \\ 0 & 0 & -2r & 0 \\ 0 & 0 & 0 & -2r \sin^2 \theta \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & r & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2r^2 \sin \theta \cos \theta \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & r \sin^2 \theta \\ 0 & 0 & 0 & r^2 \sin \theta \cos \theta \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

one obtains

$$\begin{aligned} \hat{\mathbf{I}}_0 &= (\mathbf{I}_4 \otimes [1/A(r)|0|0|0]) \left[ \mathbf{I}_{16} - \frac{1}{2} \mathbf{U}_{4 \times 4} \right] \begin{pmatrix} \frac{\partial \mathbf{G}}{\partial t} \\ \frac{\partial \mathbf{G}}{\partial r} \\ \frac{\partial \mathbf{G}}{\partial \theta} \\ \frac{\partial \mathbf{G}}{\partial \varphi} \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ \frac{A'}{A} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned} \quad (3.5)$$

Accordingly, we obtain

$$\begin{aligned} \hat{\mathbf{I}}_1 &= \begin{pmatrix} \frac{A'}{2B} & 0 & 0 & 0 \\ 0 & \frac{B'}{2B} & 0 & 0 \\ 0 & 0 & -\frac{r}{B} & 0 \\ 0 & 0 & 0 & -\frac{r}{B} \sin^2 \theta \end{pmatrix}, \\ \hat{\mathbf{I}}_2 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{2}{r} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\sin \theta \cos \theta \end{pmatrix} \end{aligned}$$

and

$$\hat{\mathbf{I}}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{2}{r} \\ 0 & 0 & 0 & 2 \cot \theta \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Symmetrizing according to

$$\mathbf{\Gamma}_k = \frac{1}{2} (\hat{\mathbf{I}}_k + \hat{\mathbf{I}}_k^\top)$$

yields the symmetric matrices

$$\begin{aligned} \mathbf{\Gamma}_0 &= \begin{pmatrix} 0 & \frac{A'}{(2A)} & 0 & 0 \\ \frac{A'}{(2A)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\ \mathbf{\Gamma}_1 &= \hat{\mathbf{I}}_1, \\ \mathbf{\Gamma}_2 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{r} & 0 \\ 0 & \frac{1}{r} & 0 & 0 \\ 0 & 0 & 0 & -\sin \theta \cos \theta \end{pmatrix} \end{aligned}$$

and

$$\Gamma_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{r} \\ 0 & 0 & 0 & \cot\theta \\ 0 & \frac{1}{r} & \cot\theta & 0 \end{pmatrix},$$

which put altogether give

$$\Gamma = \begin{pmatrix} 0 & \frac{A'}{2A} & 0 & 0 \\ \frac{A'}{2A} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline \frac{A'}{2B} & 0 & 0 & 0 \\ 0 & \frac{B'}{2B} & 0 & 0 \\ 0 & 0 & -\frac{r}{B} & 0 \\ 0 & 0 & 0 & -\frac{r}{B} \sin^2\theta \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{r} & 0 \\ 0 & \frac{1}{r} & 0 & 0 \\ 0 & 0 & 0 & -\sin\theta \cos\theta \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{r} \\ 0 & 0 & 0 & \cot\theta \\ 0 & \frac{1}{r} & \cot\theta & 0 \end{pmatrix}. \quad (3.6)$$

The nonzero Christoffel elements can be read from the matrices as:

$$\begin{aligned} \Gamma_{10}^0 = \Gamma_{01}^0 = \frac{A'}{2A}, \quad \Gamma_{00}^1 = \frac{A'}{2B}, \quad \Gamma_{11}^1 = \frac{B'}{2B}, \quad \Gamma_{22}^1 = -\frac{r}{B}, \\ \Gamma_{33}^1 = -\frac{r}{B} \sin^2\theta, \quad \Gamma_{12}^2 = \Gamma_{21}^2 = \frac{1}{r}, \quad \Gamma_{33}^2 = -\sin\theta \cos\theta, \\ \Gamma_{13}^3 = \Gamma_{31}^3 = \frac{1}{r} \quad \text{and} \quad \Gamma_{23}^3 = \Gamma_{32}^3 = \cot\theta. \end{aligned}$$

### 3.1.2 Ricci Matrix $R_{\text{Ric}}$

The Ricci matrix  $R_{\text{Ric}}$  is obtained from the Riemannian curvature matrix

$$\mathbf{R} = \left[ \frac{\partial \Gamma}{\partial x^\tau} + (\bar{\Gamma} \otimes I_4)(I_4 \otimes \Gamma) \right] (U_{4 \times 4} - I_{16})$$

by summing the  $4 \times 4$ -sub-matrices on the main diagonal. For the calculation of  $\mathbf{R}$ , in addition to  $\mathbf{\Gamma}$ , we need

$$\frac{\partial \mathbf{\Gamma}}{\partial \bar{x}^\top} = \begin{pmatrix} \mathbf{0}_4 & \begin{array}{ccc|ccc} 0 & \frac{A''A-A'^2}{2A^2} & 0 & 0 & 0 & 0 \\ \frac{A''A-A'^2}{2A^2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} & \mathbf{0}_4 & \mathbf{0}_4 \\ \mathbf{0}_4 & \begin{array}{ccc|ccc} \frac{A''B-A'B'}{2B^2} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{B''B-B'^2}{2B^2} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{B-rB'}{B^2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\sin^2 \theta \frac{B-rB'}{B^2} & 0 & 0 \end{array} & \mathbf{0}_{3 \times 4} & \mathbf{0}_4 \\ \mathbf{0}_4 & \begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{r^2} & 0 & 0 & 0 \\ 0 & -\frac{1}{r^2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} & \mathbf{0}_{3 \times 4} & \mathbf{0}_4 \\ \mathbf{0}_4 & \begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} & \begin{array}{ccc|ccc} 0 & 0 & 0 & -\cos^2 \theta + \sin^2 \theta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} & \mathbf{0}_4 \end{pmatrix}$$

and

$$\bar{\mathbf{\Gamma}} = \begin{pmatrix} \begin{array}{ccc|ccc} 0 & \frac{A'}{2A} & 0 & 0 & \frac{A'}{2A} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{A'}{2B} & 0 & 0 & 0 & 0 & \frac{B'}{2B} & 0 & 0 & 0 & -\frac{r}{B} & 0 & 0 & 0 & 0 & -\frac{r}{B} \sin^2 \theta \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{r} & 0 & 0 & \frac{1}{r} & 0 & 0 & 0 & 0 & -\sin \theta \cos \theta \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{r} & 0 & 0 & 0 & \cot \theta & 0 & \frac{1}{r} & \cot \theta & 0 \end{array} \end{pmatrix}.$$

For the Ricci matrix  $\mathbf{R}_{\text{Ric}}$  the  $16 \times 16$ -matrix  $\frac{\partial \mathbf{\Gamma}}{\partial \bar{x}^\top}(\mathbf{U}_{4 \times 4} - \mathbf{I}_{16})$  contributes the term

$$\begin{pmatrix} \begin{array}{ccc|ccc} \frac{A'B'-A''B}{2B^2} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{A''A-A'^2}{2A^2} - \frac{2}{r^2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{B-rB'}{B^2} - \frac{1}{\sin^2 \theta} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sin^2 \theta \frac{B-rB'}{B^2} + \cos^2 \theta - \sin^2 \theta & 0 & 0 \end{array} \end{pmatrix}$$

and the matrix  $(\bar{\mathbf{\Gamma}} \otimes \mathbf{I}_4)(\mathbf{I}_4 \otimes \mathbf{\Gamma})(\mathbf{U}_{4 \times 4} - \mathbf{I}_{16})$  the term

$$\begin{pmatrix} \begin{array}{ccc|ccc} \frac{A^2}{4AB} - \frac{A'B'}{4B^2} - \frac{A'}{rB} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{A^2}{4A^2} - \frac{A'B'}{4AB} - \frac{B'}{rB} + \frac{2}{r^2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{rB'}{B^2} + \frac{rA'}{2AB} + \cot^2 \theta & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{rB'}{B^2} \sin^2 \theta + \frac{rA'}{2AB} \sin^2 \theta - \cos^2 \theta & 0 & 0 \end{array} \end{pmatrix}.$$

Finally, we obtain for the four elements on the main diagonal of the Ricci matrix  $\mathbf{R}_{\text{Ric}}$ :

$$\underline{\underline{R_{\text{Ric},00}}} = -\frac{A''}{2B} + \frac{A'}{4B} \left( \frac{A'}{A} + \frac{B'}{B} \right) - \frac{A'}{rB}, \quad (3.7)$$

$$\underline{\underline{R_{\text{Ric},11}}} = \frac{A''}{2A} - \frac{A'}{4A} \left( \frac{A'}{A} + \frac{B'}{B} \right) - \frac{B'}{rB}, \quad (3.8)$$

$$\underline{\underline{R_{\text{Ric},22}}} = \frac{1}{B} + \frac{r}{2B} \left( \frac{A'}{A} - \frac{B'}{B} \right) - 1 \quad (3.9)$$

and

$$\underline{\underline{R_{\text{Ric},33}}} = \sin^2 \theta R_{\text{Ric},22}. \quad (3.10)$$

The other matrix elements are equal to zero:  $R_{\text{Ric},\nu\mu} = 0$  for  $\nu \neq \mu$ .

### 3.1.3 The Factors $A(r)$ and $B(r)$

To finally write down the metric matrix  $\mathbf{G}$ , the two factors  $A(r)$  and  $B(r)$  are required. Since we search for the solution of Einstein's field equation outside the sphere containing the mass and having the energy-momentum matrix  $\mathbf{T}$  equal to the zero matrix, the Ansatz  $\mathbf{R}_{\text{Ric}} = \mathbf{0}$  thus yields

$$-\frac{A''}{2B} + \frac{A'}{4B} \left( \frac{A'}{A} + \frac{B'}{B} \right) - \frac{A'}{rB} = 0, \quad (3.11)$$

$$\frac{A''}{2A} - \frac{A'}{4A} \left( \frac{A'}{A} + \frac{B'}{B} \right) - \frac{B'}{rB} = 0, \quad (3.12)$$

and

$$\frac{1}{B} + \frac{r}{2B} \left( \frac{A'}{A} - \frac{B'}{B} \right) - 1 = 0. \quad (3.13)$$

Adding to (3.11) divided by  $A$  (3.12) divided by  $B$ , we obtain the condition

$$A'B + AB' = 0, \quad (3.14)$$

which means nothing else than that

$$AB = \text{constant}. \quad (3.15)$$

However, since  $A(\infty) = c^2$  and  $B(\infty) = 1$  as  $r \rightarrow \infty$ , the constant must be equal to  $c^2$ , so

$$A(r)B(r) = c^2 \quad \text{and} \quad B(r) = \frac{c^2}{A(r)}.$$

Substituting this into (3.13), we obtain  $A + rA' = c^2$ , which we also can write

$$\frac{d(rA)}{dr} = c^2. \quad (3.16)$$

Integrating this equation yields

$$rA = c^2(r + K), \quad (3.17)$$

so

$$A(r) = c^2 \left(1 + \frac{K}{r}\right) \quad \text{and} \quad B(r) = \left(1 + \frac{K}{r}\right)^{-1}. \quad (3.18)$$

$K$  still remains to be determined. In Sect. 2.8, a uniform rotating system was treated. In (2.85), we obtained the element  $g_{00}$  of the metric matrix  $\mathbf{G}$  as

$$g_{00} = 1 - \frac{r^2\omega^2}{c^2} \quad (3.19)$$

and in (2.88) the centrifugal force  $mr\omega^2$ . According to Sect. 2.8, the centrifugal force  $\mathbf{f}$  depends, on the other hand, on the centrifugal potential  $\varphi$ , so together

$$\mathbf{f} = -m\nabla\varphi. \quad (3.20)$$

Taking

$$\varphi = -\frac{r^2\omega^2}{2} \quad (3.21)$$

(3.20) provides just the above centrifugal force. Equation (3.21) inserted into (3.19) provides the general connection

$$g_{00} = 1 + \frac{2\varphi}{c^2}. \quad (3.22)$$

Outside of our spherical mass  $M$ , as Newton's approximation the gravitational potential

$$\varphi = -\frac{GM}{r} \quad (3.23)$$

is obtained. This used in (3.22) and compared with (3.18) yields

$$K = -\frac{2GM}{c^2}, \quad (3.24)$$

so that we finally obtain the Schwarzschild's metric in matrix form:

$$ds^2 = d\vec{x}^\top \mathbf{G} d\vec{x} = d\vec{x}^\top \begin{pmatrix} 1 - \frac{2GM}{c^2 r} & 0 & 0 & 0 \\ 0 & -(1 - \frac{2GM}{c^2 r})^{-1} & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2 \sin^2 \theta \end{pmatrix} d\vec{x}, \quad (3.25)$$

with

$$d\vec{x} \stackrel{\text{def}}{=} \begin{pmatrix} c dt \\ dr \\ d\theta \\ d\varphi \end{pmatrix},$$

and the calculated Schwarzschild's metric is then

$$ds^2 = \left(1 - \frac{2GM}{c^2 r}\right) c^2 dt^2 - \left(1 - \frac{2GM}{c^2 r}\right)^{-1} dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\varphi^2). \quad (3.26)$$

For the so-called Schwarzschild's radius

$$r_S \stackrel{\text{def}}{=} \frac{2GM}{c^2}, \quad (3.27)$$

the Schwarzschild's metric shows a singularity as for  $r = 0$ . Of these, however, only the singularity at  $r = 0$  is a real singularity. This is suggested even by the fact that, when  $r = r_S$ , the determinant of  $\mathbf{G}$ , namely  $g = -r^4 \sin^2 \theta$ , has no singularity. The singularity at  $r = r_S$  is not a *physical singularity* but a *coordinate singularity*, which depends entirely on the choice of the coordinate system, i.e. there would exist no singularity at this place if a different coordinate system were chosen! More about this in the next section on black holes.

## 3.2 Influence of a Massive Body on the Environment

### 3.2.1 Introduction

The Schwarzschild's solution (3.26) is valid only outside the solid sphere of radius  $r_M$  generating the gravity, so only for  $r_M < r < \infty$ . Because for the Schwarzschild's radius  $r_S$  the element  $g_{11} = (1 - \frac{r_S}{r})^{-1}$  goes to infinity,  $r_S$  is also a limit. In general,  $r_S \ll r_M$ , but for the so-called *Black Holes* one has  $r_M < r_S$ , and in this case the solution is limited to  $r_S < r < \infty$ .



### 3.2.2 Changes to Length and Time

How does a length change in the environment of a massive body? For a constant time, i.e. for  $dt = 0$  one receives, due to (3.26),

$$d\ell^2 = \left(1 - \frac{2GM}{c^2 r}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2). \quad (3.28)$$

On the surface of a sphere of radius  $r > r_M$  with the centre being the centre of mass, as also for  $dr = 0$ , the tangential infinitely small distances are given by

$$dL = r(d\theta^2 + \sin^2\theta d\varphi^2)^{1/2}. \quad (3.29)$$

This is a result that is valid on every sphere, whether with or without gravity. But what happens with a distance in the radial direction? In this case,  $d\theta$  and  $d\varphi$  are equal to zero, thus for infinitely small distances in the radial direction, according to (3.28), one obtains

$$\boxed{dR = \left(1 - \frac{2GM}{c^2 r}\right)^{-1/2} dr}, \quad (3.30)$$

so  $dR > dr$ , and so the length is much longer, the larger the mass  $M$  and the smaller the distance  $r$  from the mass!  $dr$  is through the mass ‘elongated’, caused by a curvature of space.

Let us now turn to the time. For a clock at the point, where  $r$ ,  $\theta$ , and  $\varphi$  are constant, we obtain from the Schwarzschild’s metric (3.26)

$$ds^2 = c^2 d\tau^2 = c^2 \left(1 - \frac{2GM}{c^2 r}\right) dt^2,$$

so

$$\boxed{d\tau = \left(1 - \frac{2GM}{c^2 r}\right)^{1/2} dt}. \quad (3.31)$$

If  $d\tau < dt$ , the closer one is located to the mass, the shorter are the time intervals, i.e. the slower the time passes! For an observer the time goes slower the closer he is to the mass. Particularly large is the time dilation near a *Black Hole* in which the mass is so concentrated that the body radius is smaller than its own Schwarzschild’s radius. For a black hole with a mass of ten solar masses, the Schwarzschild’s radius is  $r_S = 30$  km. At a distance of 1 cm from the so-called horizon, which is the spherical shell around the centre of mass with radius  $r_S$ ,  $\gamma = \left(1 - \frac{2GM}{c^2 r}\right)^{1/2} = 1.826 \cdot 10^{-5}$ , so the time goes by about 55000 times slower than far away: after a year near the horizon of the black hole, in a place far away, e.g. on the home planet of an astronaut, 55000 years would have gone by! All is summed up into a popular statement:

**If you climb a mountain, you are smaller and age faster!**

The relations (3.30) and (3.31) are very similar to the relationships of space and time contractions for the relatively to each other moving reference frames in special relativity; see Chap. 1. There with

$$\gamma \stackrel{\text{def}}{=} \left(1 - \frac{v^2}{c^2}\right)^{-1/2}$$

for the length contraction we had

$$d\ell = \frac{1}{\gamma} d\ell_0$$

and for the time dilation

$$dt = \gamma d\tau.$$

If we introduce the pseudo-velocity

$$v_G^2(r) \stackrel{\text{def}}{=} \frac{2GM}{r}$$

then with

$$\gamma_G(r) \stackrel{\text{def}}{=} \left(1 - \frac{v_G^2(r)}{c^2}\right)^{-1/2}$$

the above gravitational relationships can be written as follows:

$$dr = \frac{1}{\gamma_G(r)} dR \quad \text{and} \quad dt = \gamma_G(r) d\tau.$$

By the way,  $v_G(r)$  is the escape velocity of a planet with the diameter  $2r$  and mass  $M$ .

### 3.2.3 Redshift of Spectral Lines

Suppose in the gravitational field of a mass a light signal is sent from a transmitter at a fixed point  $\mathbf{x}_T = [r_T, \theta_T, \varphi_T]^T$  and it drifts along a geodesic line to a fixed receiver at the point  $\mathbf{x}_R = [r_R, \theta_R, \varphi_R]^T$ . Let, furthermore,  $\lambda$  be any parameter along the geodesic line, with  $\lambda = \lambda_T$  for the sent event and  $\lambda = \lambda_R$  for the received event. Then for a photon, by (2.11), the following is valid:

$$\frac{d\vec{x}^T}{d\lambda} \mathbf{G} \frac{d\vec{x}}{d\lambda} = 0,$$

so here

$$c^2 \left(1 - \frac{2GM}{c^2 r}\right) \left(\frac{dt}{d\lambda}\right)^2 - \left(1 - \frac{2GM}{c^2 r}\right)^{-1} \left(\frac{dr}{d\lambda}\right)^2 - r^2 \left( \left(\frac{d\theta}{d\lambda}\right)^2 + \sin^2 \left(\frac{d\varphi}{d\lambda}\right)^2 \right) = 0,$$

i.e.

$$\frac{dt}{d\lambda} = \frac{1}{c} \left[ \left(1 - \frac{2GM}{c^2 r}\right)^{-2} \left(\frac{dr}{d\lambda}\right)^2 + \left(1 - \frac{2GM}{c^2 r}\right)^{-1} r^2 \left( \left(\frac{d\theta}{d\lambda}\right)^2 + \sin^2 \left(\frac{d\varphi}{d\lambda}\right)^2 \right) \right]^{1/2}.$$

From this we obtain the signal transmission time

$$t_R - t_T = \frac{1}{c} \int_{\lambda_T}^{\lambda_R} \left[ \left(1 - \frac{2GM}{c^2 r}\right)^{-2} \left(\frac{dr}{d\lambda}\right)^2 + \left(1 - \frac{2GM}{c^2 r}\right)^{-1} r^2 \left( \left(\frac{d\theta}{d\lambda}\right)^2 + \sin^2 \left(\frac{d\varphi}{d\lambda}\right)^2 \right) \right]^{1/2} d\lambda.$$

This time depends only on the path that the light takes between the spatially fixed transmitter and the spatially fixed receiver. So for two consecutive transmitted signals 1 and 2 the duration is equal:

$$t_{E,1} - t_{S,1} = t_{E,2} - t_{S,2},$$

and also

$$\Delta t_R \stackrel{\text{def}}{=} t_{E,2} - t_{E,1} = t_{S,2} - t_{S,1} \stackrel{\text{def}}{=} \Delta t_T,$$

so the Schwarzschild's time difference at the transmitter is equal to the Schwarzschild's time difference at the receiver, though the clock of an observer at the transmitter location would show the proper time difference to

$$\Delta \tau_T = \left(1 - \frac{2GM}{c^2 r_T}\right)^{1/2} \Delta t_T$$

and accordingly at the receiving location the proper time difference would be

$$\Delta \tau_R = \left(1 - \frac{2GM}{c^2 r_R}\right)^{1/2} \Delta t_R.$$

Since  $\Delta t_R = \Delta t_T$ , one obtains the ratio

$$\frac{\Delta \tau_R}{\Delta \tau_T} = \left( \frac{1 - \frac{2GM}{c^2 r_R}}{1 - \frac{2GM}{c^2 r_T}} \right)^{1/2}.$$

If the transmitter is a pulsating atom, emitting  $N$  pulses in the proper time interval  $\Delta \tau_T$ , an observer at the transmitter would assign to the signal a proper frequency

$\nu_T \stackrel{\text{def}}{=} \frac{N}{\Delta\tau_T}$ . An observer at the receiver will see the  $N$  pulses arriving in the proper time  $\Delta\tau_R$ , thus at the frequency  $\nu_R = \frac{N}{\Delta\tau_R}$ . Then the frequency ratio is obtained as

$$\boxed{\frac{\nu_R}{\nu_T} = \left( \frac{1 - \frac{2GM}{c^2 r_T}}{1 - \frac{2GM}{c^2 r_R}} \right)^{1/2}} \quad (3.32)$$

If the transmitter is closer to the mass compared to the receiver ( $r_T < r_R$ ), then a shift to the red in the “colour” of the signal takes place. Conversely, if the transmitter is further away than the receiver, one gets a blueshift. For  $r_S, r_E \gg 2GM$  one obtains the approximation

$$\frac{\nu_R}{\nu_T} \approx 1 + \frac{GM}{c^2} \left( \frac{1}{r_R} - \frac{1}{r_T} \right)$$

and the relative frequency change

$$\frac{\Delta\nu}{\nu_T} \stackrel{\text{def}}{=} \frac{\nu_R - \nu_T}{\nu_T} \approx \frac{GM}{c^2} \left( \frac{1}{r_R} - \frac{1}{r_T} \right).$$

If the transmitter (e.g. a radiating atom) is on the solar surface and the observing receiver on the surface of the Earth, then

$$\frac{\Delta\nu}{\nu_T} \approx 2 \cdot 10^{-6}.$$

This effect is indeed poorly measurable due to a variety of disturbances by the atmosphere. However, this redshift is measurable using the Mössbauer’s effect (see the relevant Physics literature).

### 3.2.4 Deflection of Light

According to (2.11), for light one has

$$\vec{x}^\top \mathbf{G} \vec{x} = 0. \quad (3.33)$$

Differentiating (3.33) with respect to the orbital parameter  $\lambda$ , one obtains

$$\frac{\partial \vec{x}^\top}{\partial \lambda} \mathbf{G} \frac{\partial \vec{x}}{\partial \lambda} = 0. \quad (3.34)$$

Setting for  $\mathbf{G}$  the Schwarzschild’s metric, we obtain

$$c^2 \left( 1 - \frac{r_S}{r} \right) \left( \frac{\partial t}{\partial \lambda} \right)^2 - \left( 1 - \frac{r_S}{r} \right)^{-1} \left( \frac{\partial r}{\partial \lambda} \right)^2 - r^2 \left( \frac{\partial \theta^2}{\partial \lambda} + \sin^2 \theta \left( \frac{\partial \varphi}{\partial \lambda} \right)^2 \right) = 0. \quad (3.35)$$

Without loss of generality, especially in the present centrally symmetric solution, we can assume that  $\theta = \pi/2$  and  $\frac{\partial\theta}{\partial\lambda} = 0$ , and that the solution is in a plane through the centre of mass, then (3.35) simplifies to

$$c^2 \left(1 - \frac{r_S}{r}\right) \left(\frac{\partial t}{\partial \lambda}\right)^2 - \left(1 - \frac{r_S}{r}\right)^{-1} \left(\frac{\partial r}{\partial \lambda}\right)^2 - r^2 \left(\frac{\partial \varphi}{\partial \lambda}\right)^2 = 0. \quad (3.36)$$

For a light ray in the gravitational field the following holds:

$$\frac{\partial^2 \vec{x}}{\partial \lambda^2} = - \left( \mathbf{I}_4 \otimes \frac{\partial \vec{x}^\top}{\partial \lambda} \right) \mathbf{\Gamma} \frac{\partial \vec{x}}{\partial \lambda}. \quad (3.37)$$

With the Christoffel elements from Sect. 3.1.1 and  $\vec{x}^\top = [x_0 | r | \theta | \varphi]$  expanding (3.37) yields

$$\frac{d^2 x_0}{d\lambda^2} = - \frac{A'}{A} \frac{dx_0}{d\lambda} \frac{dr}{d\lambda}, \quad (3.38)$$

$$\frac{d^2 r}{d\lambda^2} = - \frac{A'}{B} \left(\frac{dx_0}{d\lambda}\right)^2 - \frac{B'}{2B} \left(\frac{dr}{d\lambda}\right)^2 + \frac{r}{B} \left(\frac{d\theta}{d\lambda}\right)^2 + \frac{r \sin^2 \theta}{B} \left(\frac{d\varphi}{d\lambda}\right)^2, \quad (3.39)$$

$$\frac{d^2 \theta}{d\lambda^2} = - \frac{2}{r} \frac{d\theta}{d\lambda} \frac{dr}{d\lambda} + \sin \theta \cos \theta \left(\frac{d\varphi}{d\lambda}\right)^2, \quad (3.40)$$

$$\frac{d^2 \varphi}{d\lambda^2} = - \frac{2}{r} \frac{d\varphi}{d\lambda} \frac{dr}{d\lambda} - \cot \theta \frac{d\theta}{d\lambda} \frac{d\varphi}{d\lambda}. \quad (3.41)$$

If the coordinate system is selected so that for the start time  $\lambda_0$  one has

$$\theta = \pi/2 \quad \text{and} \quad \frac{\partial \theta}{\partial \lambda} = 0, \quad (3.42)$$

then from (3.40)  $\frac{d^2 \theta}{d\lambda^2} = 0$ , so  $\theta(\lambda) \equiv \pi/2$ , i.e. the entire path remains in the plane through the centre of mass. With this  $\theta$  value from (3.41) one obtains

$$\frac{d^2 \varphi}{d\lambda^2} + \frac{2}{r} \frac{d\varphi}{d\lambda} \frac{dr}{d\lambda} = 0,$$

which can be summarized as follows:

$$\frac{1}{r^2} \frac{d}{d\lambda} \left( r^2 \frac{d\varphi}{d\lambda} \right) = 0. \quad (3.43)$$

It is therefore true that

$$\underline{\underline{r^2 \frac{d\varphi}{d\lambda} = \text{const.} = h.}} \quad (3.44)$$

Equation (3.38) can be converted to

$$\frac{d}{d\lambda} \left( \ln \frac{dx_0}{d\lambda} + \ln A \right) = 0 \quad (3.45)$$

and integrated to yield

$$\underline{\underline{A \frac{dx_0}{d\lambda} = \text{const.} = k.}} \quad (3.46)$$

Equations (3.42), (3.44) and (3.46) used in (3.39) provide

$$\frac{d^2 r}{d\lambda^2} + \frac{k^2 A'}{2A^2 B} + \frac{B'}{2B} \left( \frac{dr}{d\lambda} \right)^2 - \frac{h^2}{Br^3} = 0. \quad (3.47)$$

Multiplying this equation by  $2B \frac{dr}{d\lambda}$ , we obtain first

$$\frac{d}{d\lambda} \left( B \frac{dr}{d\lambda} + \frac{h^2}{r^2} - \frac{k^2}{A} \right) = 0 \quad (3.48)$$

and after integration

$$B \left( \frac{dr}{d\lambda} \right)^2 + \frac{h^2}{r^2} - \frac{k^2}{A} = \text{const.} = 0, \quad (3.49)$$

or

$$\frac{dr}{d\lambda} = \sqrt{\frac{\frac{k^2}{A} - \frac{h^2}{r^2}}{B}}. \quad (3.50)$$

But we don't look for  $r(\lambda)$ , we want  $\varphi(r)$ . Since

$$\frac{d\varphi}{dr} = \frac{d\varphi}{d\lambda} \frac{d\lambda}{dr},$$

it follows by (3.44) and (3.50) that

$$\frac{d\varphi}{dr} = \frac{h}{r^2} \sqrt{\frac{B}{\frac{k^2}{A} - \frac{h^2}{r^2}}} \quad (3.51)$$

and integrated

$$\underline{\underline{\varphi(r) = \varphi(r_0) + \int_{r_0}^r \frac{\sqrt{B(\psi)}}{\psi^2 \sqrt{\frac{k^2}{A(\psi)h^2} - \frac{1}{\psi^2}}} d\psi.}} \quad (3.52)$$

Let  $r_0$  be the smallest distance from the centre of mass, having a passing flying photon. It moves in the plane through the centre of mass with  $\theta = \pi/2$ . The coordinate system is still placed such that  $\phi(r_0) = 0$ . If gravity is not present, the

photon will continue to fly straight and as  $r \rightarrow \infty$  we will have  $\varphi(\infty) = \pi/2$ . It comes from  $\varphi(-\infty) = -\pi/2$ . Overall, therefore, the photon has covered an angle of  $\Delta\varphi = \varphi(\infty) - \varphi(-\infty) = \pi$ . Considering the gravity, the photon is “bent” to the mass. It passes from  $(r_0, \varphi(r_0))$  to  $\varphi(\infty) = \pi/2 + \alpha/2$  as  $r \rightarrow \infty$ . For reasons of symmetry, it is then  $\varphi(-\infty) = \pi/2 - \alpha/2$  as  $r \rightarrow -\infty$ , so that the photon has covered an overall angle of  $\Delta\varphi = \varphi(\infty) - \varphi(-\infty) = \pi + \alpha$ . This angle  $\alpha$  is now calculated for the solar mass  $M_\odot = 1.9891 \times 10^{33}$  g.

Since  $r_0$  should be the minimum distance, it is true that

$$\left. \frac{dr}{d\varphi} \right|_{r_0} = 0. \quad (3.53)$$

On the other hand,  $\frac{d\varphi}{dr}$  is precisely the integrand of (3.52), whence it follows

$$\frac{k^2}{h^2} = \frac{A(r_0)}{r_0^2}. \quad (3.54)$$

The integral used in (3.52) yields

$$\varphi(\infty) = \int_{r_0}^{\infty} \frac{\sqrt{B(r)}}{r \sqrt{\frac{r^2 A(r_0)}{r_0^2 A(r)} - 1}} dr. \quad (3.55)$$

For this integral no simple antiderivative exists so that it should be calculated by approximations. One has

$$A(r) = 1 - \frac{2GM_\odot}{c^2 r} \quad (3.56)$$

and

$$B(r) = \left(1 - \frac{2GM_\odot}{c^2 r}\right)^{-1} \approx 1 + \frac{2GM_\odot}{c^2 r}. \quad (3.57)$$

Furthermore,

$$\begin{aligned} \frac{A(r_0)}{A(r)} &= \left(1 - \frac{2GM_\odot}{c^2 r_0}\right) \left(1 - \frac{2GM_\odot}{c^2 r}\right)^{-1} \\ &\approx \left(1 - \frac{2GM_\odot}{c^2 r_0}\right) \left(1 + \frac{2GM_\odot}{c^2 r}\right) = 1 + \frac{2GM_\odot}{c^2} \left(\frac{1}{r} - \frac{1}{r_0}\right) \end{aligned}$$

and

$$\frac{r^2 A(r_0)}{r_0^2 A(r)} - 1 \approx \left(\frac{r^2}{r_0^2} - 1\right) \left(1 - \frac{2GM_\odot r}{c^2 r_0(r + r_0)}\right). \quad (3.58)$$

Therefore, the integral (3.55) becomes

$$\begin{aligned}\varphi(\infty) &= \int_{r_0}^{\infty} \frac{r_0}{\sqrt{r^2 - r_0^2}} \left( \frac{1}{r} + \frac{GM_{\odot}}{c^2 r^2} + \frac{GM_{\odot}}{c^2 (r + r_0)} \right) dr \\ &= \left[ \arccos \frac{r_0}{r} + \frac{GM_{\odot}}{c^2 r_0} \frac{\sqrt{r^2 - r_0^2}}{r} + \frac{GM_{\odot}}{c^2 r_0} \sqrt{\frac{r - r_0}{r + r_0}} \right]_{r_0}^{\infty},\end{aligned}\quad (3.59)$$

so

$$\underline{\underline{\varphi(\infty) = \frac{\pi}{2} + \frac{2GM_{\odot}}{c^2 r_0}}},\quad (3.60)$$

from which one can read off

$$\underline{\underline{\alpha = \frac{4GM_{\odot}}{c^2 r_0}}}. \quad (3.61)$$

If  $r_0$  is precisely the solar radius  $R_{\odot} = 6.957 \times 10^8$  m, we obtain

$$\underline{\underline{\alpha = 1.75''}}. \quad (3.62)$$

This prediction was one of the first confirmations of the Theory of General Relativity in 1919 at a solar eclipse measured by Eddington.

### 3.3 Schwarzschild's Inner Solution

For the determination of the metric inside a symmetric sphere, the right-hand side of Einstein's field equation with the energy-momentum matrix is needed. The hydro-mechanical energy-momentum matrix was deduced in Chap. 1 in (1.183) as

$$\mathbf{T}_{\text{mech}} = \left( \rho_0 + \frac{p}{c^2} \right) \vec{u} \vec{u}^{\top} - p \mathbf{M}.$$

This was done in the Lorentz basis of an inertial frame. It can thus also be the local inertial system of a general coordinate system. If for the relationship between the local coordinates  $d\vec{x}$  and the global coordinates  $d\vec{x}'$  the transformation equation is  $d\vec{x} = \mathbf{J} d\vec{x}'$ , then for the velocities one has  $\vec{u} = \mathbf{J} \vec{u}'$ . If used above this yields

$$\mathbf{T}_{\text{mech}} = \left( \rho_0 + \frac{p}{c^2} \right) \mathbf{J} \vec{u}' \vec{u}'^{\top} \mathbf{J}^{\top} - p \mathbf{M}. \quad (3.63)$$

Now multiplying this equation from the left with the matrix product  $\mathbf{J}^{\top} \mathbf{M}$  and from the right with the matrix product  $\mathbf{M} \mathbf{J}$ , we finally obtain with  $\mathbf{M} \mathbf{M} = \mathbf{I}$  and  $\mathbf{J}^{\top} \mathbf{M} \mathbf{J} = \mathbf{G}$

$$\mathbf{T}_{\text{mech, Riemann}} = \left( \rho_0 + \frac{p}{c^2} \right) \mathbf{G} \vec{u}' \vec{u}'^{\top} \mathbf{G} - p \mathbf{G}. \quad (3.64)$$



Now the metric matrix  $\mathbf{G}$  appears globally.

The Einstein's field equation is used in the form (2.237)

$$\mathbf{R}_{\text{Ric}} = \frac{8\pi G}{c^4} \left( \frac{T}{2} \mathbf{I}_4 - \mathbf{T} \right). \quad (3.65)$$

It is assumed that the masses inside the sphere do not move, the velocity components are zero:  $u'_i = 0$ , for  $i = 1, 2, 3$ . Therefore, from the condition

$$c^2 = \bar{\mathbf{u}}'^T \mathbf{G} \bar{\mathbf{u}}'$$

there remains

$$c^2 = g_{00} u_0^2.$$

If the metric matrix is chosen again as in (3.2), namely

$$ds^2 = A(r) dt^2 - B(r) dr^2 - r^2(d\theta^2 + \sin^2\theta d\varphi^2), \quad (3.66)$$

then

$$u_0 = \frac{c}{\sqrt{A(r)}}.$$

Thus we get for  $\mathbf{T}$  the diagonal matrix

$$\mathbf{T} = \text{diag}(\varrho c^2 A(r), pB(r), pr^2, pr^2 \sin^2\theta). \quad (3.67)$$

With the trace  $T = c^2\varrho - 3p$  of the local energy–momentum matrix, we obtain for the right-hand side of the Einstein's field equation

$$\begin{aligned} & \frac{8\pi G}{c^4} \left( \frac{T}{2} \mathbf{I}_4 - \mathbf{T} \right) \\ &= \frac{4\pi G}{c^4} \text{diag}((\varrho c^2 + 3p)A, (\varrho c^2 - p)B, (\varrho c^2 - p)r^2, (\varrho c^2 - p)r^2 \sin^2\theta). \end{aligned} \quad (3.68)$$

This with the elements (3.7), (3.8), (3.9) and (3.10) on the main diagonal of the Ricci matrix  $\mathbf{R}_{\text{Ric}}$  gives

$$\begin{aligned} R_{\text{Ric},00} &= -\frac{A''}{2B} + \frac{A'}{4B} \left( \frac{A'}{A} + \frac{B'}{B} \right) - \frac{A'}{rB}, \\ R_{\text{Ric},11} &= \frac{A''}{2A} - \frac{A'}{4A} \left( \frac{A'}{A} + \frac{B'}{B} \right) - \frac{B'}{rB}, \\ R_{\text{Ric},22} &= \frac{1}{B} + \frac{r}{2B} \left( \frac{A'}{A} - \frac{B'}{B} \right) - 1 \end{aligned}$$

and

$$R_{\text{Ric},33} = \sin^2 \theta R_{\text{Ric},22}$$

(the remaining matrix elements are equal to zero:  $R_{\text{Ric},\nu\mu} = 0$  for  $\nu \neq \mu$ ), combined with Einstein's field equation yielding the three determining equations for the factors  $A(r)$  and  $B(r)$ :

$$\frac{A''}{2B} - \frac{A'}{4B} \left( \frac{A'}{A} + \frac{B'}{B} \right) + \frac{A'}{rB} = \frac{4\pi G}{c^4} (\varrho c^2 + 3p)A, \quad (3.69)$$

$$-\frac{A''}{2A} + \frac{A'}{4A} \left( \frac{A'}{A} + \frac{B'}{B} \right) + \frac{B'}{rB} = \frac{4\pi G}{c^4} (\varrho c^2 - p)B \quad (3.70)$$

and

$$-\frac{1}{B} - \frac{r}{2B} \left( \frac{A'}{A} - \frac{B'}{B} \right) + 1 = \frac{4\pi G}{c^4} (\varrho c^2 - p)r^2. \quad (3.71)$$

Adding together (3.69) multiplied by  $r^2/(2A)$ , (3.70) multiplied by  $r^2/(2B)$  and (3.71), we obtain

$$\frac{B'r}{B^2} + 1 - \frac{1}{B} = \frac{8\pi G}{c^2} \varrho r^2. \quad (3.72)$$

This can also be rewritten as

$$\frac{d}{dr} \frac{r}{B(r)} = 1 - \frac{8\pi G}{c^2} \varrho r^2. \quad (3.73)$$

Integrating this equation provides

$$\frac{r}{B(r)} = \int_0^r \left( 1 - \frac{8\pi G}{c^2} \varrho(\alpha) \alpha^2 \right) d\alpha = r - \frac{2G}{c^2} \mathcal{M}(r), \quad (3.74)$$

with

$$\mathcal{M}(r) \stackrel{\text{def}}{=} 4\pi \int_0^r \varrho(\alpha) \alpha^2 d\alpha. \quad (3.75)$$

Solving (3.74) for  $B(r)$  finally yields

$$\underline{\underline{B(r) = \left( 1 - \frac{2G\mathcal{M}(r)}{c^2 r} \right)^{-1}}}. \quad (3.76)$$

The density function  $\varrho(r)$  is zero for  $r > R$ , i.e. outside the spherical mass, so  $\mathcal{M}(r) = \mathcal{M}(R) = M$  for  $r > R$ . We thus obtain the same coefficient  $B(r)$  for  $r > R$  as for the Schwarzschild's outer solution.

Now  $A(r)$  is still to be determined. For that,  $B(r)$  and

$$B'(r) = B^2(r) \left( \frac{8\pi G}{c^2} \varrho(r)r - \frac{2G\mathcal{M}(r)}{c^2 r^2} \right)$$

is used in (3.71), which gives

$$-\frac{A'(r)r}{2A(r)}\left(1 - \frac{2GM(r)}{c^2 r}\right) + \frac{4\pi G}{c^2}\varrho(r)r^2 + \frac{GM(r)}{c^2 r} = \frac{4\pi G}{c^4}(\varrho c^2 - p)r^2,$$

i.e.

$$\frac{A'(r)}{A(r)} = \frac{d}{dr}(\ln A(r) + \text{const.}) = \left(\frac{8\pi G}{c^4}p(r)r + \frac{2GM(r)}{c^2 r^2}\right)\left(1 - \frac{2GM(r)}{c^2 r}\right)^{-1} \\ \stackrel{\text{def}}{=} f(r).$$

The sum  $\ln A(r) + \text{const.}$  is thus the antiderivative for the right-hand side  $f(r)$  of this equation. Therefore,  $\ln A(r)$  equals the integral of  $f(r)$ . If we use as integration limits  $r$  and  $\infty$  and  $A(\infty) = 1$ , we finally obtain

$$\underline{\underline{A(r) = \exp\left[-\frac{2G}{c^2} \int_r^\infty \frac{1}{\alpha^2}(\mathcal{M}(\alpha) + 4\pi\alpha^3 p(\alpha)/c^2)\left(1 - \frac{2GM}{c^2\alpha}\right)^{-1} d\alpha\right]}. \quad (3.77)}}$$

For  $r > R$ , this solution is equal to the Schwarzschild's outer solution because for  $r > R$  one has  $\varrho(r) = p(r) = 0$  and  $\mathcal{M}(r) = \mathcal{M}(R) = M$ , so

$$A(r) = \exp\left[-\frac{2G}{c^2} \int_r^\infty \frac{M}{\alpha^2} \left(1 - \frac{r_S}{\alpha}\right)^{-1} d\alpha\right]. \quad (3.78)$$

With the new integration variable  $\xi = 1 - \frac{r_S}{r}$  we obtain for  $r > R$

$$A(r) = \exp\left[\int_1^{1-\xi} \frac{1}{\xi'} d\xi'\right] = 1 - \frac{r_S}{r}, \quad (3.79)$$

i.e. the same solution as the Schwarzschild's outer solution. The calculated factors  $A(r)$  and  $B(r)$  therefore apply both inside and outside the spherical mass of radius  $R$ .

## 3.4 Black Holes

### 3.4.1 Astrophysics

The three most interesting celestial objects are *white dwarfs*, *neutron stars* and *black holes*. Stars like the Sun are built up by compression of interstellar clouds, caused by gravity. The contraction process comes to a standstill when the temperature inside becomes so great that nuclear fusion starts. By the nuclear fusion, hydrogen is converted into helium. The further fate of a star depends on its initial mass: The initial mass equal to

- $\approx 0.05$  solar mass ends up as a *brown dwarf*;
- $\approx 1$  solar mass and of volume comparable to that of the Earth grows into a red giant and ends after an explosion (supernova) as a *white dwarf*;
- $\approx 3$  solar masses, with a corresponding radius of about 12 km, grows into a supergiant and ends after explosion (supernova) as a *neutron star*;
- $\approx 30$  solar masses grows into a supergiant and ends after explosion (supernova) as a *black hole*.

A white dwarf is composed mostly of carbon and oxygen, produced by nuclear fusion, that are left over when the nuclear solar fuel is consumed. A white dwarf has a mass about that of our Sun and the diameter of the Earth, so a fairly compact structure. However, a neutron star is much more compact, i.e. it has a factor of  $10^9$  higher density. It has roughly the mass of a white dwarf but a diameter of only 24 km! A black hole, on the other hand, has different masses, sizes, and densities, as is further explained below.

The Schwarzschild's solution (3.26)

$$ds^2 = \left(1 - \frac{r_S}{r}\right) c^2 dt^2 - \left(1 - \frac{r_S}{r}\right)^{-1} dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\varphi^2)$$

applies only outside the considered spherical mass. This formula gives the tiny spacetime interval  $ds$  which mass elements cover, moving from an event A to a closely adjacent event B. The Schwarzschild's radius is therefore only interesting for very large masses or masses with sufficient density, where  $r_S > R$  is the radius of the spherical mass.

However, for the solar mass  $M_\odot \approx 2 \cdot 10^{30}$  kg the Schwarzschild's radius is

$$r_{S\odot} \approx 3 \text{ km},$$

i.e. it is much smaller than the solar radius  $r_\odot \approx 7 \cdot 10^5$  km, so our solution is valid only outside the mass of the Sun! For the Earth with a mass of  $M_\oplus \approx 6 \cdot 10^{24}$  kg we even obtain a Schwarzschild's radius of

$$r_{S\oplus} \approx 9 \text{ mm!}$$

It seems that the Schwarzschild's radius only plays a role for highly concentrated matter. But that is not the case. Indeed, consider just the radius of a spherical mass  $R = r_S$ , that is,  $R = \frac{2GM}{c^2}$ . The so-called Schwarzschild's *density*  $\rho_S$  would in this case be

$$\begin{aligned} \underline{\underline{\rho_S}} &\stackrel{\text{def}}{=} \frac{M}{\frac{4}{3}\pi r_S^3} = \frac{M}{\frac{4}{3}\pi \left(\frac{2GM}{c^2}\right)^3} \\ &= \frac{3c^6}{32\pi G^3 M^2} = \underline{\underline{2.33 \cdot 10^{71} \cdot M_{[\text{kg}]^{-2} \left[ \frac{\text{kg}}{\text{cm}^3} \right]}}}. \end{aligned}$$

The required density decreases inversely proportionally with the square of the mass! If one wants to do a black hole from Earth's mass, i.e. shrink the Earth to a radius

of 9 mm, then the resulting body would have a density of  $8.58 \cdot 10^{22} \text{ kg/cm}^3$ . For the Sun a density of approximately  $8.6 \cdot 10^{10} \text{ kg/cm}^3$  would be necessary; that is, 86 million tons per cubic centimetre! So still the enormous density of the neutron liquid<sup>1</sup> is in a neutron star. In the centre of the Milky Way, there is a black hole with a mass of  $5.2 \cdot 10^{36} \text{ kg}$ , or 2.6 million solar masses, and in the centre of the Virgo galaxy cluster, a black hole with the mass of  $6 \cdot 10^{39} \text{ kg}$  (3 billion solar masses).

### 3.4.2 Further Details about “Black Holes”

A spherical mass of radius less than the Schwarzschild’s radius,  $R < r_S$ , is called a *black hole*. The name is explained as follows. The Schwarzschild’s solution (3.26) for constant angles  $\theta$  and  $\varphi$ , i.e. for  $d\theta = d\varphi = 0$ , with  $r_S = \frac{2GM}{c^2}$ , is

$$ds^2 = \left(1 - \frac{r_S}{r}\right) c^2 dt^2 - \left(1 - \frac{r_S}{r}\right)^{-1} dr^2. \quad (3.80)$$

For light, i.e. photons, one has  $ds^2 = 0$ , so from (3.80) follows

$$c^2 dt^2 = \left(1 - \frac{r_S}{r}\right)^{-2} dr^2,$$

giving

$$c dt = \pm \left(1 - \frac{r_S}{r}\right)^{-1} dr = \pm \frac{r}{r - r_S} dr,$$

which integrated yields

$$c \int_{t_0}^t dt = ct - ct_0 = \int_{r_0}^r \frac{r}{r - r_S} dr = \pm \left[ r + r_S \ln \frac{r - r_S}{r_0 - r_S} - r_0 \right]$$

and with  $c_0 \stackrel{\text{def}}{=} ct_0 - r_0 - r_S \ln(r_0 - r_S)$  finally

$$\underline{\underline{ct = \pm (r + r_S \ln(r - r_S) + c_0)}}, \quad (3.81)$$

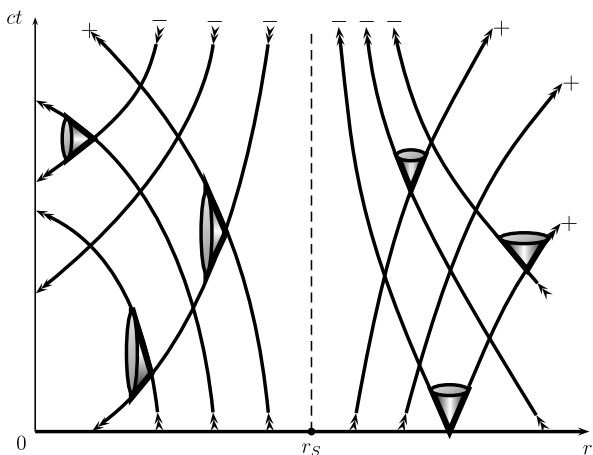
where the constant  $c_0$  depends on the start time  $t_0$  and on the starting location  $R_0$ , and  $ct$  has the dimension of a length. The substance of this formula can be pictured in the  $(ct - r)$ -half-plane, as shown in Fig. 3.1.

The light cones are to the right of the dividing line at  $r = r_S$  upward open towards the time axis, i.e. particles (including photons) in the picture fly only upwards. In particular, even photons on the minus-sign trajectories seemingly not exceed the

---

<sup>1</sup>Neutron liquid consists predominantly of neutrons, and their average density is approximately equal to that of atomic nuclei.

**Fig. 3.1** Schwarzschild's solution



dividing line  $r = r_s$ , so can never arrive at the field with  $r < r_s$ . This is particularly recognised at the speed of the particles or photons, resulting from the above equations in

$$\frac{dr}{dt} = \pm \left(1 - \frac{r_s}{r}\right) c.$$

This velocity goes to zero as  $r$  goes to  $r_s$ . For  $r < r_s$  the light cone shows the future in the direction  $r = 0$ . No particle or photon can escape from the field  $r < r_s$ ! No light can overstep the limit of  $r_s$ ; the “black hole” is, in fact, “black”! For a particle at  $r_s$  as is in Fig. 3.1, the time  $t$  decreases, so runs backwards! It is obvious that for this region, the time  $t$  is not very suitable. On the other hand, it is a fallacy that the boundary  $r = r_s$  cannot be exceeded from the outside because the formulas report only the behaviour which would be seen by an observer. An observer, flying with the particle, would normally exceed the limit  $r = r_s$  because for him the speed in the vicinity  $r = r_s$  would not tend to zero, he would fly continuously through this sphere with finite speed. This can be shown as follows:

Due to Newton's theory, we arrive at the following context. Integrating the equation

$$m_0 a(t) = m_0 \frac{dv}{dt} = G \frac{M m_0}{r^2}$$

with respect to

$$a = \frac{dv}{dt} = \frac{dv}{dr} \frac{dr}{dt} = v \frac{dv}{dr},$$

one obtains

$$\int_0^v \tilde{v} d\tilde{v} = \frac{v^2}{2} = \int_\infty^r a d\tilde{r} = \int_\infty^r \frac{GM}{\tilde{r}^2} d\tilde{r} = \frac{GM}{r},$$

or solving for the velocity

$$\underline{\underline{v = -\sqrt{\frac{2GM}{r}} = -c\sqrt{\frac{r_S}{r}}.}} \quad (3.82)$$

The mass  $m_0$  would pass the Schwarzschild's spherical shell for  $r = r_S$  and according to (3.82) with the speed of light  $c$ , which is not possible according to the special theory of relativity! For  $r < r_S$  the velocity would even be greater than light's! The velocity equations have to be modified. What happens then? If one passes through the Schwarzschild's spherical shell in the direction  $r \rightarrow 0$ , we will never be able to learn from him what he experienced because neither he nor a signal from him to us outside the Schwarzschild's spherical shell will ever reach us. We only can analyse theoretically what happens.

**Within the Schwarzschild's Radius** From (3.82) for the proper time we have

$$d\tau = -\frac{1}{c} \left( \frac{r}{r_S} \right)^{1/2} dr. \quad (3.83)$$

From this we get the time  $\tau$  needed when flying through the Schwarzschild's spherical shell to reach the singularity  $r = 0$ . It is

$$\tau_S \stackrel{\text{def}}{=} -\frac{1}{c} \int_{r_S}^0 \left( \frac{r}{r_S} \right)^{1/2} dr = \frac{2}{3} \frac{r_S}{c}, \quad (3.84)$$

therefore, a finite time.

If a particle is located to the left of the parting line  $r = r_S$ , then the particle flies left in the direction  $r \rightarrow 0$ . Photons, starting left of the parting line cannot also exceed this border. Therefore, no light (photon) is coming outward from inside the sphere of radius  $r_S$ , a "black hole" is given there. Further, it is interesting that an ensemble of boundary trajectories (each characterized with a minus sign) proceeds in the negative direction of time, thus decreasing time. If one starts, for instance, at the point  $(t_0 = 0, r = r_0 < r_S)$  on a 'minus-trajectory', then one travels to the past! The result, however, contains no logical contradiction because one cannot act from the inside of the sphere with the radius  $r_S$  on the outside: one only passively receives signals, i.e. one can "look into the past".

### 3.4.3 Singularities

The solution (3.26) has two singularities: a real singularity at  $r = 0$  and an apparent one at  $r = r_S$ . The apparent singularity is a so-called *coordinate singularity* which could have been avoided by a better coordinate choice. Such a favourable coordinate

system would have been obtained, for example, if we had introduced a different radial coordinate  $r^*$  as follows:

$$r = \left(1 + \frac{r_S}{4r^*}\right)^2 r^*.$$

Then

$$\frac{dr}{dr^*} = \left(1 - \frac{r_S}{4r^*}\right) \left(1 + \frac{r_S}{4r^*}\right),$$

so

$$dr = \left(1 - \frac{r_S}{4r^*}\right) \left(1 + \frac{r_S}{4r^*}\right) dr^*.$$

Used in (3.26) this yields

$$ds^2 = \left(\frac{1 - \frac{r_S}{4r^*}}{1 + \frac{r_S}{4r^*}}\right)^2 c^2 dt^2 - \left(1 + \frac{r_S}{4r^*}\right)^4 (dr^{*2} + r^{*2}(d\theta^2 + \sin^2\theta d\varphi^2)).$$

In these coordinates, in fact, only a singularity at  $r^* = 0$  exists!

The authenticity of the singularity at  $r = 0$  can be seen by examining the invariants. A function of coordinates is invariant under a transformation if it remains unchanged applying the transformation to the coordinates. This is precisely the mark of a true singularity, which does not depend on the randomly selected coordinate system. Such invariants were already encountered in the investigation of electromagnetic fields and the Lorentz transformation, e.g.

$$-\frac{1}{2}\text{trace}(\mathbf{F}_{B,e}^* \mathbf{F}_{B,e}) = b^2 - e^2 = b'^2 - e'^2$$

and

$$-\frac{1}{4}\text{trace}(\mathbf{F}_{B,e}^* \mathbf{F}_{E,b}) = \mathbf{e}^\top \mathbf{b} = \mathbf{e}'^\top \mathbf{b}'.$$

Such an invariant for our problem now is the so-called Kretschmann's invariant, which is defined with the aid of the modified symmetric Riemannian curvature matrix

$$\mathbf{R}^* \stackrel{\text{def}}{=} (\mathbf{I}_4 \otimes \mathbf{G}^{-1}) \mathbf{R}$$

as follows:

$$I_K \stackrel{\text{def}}{=} \text{trace}(\mathbf{R}^* \mathbf{R}^*). \quad (3.85)$$

Because the Ricci matrix  $\mathbf{R}_{\text{Ric}}$  equals the zero matrix, it is unsuitable for an invariant. The newly defined matrix  $\mathbf{R}^*$  will now be derived step by step for the



Schwarzschild's metric. First, we have

$$\Gamma = \begin{pmatrix} 0 & \frac{m}{r^2 h} & 0 & 0 \\ \frac{m}{r^2 h} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline \frac{m h}{r^2} & 0 & 0 & 0 \\ 0 & -\frac{m}{r^2 h} & 0 & 0 \\ 0 & 0 & -r h & 0 \\ 0 & 0 & 0 & -r h \sin^2 \theta \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{r} & 0 \\ 0 & \frac{1}{r} & 0 & 0 \\ 0 & 0 & 0 & -\sin \theta \cos \theta \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{r} \\ 0 & 0 & 0 & \cot \theta \\ 0 & \frac{1}{r} & \cot \theta & 0 \end{pmatrix},$$

where

$$m \stackrel{\text{def}}{=} \frac{GM}{c^2} \quad \text{and} \quad h \stackrel{\text{def}}{=} 1 - \frac{2m}{r}.$$

With the help of (2.165)  $\mathbf{R}$  can be calculated. The result is

$$\mathbf{R} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{2m}{r^3 h} & 0 & 0 & \frac{2m}{r^3 h} & 0 & 0 & 0 \\ 0 & 0 & \frac{m}{r} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{m \sin^2 \theta}{r} & 0 & 0 & 0 & 0 \\ \hline 0 & -\frac{2m h}{r^3} & 0 & 0 & \frac{2m h}{r^3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{m}{r} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{m \sin^2 \theta}{r} \\ \hline 0 & 0 & \frac{m h}{r^3} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{m}{r^3 h} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & \frac{m h}{r^3} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{m}{r^3 h} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\left( \begin{array}{cccc|cccc}
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 -\frac{m}{r} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & -\frac{m \sin^2 \theta}{r} & 0 & 0 & 0 \\
 \hline
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & -\frac{m}{r} & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & -\frac{m \sin^2 \theta}{r} & 0 & 0 \\
 \hline
 -\frac{mh}{r^3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & \frac{m}{r^3 h} & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & -\frac{2m \sin^2 \theta}{r} & 0 & 0 & \frac{2m \sin^2 \theta}{r} & 0 \\
 \hline
 0 & 0 & 0 & 0 & -\frac{mh}{r^3} & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & \frac{m}{r^3 h} & 0 & 0 \\
 0 & 0 & 0 & \frac{2m}{r} & 0 & 0 & -\frac{2m}{r} & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
 \end{array} \right) .$$

The first, (4 + 2)th, (8 + 3)th and 16th row/column are zero rows/columns, as it must always be in the matrix  $\mathbf{R}$ . With  $\mathbf{R}$  the following symmetric matrix is obtained

$$(\mathbf{I}_4 \otimes \mathbf{G}^{-1})\mathbf{R}$$

$$= \left( \begin{array}{cccc|cccc}
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & \frac{2m}{r^3} & 0 & 0 & -\frac{2m}{r^3} & 0 & 0 & 0 \\
 0 & 0 & -\frac{m}{r^3} & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & -\frac{m}{r^3} & 0 & 0 & 0 & 0 \\
 \hline
 0 & -\frac{2m}{r^3} & 0 & 0 & \frac{2m}{r^3} & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & -\frac{m}{r^3} & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{m}{r^3} \\
 \hline
 0 & 0 & \frac{m}{r^3} & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & \frac{m}{r^3} & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 \hline
 0 & 0 & 0 & \frac{m}{r^3} & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{m}{r^3} \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
 \end{array} \right)$$

$$\left( \begin{array}{cccc|cccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{m}{r^3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{m}{r^3} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{m}{r^3} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{m}{r^3} & 0 & 0 \\ \hline -\frac{m}{r^3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{m}{r^3} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{2m}{r^3} & 0 & 0 & -\frac{2m}{r^3} & 0 \\ \hline 0 & 0 & 0 & 0 & -\frac{m}{r^3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{m}{r^3} & 0 & 0 \\ 0 & 0 & 0 & -\frac{2m}{r^3} & 0 & 0 & \frac{2m}{r^3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right),$$

and finally the Kretschmann’s invariant is

$$\underline{\underline{I_K = \text{trace}(\mathbf{R}^* \mathbf{R}^*) = 48 \frac{m^2}{r^6} = 12 \frac{r_S^2}{r^6}}}$$

Hence, it is now quite apparent that the only real singularity occurs at  $r = 0$ !

**Event Horizon Detector** In [13], a possibility is given, how to determine if one approaches the event horizon (the spherical shell of radius  $r_S$ ) of a black hole or even exceeds it. To this end, the authors give the invariant

$$I_1 = -\frac{720M^2(2M - r)}{r^9},$$

which for  $r = 2M = r_S$  is zero and outside the horizon is positive with a maximum at  $r = 9M/4$ . An observer falling to the black hole can detect the presence of a horizon by observation of  $I_1$ . If the event horizon is crossed, it is too late for the observer. But he may, if the maximum is exceeded, use it as a warning and quickly initiate the trajectory reversal.

### 3.4.4 Eddington’s Coordinates

The coordinate singularity at  $r = r_S$  of the Schwarzschild’s metric can be eliminated, for example, by using the following coordinate transformation.

In (3.81), from the Schwarzschild’s solution we derived

$$ct = \pm(r + r_S \ln|r - r_S| + c_0).$$

If we define

$$r^* \stackrel{\text{def}}{=} r_S \ln(r - r_S), \quad (3.86)$$

we get from

$$\frac{dr^*}{dr} = \frac{r_S}{r - r_S}$$

the differential

$$dr^* = \frac{r_S}{r} \frac{r}{r - r_S} dr = \frac{r_S}{r} \left( \frac{1}{1 - \frac{r_S}{r}} \right)^{-1} dr. \quad (3.87)$$

It looks like the term

$$-\left( \frac{1}{1 - \frac{r_S}{r}} \right)^{-1} dr^2$$

in the Schwarzschild's metric, which was the reason for the coordinate singularity! If (3.87) were squared, indeed  $dr^2$  would appear, but the parenthesis would have too much negative power. If one installs, however, the term  $r^*$  into a new time coordinate  $t^*$ , then within the Schwarzschild's metric it would again be multiplied by  $\frac{r}{r - r_S}$ , i.e. have the "right" power! Now the new time coordinate is set as

$$ct^* \stackrel{\text{def}}{=} ct + r^* = ct + r_S \ln|r - r_S|. \quad (3.88)$$

Differentiation with respect to  $r$  gives

$$c \frac{dt^*}{dr} = c \frac{dt}{dr} + \frac{dr^*}{dr}, \quad (3.89)$$

so

$$c dt^* = c dt + dr^* = c dt + \frac{r_S}{r - r_S} dr,$$

i.e.

$$c dt = c dt^* - \frac{r_S}{r - r_S} dr,$$

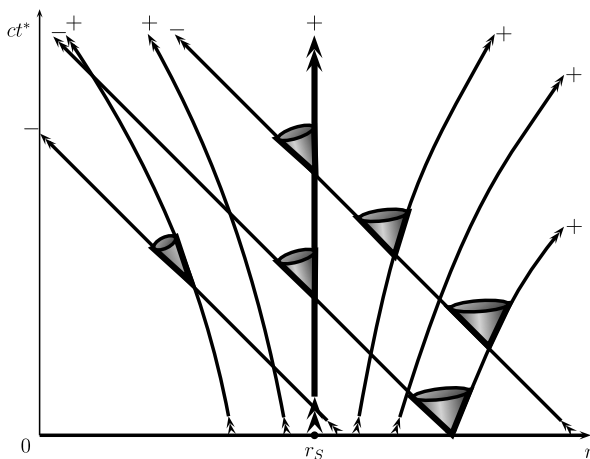
and squared

$$c^2 dt^2 = c^2 dt^{*2} - 2 \frac{r_S}{r - r_S} c dt^* dr + \left( \frac{r_S}{r - r_S} \right)^2 dr^2.$$

Used in the Schwarzschild's metric this yields

$$ds^2 = \left( 1 - \frac{r_S}{r} \right) c^2 dt^{*2} - 2 \frac{r_S}{r} c dt^* dr + \frac{r - r_S}{r} \left( \frac{r_S}{r - r_S} \right)^2 dr^2 - \left( \frac{r}{r - r_S} \right) dr^2,$$

**Fig. 3.2** Eddington's solution



so finally,

$$\underline{\underline{ds^2 = \frac{r - r_S}{r} c^2 dt^{*2} - 2 \frac{r_S}{r} c dt^* dr - \frac{r + r_S}{r} dr^2.}} \quad (3.90)$$

When we remove the restriction  $d\theta = d\varphi = 0$ , we finally get the full Schwarzschild's metric in Eddington's coordinates

$$\underline{\underline{ds^2 = \frac{r - r_S}{r} c^2 dt^{*2} - 2 \frac{r_S}{r} c dt^* dr - \frac{r + r_S}{r} dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\varphi^2).}} \quad (3.91)$$

This metric has, in fact, a singularity only at  $r = 0$ !

How will we see movements in the new  $(ct^*, r)$ -plane? For photons one has  $ds^2 = 0$ ; so, if we divide (3.90) by  $c^2 dt^{*2}$ , we receive

$$0 = \frac{r - r_S}{r} - 2 \frac{r_S}{r} \frac{dr}{c dt^*} - \frac{r + r_S}{r} \left( \frac{dr}{c dt^*} \right)^2.$$

This quadratic equation has the solutions

$$\frac{dr}{c dt^*} = \frac{-r_S \pm r}{r + r_S} = \frac{r - r_S}{r + r_S} \text{ and } -1, \quad (3.92)$$

or

$$\frac{dr}{dt^*} = \frac{-r_S \pm r}{r + r_S} c = \frac{r - r_S}{r + r_S} c \text{ and } -c. \quad (3.93)$$

The first solution provides for  $r = r_S$  the zero slope, i.e. a vertical line; for  $r < r_S$  the slope is negative, and it is positive for  $r > r_S$ . The second solution in the  $(ct^*, r)$ -plane is an inclined at  $45^\circ$  straight line. It shows that the velocity of an ingoing photon is constant and equal to  $c$ . Overall, one obtains Fig. 3.2.

## 3.5 Rotating Masses

### 3.5.1 Ansatz for the Metric Matrix $G$

From a non-rotating mass of radius  $R$  generated gravitation effect on the outside, i.e. for  $r > R$ , is described by the Schwarzschild's metric. A non-rigid rotating mass bulges out along its equator, so it cannot be described by a Schwarzschild's metric, which assumes a symmetric spherical mass. A rotating mass has in addition to the mass  $M$  also an angular momentum  $J$ , which is directly proportional to the angular velocity  $\omega$ .  $M$  and  $J$  are the only physical quantities for a rotating black hole that are required for the physical description of the gravitational field and how powerful the rotating mass is (there are possibly millions of solar masses in a black hole in the centre of a galaxy)! The Nobel-prize laureate Chandrasekhar has expressed it as:

Rotating black holes are the most perfect macroscopic objects in the universe. And as the theory of general relativity provides a single, unique, two-parameter solution for their description, they are also the simplest objects.

For the mathematical description, spatial polar coordinates  $r, \theta$  and  $\varphi$  are selected. The polar axis is the axis of rotation around which the body rotates with constant angular velocity  $\omega$ . The elements of the metric matrix  $G$  are allowed to depend neither on the angle  $\theta$  nor on time  $t$ . One can also use a priori, for reasons of symmetry, some matrix elements equal to zero. Because if one takes a time reflection  $t \rightarrow -t$ , this also changes the direction of rotation  $\omega = \frac{d\varphi}{dt} \rightarrow -\omega = -\frac{d\varphi}{dt} = \frac{d(-\varphi)}{dt}$ . But if we introduce two transformations  $t \rightarrow -t$  and  $\varphi \rightarrow -\varphi$  simultaneously, then the gravitational field is not allowed to change at all, also the metric matrix  $G$  is not changed. So that being the case, the matrix elements  $g_{t\theta}$  and  $g_{tr}$  must be zero because in the calculation of the metric here—with the indicated transformations— $dt d\theta \rightarrow -dt d\theta$  and  $dt dr \rightarrow -dt dr$ . And so one can start from the following the metric matrix which must always be symmetric:

$$G = \begin{pmatrix} g_{tt} & 0 & 0 & g_{t\varphi} \\ 0 & g_{rr} & g_{r\theta} & 0 \\ 0 & g_{r\theta} & g_{\theta\theta} & 0 \\ g_{t\varphi} & 0 & 0 & g_{\varphi\varphi} \end{pmatrix}.$$

The symmetric  $2 \times 2$ -matrix in the centre of the matrix  $G$  can be brought to diagonal form with the help of a similarity transformation so that, without loss of generality, one may start with the symmetric matrix:

$$G = \begin{pmatrix} g_{tt} & 0 & 0 & g_{t\varphi} \\ 0 & g_{rr} & 0 & 0 \\ 0 & 0 & g_{\theta\theta} & 0 \\ g_{t\varphi} & 0 & 0 & g_{\varphi\varphi} \end{pmatrix}.$$

### 3.5.2 Kerr's Solution in Boyer–Lindquist Coordinates

Without derivation we here directly give as solution the so-called Kerr's metric in Boyer–Lindquist coordinates:

$$\begin{aligned} ds^2 = & \left(1 - \frac{2mr}{\rho^2}\right) c^2 dt^2 + 4ma \frac{r \sin^2 \theta}{\rho^2} c dt d\varphi - \frac{\rho^2}{\Delta} dr^2 \\ & - \rho^2 d\theta^2 - \left(r^2 + a^2 + \frac{2mr}{r^2} a^2 \sin^2 \theta\right) \sin^2 \theta d\varphi^2, \end{aligned} \quad (3.94)$$

with

$$\begin{aligned} \Delta & \stackrel{\text{def}}{=} r^2 - 2mr + a^2, \\ \rho^2 & \stackrel{\text{def}}{=} r^2 + a^2 \cos^2 \theta, \\ m & \stackrel{\text{def}}{=} \frac{MG}{c^2} \end{aligned}$$

and the angular momentum (spin)  $J$  per unit mass

$$a \stackrel{\text{def}}{=} \frac{J}{m}.$$

As system parameters, in fact, only the two physical parameters occur: the modified mass,  $m$ , and the angular momentum per unit mass,  $a$ ! For  $J = 0$ , i.e. for a non-rotating mass, one obtains, of course, the Schwarzschild's solution.

A simpler form of the solution is obtained for  $M/r \ll 1$  and  $a/r \ll 1$ , i.e. for weak fields and slow rotation, namely

$$\begin{aligned} ds^2 \cong & \left(1 - \frac{2m}{r}\right) c^2 dt^2 + \frac{4J}{r} \sin^2 \theta c dt d\varphi - \left(1 + \frac{2m}{r}\right) dr^2 \\ & - r^2 (d\theta^2 + \sin^2 \theta d\varphi^2). \end{aligned}$$

### 3.5.3 The Lense–Thirring Effect

Two electric charges  $q_1$  and  $q_2$  of opposite polarity attract according to the Coulomb's law with the force

$$f_{\text{electr}} = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r_{12}^2} \frac{\mathbf{r}_{12}}{r_{12}}.$$

Here the vector  $\mathbf{r}_{12} \in \mathbb{R}^3$  shows from  $q_1$  to  $q_2$ . Newton's law of attraction between two masses  $m_1$  and  $m_2$  has almost the same shape, namely

$$\mathbf{f}_{\text{mech}} = G \frac{m_1 m_2}{r_{12}^2} \frac{\mathbf{r}_{12}}{r_{12}}.$$

Going over in both laws from the action at a distance theory to the theory of contiguous action, i.e. the field theory, with the electric field strength vector

$$\mathbf{e}_e \stackrel{\text{def}}{=} \frac{1}{4\pi\epsilon_0} \frac{q_1}{r_{12}^2} \frac{\mathbf{r}_{12}}{r_{12}}$$

one can also write the electric force as

$$\mathbf{f}_{\text{electr}} = \mathbf{e}_e q_2.$$

With the appropriate nomenclature for the mechanical parameters and the field strength

$$\mathbf{e}_m \stackrel{\text{def}}{=} G \frac{m_1}{r_{12}^2} \frac{\mathbf{r}_{12}}{r_{12}},$$

one finally obtains

$$\mathbf{f}_{\text{mech}} = \mathbf{e}_m m_2.$$

If an electric charge is moving, then there is in addition a magnetic field  $\mathbf{b}$ , which depends on the speed and the load of the charge. The magnetic field  $\mathbf{b}$  acts on the moving charge  $q$  having speed  $\mathbf{v}$ , generating, together with the electric field  $\mathbf{e}$ , the force:

$$\mathbf{f} = q \left( \mathbf{e} + \frac{1}{c} \mathbf{v} \times \mathbf{b} \right).$$

The question now is: Does a moving *mass* have a similar additional effect to another moving mass? This is indeed the case, for example, the precession of gyroscopes in the vicinity of large rotating masses like the Earth. This first was treated by Föppl [11]. The Austrian physicists Lense and Thirring calculated exactly this effect from the gravitational equations of Einstein in 1918.

One speaks in this context of a so-called *gravitomagnetic* field in analogy with classical electromagnetism. To this end, we first define with the help of the angular momentum  $\mathbf{J}$  of a spherical rotating mass the field

$$\mathbf{h}(\mathbf{r}) \stackrel{\text{def}}{=} -2 \frac{\mathbf{J} \times \mathbf{r}}{r^3} = -\frac{4GM R^2}{5c^3} \frac{\boldsymbol{\omega} \times \mathbf{r}}{r^3}, \quad (3.95)$$

where  $\boldsymbol{\omega}$  is the angular velocity of the rotating mass  $M$  with the radius  $R$ , and  $r > R$  is the distance from the centre of mass. Next we define the *gravitomagnetic field*:

$$\bar{\mathbf{h}} \stackrel{\text{def}}{=} \nabla \times \mathbf{h} = 2 \frac{\mathbf{J} - 3(\mathbf{r}^\top \mathbf{J})\mathbf{r}}{r^3} = \frac{2GM R^2}{5c^2} \frac{3(\boldsymbol{\omega}^\top \mathbf{r})\mathbf{r} - \boldsymbol{\omega} r^2}{r^5}. \quad (3.96)$$



The angular momentum here thus plays the same role as the magnetic dipole moment in electrodynamics, and the vector  $\mathbf{h}$  plays the same role as the vector potential. In summary, we obtain an analogue of the Lorentz force, namely

$$\underline{\underline{m \frac{d^2 \mathbf{x}}{dt^2} = m \left( \mathbf{e}_{\text{mech}} + \frac{d\mathbf{x}}{dt} \times \bar{\mathbf{h}} \right)}}. \quad (3.97)$$

In the case of stars orbiting close to a spinning, supermassive black hole, the Lense–Thirring effect should cause the star’s orbital plane to precess about the black hole’s spin axis. This effect should be detectable within the next few years via astrometric monitoring of stars at the centre of the Milky Way galaxy.

### 3.6 Summary of Results for the Gravitation of a Spherical Mass

The solution of Einstein’s field equation for the *outside* of a spherically symmetric, uniform, time-invariant mass distribution is this Schwarzschild’s metric (3.26):

$$ds^2 = \left(1 - \frac{2GM}{c^2 r}\right) c^2 dt^2 - \left(1 - \frac{2GM}{c^2 r}\right)^{-1} dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\varphi^2),$$

and the so-called Schwarzschild’s radius is (3.28):

$$r_S \stackrel{\text{def}}{=} \frac{2GM}{c^2}.$$

The Schwarzschild’s metric in matrix form is (3.25):

$$ds^2 = d\bar{\mathbf{x}}^\top \mathbf{G} d\bar{\mathbf{x}} = d\bar{\mathbf{x}}^\top \begin{pmatrix} 1 - \frac{r_S}{r} & 0 & 0 & 0 \\ 0 & -(1 - \frac{r_S}{r})^{-1} & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2 \sin^2 \theta \end{pmatrix} d\bar{\mathbf{x}},$$

with

$$d\bar{\mathbf{x}} \stackrel{\text{def}}{=} \begin{pmatrix} c dt \\ dr \\ d\theta \\ d\varphi \end{pmatrix}.$$

This changes the length (3.30) and time (3.31) according to:

$$dR = \left(1 - \frac{2GM}{c^2 r}\right)^{-1/2} dr,$$

and

$$d\tau = \left(1 - \frac{2GM}{c^2 r}\right)^{1/2} dt.$$

For the redshift of spectral lines the frequency ratio is obtained (3.32):

$$\frac{\nu_R}{\nu_T} = \left(\frac{1 - \frac{2GM}{c^2 r_T}}{1 - \frac{2GM}{c^2 r_R}}\right)^{1/2}.$$

The necessary Schwarzschild's density  $\rho_S$  for the existence of a *black hole* is

$$\begin{aligned} \rho_S &\stackrel{\text{def}}{=} \frac{M}{\frac{4}{3}\pi r_S^3} = \frac{M}{\frac{4}{3}\pi \left(\frac{2GM}{c^2}\right)^3} \\ &= \frac{3c^6}{32\pi G^3 M^2} = 2.33 \cdot 10^{71} \cdot M_{[\text{kg}]}^{-2} \left[\frac{\text{kg}}{\text{cm}^3}\right]. \end{aligned}$$

The Eddington's coordinates for a black hole are given with the new time coordinate (3.88)

$$ct^* \stackrel{\text{def}}{=} ct + r_S \ln|r - r_S|,$$

so

$$c dt^* = c dt + \frac{r_S}{r - r_S} dr.$$

The Schwarzschild's metric in Eddington's coordinates is (3.91):

$$ds^2 = \frac{r - r_S}{r} c^2 dt^{*2} - 2 \frac{r_S}{r} c dt^* dr - \frac{r + r_S}{r} dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\varphi^2).$$

The Lense–Thirring effect for a rotating mass generates a *gravitomagnetic* field in analogy with classical electromagnetism. With the help of the angular momentum  $\mathbf{J}$  of a spherical rotating mass, the field is (3.95):

$$\mathbf{h}(\mathbf{r}) \stackrel{\text{def}}{=} -2 \frac{\mathbf{J} \times \mathbf{r}}{r^3} = -\frac{4GMR^2}{5c^3} \frac{\boldsymbol{\omega} \times \mathbf{r}}{r^3},$$

where  $\boldsymbol{\omega}$  is the angular velocity of the rotating mass  $M$  with the radius  $R$ , and  $r > R$  is the distance from the centre of mass. Then the *gravitomagnetic field* is defined as (3.96):

$$\bar{\mathbf{h}} \stackrel{\text{def}}{=} \nabla \times \mathbf{h} = 2 \frac{\mathbf{J} - 3(\mathbf{r}^\top \mathbf{J})\mathbf{r}}{r^3} = \frac{2GMR^2}{5c^2} \frac{3(\boldsymbol{\omega}^\top \mathbf{r})\mathbf{r} - \omega r^2}{r^5}.$$

An analogue of the Lorentz force is obtained (3.97):

$$m \frac{d^2 \mathbf{x}}{dt^2} = m \left( \mathbf{e}_{\text{mech}} + \frac{d\mathbf{x}}{dt} \times \bar{\mathbf{h}} \right).$$

### 3.7 Concluding Remark

In the theory of *Special Relativity*, the effects are most clearly visible when the masses move quickly; however, in the theory of *General Relativity*, the effects are the greatest when the mass densities are very large and thus the spatial curvature is very pronounced.

# Appendix A

## Vectors and Matrices

### A.1 Vectors and Matrices

If the speed of a body is given, then its size and the direction need to be identified. For the description of such a directional quantity, *vectors* are used. These vectors in the three dimensional space require three components which, e.g. in a *column vector*, are summarized as follows:

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}. \tag{A.1}$$

A second possibility is to present a *transposed* column vector

$$\mathbf{v}^T = (v_1 \quad v_2 \quad v_3), \tag{A.2}$$

a *row vector*.

In another way, we get the concept of the vector when the following purely mathematical problem is considered: Find the solutions of the three coupled equations with four unknowns  $x_1$ ,  $x_2$ ,  $x_3$  and  $x_4$ :

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 = y_1, \tag{A.3}$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 = y_2, \tag{A.4}$$

$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 = y_3. \tag{A.5}$$

The four unknowns are summarized in the vector

$$\mathbf{x} \stackrel{\text{def}}{=} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}, \tag{A.6}$$

the variables  $y_1$ ,  $y_2$  and  $y_3$  form the vector

$$\mathbf{y} \stackrel{\text{def}}{=} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}, \quad (\text{A.7})$$

and the coefficients  $a_{ij}$  are included into the *matrix*

$$\mathbf{A} \stackrel{\text{def}}{=} \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}. \quad (\text{A.8})$$

With the two vectors  $\mathbf{x}$  and  $\mathbf{y}$  and the matrix  $\mathbf{A}$ , the system of equations can compactly be written as

$$\mathbf{Ax} = \mathbf{y}. \quad (\text{A.9})$$

If the two systems of equations

$$a_{11}x_1 + a_{12}x_2 = y_1, \quad (\text{A.10})$$

$$a_{21}x_1 + a_{22}x_2 = y_2 \quad (\text{A.11})$$

and

$$a_{11}z_1 + a_{12}z_2 = v_1, \quad (\text{A.12})$$

$$a_{21}z_1 + a_{22}z_2 = v_2 \quad (\text{A.13})$$

are added, one obtains

$$a_{11}(x_1 + z_1) + a_{12}(x_2 + z_2) = (y_1 + v_1), \quad (\text{A.14})$$

$$a_{21}(x_1 + z_1) + a_{22}(x_2 + z_2) = (y_2 + v_2). \quad (\text{A.15})$$

With the aid of vectors and matrices, the two systems of equations can be written as

$$\mathbf{Ax} = \mathbf{y} \quad \text{and} \quad \mathbf{Az} = \mathbf{v}. \quad (\text{A.16})$$

Adding the two equations in (A.16) is formally accomplished as

$$\mathbf{Ax} + \mathbf{Az} = \mathbf{A}(\mathbf{x} + \mathbf{z}) = \mathbf{y} + \mathbf{v}. \quad (\text{A.17})$$

A comparison of (A.17) with (A.14) and (A.15) suggests the following definition of the *addition of vectors*:

**Definition:**

$$\mathbf{y} + \mathbf{v} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \stackrel{\text{def}}{=} \begin{pmatrix} y_1 + v_1 \\ y_2 + v_2 \\ \vdots \\ y_n + v_n \end{pmatrix}. \quad (\text{A.18})$$

Accordingly, the product of a vector and a real or complex number  $c$  is defined by

**Definition:**

$$c \cdot \mathbf{x} \stackrel{\text{def}}{=} \begin{pmatrix} c \cdot x_1 \\ \vdots \\ c \cdot x_n \end{pmatrix}. \quad (\text{A.19})$$

## A.2 Matrices

### A.2.1 Types of Matrices

In the first section, the concept of the matrix has been introduced.

**Definition:** If a matrix  $\mathbf{A}$  has  $n$  rows and  $m$  columns, it is called an  $n \times m$  matrix and denoted  $\mathbf{A} \in \mathbb{R}^{n \times m}$ .

**Definition:** If  $\mathbf{A}$  (with the elements  $a_{ij}$ ) is an  $n \times m$  matrix, then the *transpose* of  $\mathbf{A}$ , denoted by  $\mathbf{A}^\top$ , is the  $m \times n$  matrix with the elements  $a_{ij}^\top = a_{ji}$ .

So the matrix (A.8) has the matrix transpose

$$\mathbf{A}^\top = \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \\ a_{14} & a_{24} & a_{34} \end{pmatrix}. \quad (\text{A.20})$$

In a *square matrix*, one has  $n = m$ ; and in an  $n \times n$  *diagonal matrix*, all the elements  $a_{ij}$ ,  $i \neq j$ , outside the main diagonal are equal to zero. An *identity matrix*  $\mathbf{I}$  is a diagonal matrix where all elements on the main diagonal are equal to one. An  $r \times r$  identity matrix is denoted by  $\mathbf{I}_r$ . If a transposed matrix  $\mathbf{A}^\top$  is equal to the original matrix  $\mathbf{A}$ , such a matrix is called *symmetric*. In this case,  $a_{ij} = a_{ji}$ .

### A.2.2 Matrix Operations

If the two systems of equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= y_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m &= y_2, \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m &= y_n \end{aligned}$$

and

$$\begin{aligned} b_{11}x_1 + b_{12}x_2 + \cdots + b_{1m}x_m &= z_1, \\ b_{21}x_1 + b_{22}x_2 + \cdots + b_{2m}x_m &= z_2, \\ &\vdots \\ b_{n1}x_1 + b_{n2}x_2 + \cdots + b_{nm}x_m &= z_n \end{aligned}$$

are added, one obtains

$$\begin{aligned} (a_{11} + b_{11})x_1 + (a_{12} + b_{12})x_2 + \cdots + (a_{1m} + b_{1m})x_m &= (y_1 + z_1), \\ (a_{21} + b_{21})x_1 + (a_{22} + b_{22})x_2 + \cdots + (a_{2m} + b_{2m})x_m &= (y_2 + z_2), \\ &\vdots \\ (a_{n1} + b_{n1})x_1 + (a_{n2} + b_{n2})x_2 + \cdots + (a_{nm} + b_{nm})x_m &= (y_n + z_n), \end{aligned}$$

or, in vector–matrix notation, with

$$\mathbf{Ax} = \mathbf{y} \quad \text{and} \quad \mathbf{Bx} = \mathbf{z},$$

the same can also be written symbolically as

$$(\mathbf{A} + \mathbf{B})\mathbf{x} = \mathbf{y} + \mathbf{z}. \tag{A.21}$$

A comparison of the last equations suggests the following definition:

**Definition:** The *sum* of two  $n \times m$  matrices  $\mathbf{A}$  and  $\mathbf{B}$  is defined by

$$\begin{aligned} \mathbf{A} + \mathbf{B} &= \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nm} \end{pmatrix} \\ &\stackrel{\text{def}}{=} \begin{pmatrix} (a_{11} + b_{11}) & \cdots & (a_{1m} + b_{1m}) \\ \vdots & & \vdots \\ (a_{n1} + b_{n1}) & \cdots & (a_{nm} + b_{nm}) \end{pmatrix}. \end{aligned} \tag{A.22}$$

The sum of two matrices can only be formed when both matrices have the same number of rows and the same number of columns.

If the relations

$$\mathbf{y} = \mathbf{A}\mathbf{x} \quad \text{and} \quad \mathbf{x} = \mathbf{B}\mathbf{z} \quad (\text{A.23})$$

are given, what is the connection between the two vectors  $\mathbf{y}$  and  $\mathbf{z}$ ? One may write

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= y_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m &= y_2, \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m &= y_n \end{aligned}$$

and

$$\begin{aligned} b_{11}z_1 + b_{12}z_2 + \cdots + b_{1\ell}z_\ell &= x_1, \\ b_{21}z_1 + b_{22}z_2 + \cdots + b_{2\ell}z_\ell &= x_2, \\ &\vdots \\ b_{m1}z_1 + b_{m2}z_2 + \cdots + b_{m\ell}z_\ell &= x_m, \end{aligned}$$

then one obtains, by inserting the  $x_i$ 's from the latter system of equations into the former system,

$$\begin{aligned} a_{11}(b_{11}z_1 + \cdots + b_{1\ell}z_\ell) + \cdots + a_{1m}(b_{m1}z_1 + \cdots + b_{m\ell}z_\ell) &= y_1, \\ a_{21}(b_{11}z_1 + \cdots + b_{1\ell}z_\ell) + \cdots + a_{2m}(b_{m1}z_1 + \cdots + b_{m\ell}z_\ell) &= y_2, \\ &\vdots \\ a_{n1}(b_{11}z_1 + \cdots + b_{1\ell}z_\ell) + \cdots + a_{nm}(b_{m1}z_1 + \cdots + b_{m\ell}z_\ell) &= y_n. \end{aligned}$$

Combining the terms with  $z_i$ , we obtain

$$\begin{aligned} (a_{11}b_{11} + \cdots + a_{1m}b_{m1})z_1 + \cdots + (a_{11}b_{1\ell} + \cdots + a_{1m}b_{m\ell})z_\ell &= y_1, \\ (a_{21}b_{11} + \cdots + a_{2m}b_{m1})z_1 + \cdots + (a_{21}b_{1\ell} + \cdots + a_{2m}b_{m\ell})z_\ell &= y_2, \\ &\vdots \\ (a_{n1}b_{11} + \cdots + a_{nm}b_{m1})z_1 + \cdots + (a_{n1}b_{1\ell} + \cdots + a_{nm}b_{m\ell})z_\ell &= y_n. \end{aligned}$$

If, on the other hand, we formally insert the right-hand side of (A.23) into the left equation, we obtain

$$\mathbf{y} = \mathbf{A}\mathbf{B}\mathbf{z} \stackrel{\text{def}}{=} \mathbf{C}\mathbf{z}. \quad (\text{A.24})$$



**Definition:** The *product* of an  $n \times m$  matrix  $A$  with an  $m \times \ell$  matrix  $B$  is the  $n \times \ell$  matrix  $C$  with the matrix elements

$$c_{ij} = \sum_{k=1}^m a_{ik} b_{kj}, \quad (\text{A.25})$$

for  $i = 1, 2, \dots, n$  and  $j = 1, 2, \dots, \ell$ .

The element  $c_{ij}$  of the product matrix  $C$  is obtained by multiplying the elements of the  $i$ th row of the first matrix  $A$  with the elements of the  $j$ th column of the second matrix  $B$  and adding these products. It follows that the number of columns of the first matrix must be equal to the number of rows of the second matrix, so that the matrix multiplication can be executed at all. The product matrix has as many rows as the first matrix and as many columns as the second matrix. It follows that, in general,  $AB \neq BA$ .

We get another matrix operation through the following problem. In

$$Ax = b, \quad (\text{A.26})$$

the  $3 \times 3$  matrix  $A$  and the  $3 \times 1$  vector  $b$  shall be given. Wanted is the  $3 \times 1$  vector  $x$  that satisfies the system of equations (A.26). Written out this is the linear system of equations:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1,$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2,$$

$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3.$$

Denoting the determinant of the square matrix  $A$  by  $\det(A)$ , the solutions are obtained by using Cramer's rule

$$x_1 = \frac{1}{\det(A)} \det \begin{pmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{pmatrix}, \quad (\text{A.27})$$

$$x_2 = \frac{1}{\det(A)} \det \begin{pmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{pmatrix}, \quad (\text{A.28})$$

$$x_3 = \frac{1}{\det(A)} \det \begin{pmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{pmatrix}. \quad (\text{A.29})$$

If we develop in (A.27) the determinant in the numerator with respect to the first column, we obtain

$$\begin{aligned} x_1 &= \frac{1}{\det(\mathbf{A})} \left( b_1 \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - b_2 \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} + b_3 \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix} \right) \\ &= \frac{1}{\det(\mathbf{A})} (b_1 A_{11} + b_2 A_{21} + b_3 A_{31}) \\ &= \frac{1}{\det(\mathbf{A})} (A_{11} \ A_{21} \ A_{31}) \mathbf{b}. \end{aligned} \quad (\text{A.30})$$

Accordingly, we obtain from (A.28) and (A.29)

$$x_2 = \frac{1}{\det(\mathbf{A})} (A_{12} \ A_{22} \ A_{32}) \mathbf{b} \quad (\text{A.31})$$

and

$$x_3 = \frac{1}{\det(\mathbf{A})} (A_{13} \ A_{23} \ A_{33}) \mathbf{b}. \quad (\text{A.32})$$

Here the *adjuncts*  $A_{ij}$  are the determinants which are obtained when the  $i$ th row and the  $j$ th column of the matrix  $\mathbf{A}$  are removed, and from the remaining matrix the determinant is computed and this is multiplied by the factor  $(-1)^{i+j}$ .

**Definition:** The adjuncts are summarized in the *adjoint matrix*

$$\text{adj}(\mathbf{A}) \stackrel{\text{def}}{=} \begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix}. \quad (\text{A.33})$$

With this matrix, the three equations (A.30) to (A.32) can be written as one equation

$$\mathbf{x} = \frac{\text{adj}(\mathbf{A})}{\det(\mathbf{A})} \mathbf{b}. \quad (\text{A.34})$$

**Definition:** The  $n \times n$  matrix (whenever  $\det(\mathbf{A}) \neq 0$ )

$$\mathbf{A}^{-1} \stackrel{\text{def}}{=} \frac{\text{adj}(\mathbf{A})}{\det(\mathbf{A})} \quad (\text{A.35})$$

is called the *inverse matrix* of the square  $n \times n$  matrix  $\mathbf{A}$ .

For a matrix product, the inverse matrix is obtained as

$$(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} \quad (\text{A.36})$$

because

$$(\mathbf{A}\mathbf{B})(\mathbf{B}^{-1}\mathbf{A}^{-1}) = \mathbf{A}(\mathbf{B}\mathbf{B}^{-1})\mathbf{A}^{-1} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}.$$

### A.2.3 Block Matrices

Often large matrices have a certain structure, e.g. when one or more sub-arrays are zero matrices. On the other hand, one can make a *block matrix* from each matrix by drawing vertical and horizontal lines. For a system of equations, one then, for example, obtains

$$\left( \begin{array}{c|c|c|c} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1n} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2n} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \mathbf{A}_{m1} & \mathbf{A}_{m2} & \cdots & \mathbf{A}_{mn} \end{array} \right) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}. \quad (\text{A.37})$$

The  $A_{ij}$ 's are called *sub-matrices* and the vectors  $x_i$  and  $y_i$  *sub-vectors*. For two appropriately partitioned block matrices, the product may be obtained by simply carrying out the multiplication as if the sub-matrices were themselves elements:

$$\left( \begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} \end{array} \right) \left( \begin{array}{c|c} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \hline \mathbf{B}_{21} & \mathbf{B}_{22} \end{array} \right) = \left( \begin{array}{c|c} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \hline \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{array} \right).$$

In particular, the subdivision into blocks of matrices is of benefit for the calculation of the inverse matrix. Looking at the system of equations

$$\mathbf{A}x_1 + \mathbf{B}x_2 = y_1, \quad (\text{A.38})$$

$$\mathbf{C}x_1 + \mathbf{D}x_2 = y_2, \quad (\text{A.39})$$

or combined into the form

$$\left( \begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array} \right) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad (\text{A.40})$$

the inverse of the matrix

$$\mathbf{M} = \left( \begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array} \right) \quad (\text{A.41})$$

can be expressed by more easily calculated inverse sub-matrices. When the matrix  $\mathbf{A}$  is invertible, one gets from (A.38)

$$x_1 = \mathbf{A}^{-1}y_1 - \mathbf{A}^{-1}\mathbf{B}x_2. \quad (\text{A.42})$$

This, used in (A.39), yields

$$y_2 = \mathbf{C}\mathbf{A}^{-1}y_1 - (\mathbf{C}\mathbf{A}^{-1}\mathbf{B} - \mathbf{D})x_2 \quad (\text{A.43})$$

and, solving for  $x_2$ ,

$$x_2 = (\mathbf{C}\mathbf{A}^{-1}\mathbf{B} - \mathbf{D})^{-1}(\mathbf{C}\mathbf{A}^{-1}y_1 - y_2). \quad (\text{A.44})$$

Equation (A.44) used in (A.42) yields

$$\mathbf{x}_1 = [\mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}\mathbf{A}^{-1}\mathbf{B} - \mathbf{D})^{-1}\mathbf{C}\mathbf{A}^{-1}]\mathbf{y}_1 + \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}\mathbf{A}^{-1}\mathbf{B} - \mathbf{D})^{-1}\mathbf{y}_2. \quad (\text{A.45})$$

Thus, the solution of the system of equations (A.40) is obtained, namely

$$\begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} = \mathbf{M}^{-1} \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} \quad (\text{A.46})$$

with

$$\mathbf{M}^{-1} = \left( \begin{array}{c|c} \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}\mathbf{A}^{-1}\mathbf{B} - \mathbf{D})^{-1}\mathbf{C}\mathbf{A}^{-1} & \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}\mathbf{A}^{-1}\mathbf{B} - \mathbf{D})^{-1} \\ \hline (\mathbf{C}\mathbf{A}^{-1}\mathbf{B} - \mathbf{D})^{-1}\mathbf{C}\mathbf{A}^{-1} & -(\mathbf{C}\mathbf{A}^{-1}\mathbf{B} - \mathbf{D})^{-1} \end{array} \right), \quad (\text{A.47})$$

and the inverse matrix of  $\mathbf{M}$  can be calculated by using the inverse matrices of the smaller sub-matrices  $\mathbf{A}$  and  $(\mathbf{C}\mathbf{A}^{-1}\mathbf{B} - \mathbf{D})$ . If the sub-matrix  $\mathbf{D}$  is invertible, (A.39) can be solved for  $\mathbf{x}_2$ , and then, in a similar way, also the inverse matrix of  $\mathbf{M}$  can be calculated. One then obtains a different form of the inverted matrix, namely

$$\mathbf{M}^{-1} = \left( \begin{array}{c|c} -(\mathbf{B}\mathbf{D}^{-1}\mathbf{C} - \mathbf{A})^{-1} & (\mathbf{B}\mathbf{D}^{-1}\mathbf{C} - \mathbf{A})^{-1}\mathbf{B}\mathbf{D}^{-1} \\ \hline \mathbf{D}^{-1}\mathbf{C}(\mathbf{B}\mathbf{D}^{-1}\mathbf{C} - \mathbf{A})^{-1} & \mathbf{D}^{-1} - \mathbf{D}^{-1}\mathbf{C}(\mathbf{B}\mathbf{D}^{-1}\mathbf{C} - \mathbf{A})^{-1}\mathbf{B}\mathbf{D}^{-1} \end{array} \right). \quad (\text{A.48})$$

So there are two different results available for the same matrix. It follows that the corresponding sub-matrices must be equal. From the comparison of the *northwestern* blocks, if  $\mathbf{D}$  is replaced by  $-\mathbf{D}$ , the known matrix-inversion lemma follows:

$$\boxed{(\mathbf{A} + \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}\mathbf{A}^{-1}\mathbf{B} + \mathbf{D})^{-1}\mathbf{C}\mathbf{A}^{-1}.} \quad (\text{A.49})$$

A special case occurs when we have a block-triangular matrix, for example, if  $\mathbf{C} = \mathbf{O}$ . Then we obtain

$$\left( \begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{O} & \mathbf{D} \end{array} \right)^{-1} = \left( \begin{array}{c|c} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{D}^{-1} \\ \hline \mathbf{O} & \mathbf{D}^{-1} \end{array} \right). \quad (\text{A.50})$$

## A.3 The Kronecker-Product

### A.3.1 Definitions

**Definition:** The *Kronecker-product* of two matrices  $A \in \mathbb{R}^{n \times m}$  and  $B \in \mathbb{R}^{p \times q}$  is a matrix  $C \in \mathbb{R}^{np \times mq}$ , denoted as

$$A \otimes B = C.$$

Here, the sub-matrix  $C_{ij} \in \mathbb{R}^{p \times q}$ , for  $i = 1$  to  $n$  and  $j = 1$  to  $m$ , is defined by

$$C_{ij} \stackrel{\text{def}}{=} a_{ij} B,$$

so that, overall, the matrix  $C$  has the form

$$C = \begin{pmatrix} a_{11} B & a_{12} B & \dots & a_{1m} B \\ a_{21} B & a_{22} B & \dots & a_{2m} B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} B & a_{n2} B & \dots & a_{nm} B \end{pmatrix}.$$

The matrix elements of the product matrix  $C$  can be directly calculated using the following formula

$$c_{i,j} = a_{\lfloor \frac{i-1}{p} \rfloor + 1, \lfloor \frac{j-1}{q} \rfloor + 1} \cdot b_{i - \lfloor \frac{i-1}{p} \rfloor p, j - \lfloor \frac{j-1}{q} \rfloor q},$$

where  $\lfloor x \rfloor$  is the integer part of  $x$ .

**Definition:** If a matrix  $A$  is composed of the  $m$  columns  $a_i \in \mathbb{C}^n$ ,

$$A = (a_1 \quad a_2 \quad \dots \quad a_m),$$

the *vec*-operator is defined as follows:

$$\text{vec}(A) \stackrel{\text{def}}{=} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} \in \mathbb{C}^{nm}.$$

### A.3.2 Some Theorems

The following is very interesting:

$$\text{Lemma: } \text{vec}(AXB) = (B^T \otimes A) \text{vec}(X). \quad (\text{A.51})$$

*Proof* Let  $\mathbf{B} \in \mathbb{C}^{n \times m}$ , then one has

$$\begin{aligned} \mathbf{A}\mathbf{X}\mathbf{B} &= (\mathbf{A}\mathbf{x}_1 \quad \mathbf{A}\mathbf{x}_2 \quad \dots \quad \mathbf{A}\mathbf{x}_n) \mathbf{B} \\ &= (\mathbf{A}\mathbf{x}_1 \quad \mathbf{A}\mathbf{x}_2 \quad \dots \quad \mathbf{A}\mathbf{x}_n) (\mathbf{b}_1 \quad \mathbf{b}_2 \quad \dots \quad \mathbf{b}_m) \\ &= \begin{pmatrix} (b_{11}\mathbf{A}\mathbf{x}_1 + b_{21}\mathbf{A}\mathbf{x}_2 + \dots + b_{n1}\mathbf{A}\mathbf{x}_n) & \dots \\ (b_{1m}\mathbf{A}\mathbf{x}_1 + b_{2m}\mathbf{A}\mathbf{x}_2 + \dots + b_{nm}\mathbf{A}\mathbf{x}_n) & \dots \end{pmatrix}. \end{aligned}$$

Applying the *vec*-operator to the last equation, we obtain

$$\begin{aligned} \mathit{vec}(\mathbf{A}\mathbf{X}\mathbf{B}) &= \begin{pmatrix} (b_{11}\mathbf{A}\mathbf{x}_1 + b_{21}\mathbf{A}\mathbf{x}_2 + \dots + b_{n1}\mathbf{A}\mathbf{x}_n) \\ \vdots \\ (b_{1m}\mathbf{A}\mathbf{x}_1 + b_{2m}\mathbf{A}\mathbf{x}_2 + \dots + b_{nm}\mathbf{A}\mathbf{x}_n) \end{pmatrix} \\ &= \begin{pmatrix} b_{11}\mathbf{A} & \dots & b_{n1}\mathbf{A} \\ \vdots & \ddots & \vdots \\ b_{1m}\mathbf{A} & \dots & b_{nm}\mathbf{A} \end{pmatrix} \mathit{vec}(\mathbf{X}) \\ &= (\mathbf{B}^\top \otimes \mathbf{A})\mathit{vec}(\mathbf{X}). \quad \square \end{aligned}$$

As corollaries of this lemma, we get the following results:

$$\boxed{\mathit{vec}(\mathbf{A}\mathbf{X}) = (\mathbf{I} \otimes \mathbf{A})\mathit{vec}(\mathbf{X}).} \quad (\text{A.52})$$

*Proof* Set  $\mathbf{B} = \mathbf{I}$  in the lemma. □

$$\boxed{\mathit{vec}(\mathbf{X}\mathbf{B}) = (\mathbf{B}^\top \otimes \mathbf{I})\mathit{vec}(\mathbf{X}).} \quad (\text{A.53})$$

*Proof* Set  $\mathbf{A} = \mathbf{I}$  in the lemma. □

$$\boxed{\mathit{vec}(\mathbf{b}\mathbf{a}^\top) = (\mathbf{a} \otimes \mathbf{b}).} \quad (\text{A.54})$$

*Proof* Simply write  $\mathit{vec}(\mathbf{b}\mathbf{a}^\top) = \mathit{vec}(\mathbf{b}\mathbf{1}\mathbf{a}^\top) = (\mathbf{a} \otimes \mathbf{b})\mathit{vec}(\mathbf{1}) = \mathbf{a} \otimes \mathbf{b}$ . □

### A.3.3 The Permutation Matrix $U_{p \times q}$

**Definition:** The *permutation matrix*

$$U_{p \times q} \stackrel{\text{def}}{=} \sum_i^p \sum_k^q E_{ik}^{p \times q} \otimes E_{ki}^{q \times p} \in \mathbb{R}^{pq \times qp} \quad (\text{A.55})$$

has just one 1 in each column and each row. In the formation of matrix

$$E_{ik}^{p \times q} \stackrel{\text{def}}{=} \mathbf{e}_i \mathbf{e}_k^\top, \quad (\text{A.56})$$

$\mathbf{e}_i$  is the  $i$ th column of  $\mathbf{I}_p$  and  $\mathbf{e}_k$  is the  $k$ th column of  $\mathbf{I}_q$ . However, only the matrix element  $E_{ik} = 1$ ; the other matrix elements are zeros.

For example, the permutation matrix  $U_{4 \times 4} \in \mathbb{R}^{16 \times 16}$ —often used in this book—has the form

$$U_{4 \times 4} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}. \quad (\text{A.57})$$

Permutation matrices have the following characteristics [4]:

$$\mathbf{U}_{p \times q}^\top = \mathbf{U}_{q \times p}, \quad (\text{A.58})$$

$$\mathbf{U}_{p \times q}^{-1} = \mathbf{U}_{q \times p}, \quad (\text{A.59})$$

$$\mathbf{U}_{p \times 1} = \mathbf{U}_{1 \times p} = \mathbf{I}_p, \quad (\text{A.60})$$

$$\mathbf{U}_{n \times n} = \mathbf{U}_{n \times n}^\top = \mathbf{U}_{n \times n}^{-1}. \quad (\text{A.61})$$

Permutation matrices are mainly used to change the order of the factors in a Kronecker-product because

$$\mathbf{U}_{s \times p}(\mathbf{B} \otimes \mathbf{A})\mathbf{U}_{q \times t} = \mathbf{A} \otimes \mathbf{B} \quad \text{if } \mathbf{A} \in \mathbb{R}^{p \times q} \text{ and } \mathbf{B} \in \mathbb{R}^{s \times t}. \quad (\text{A.62})$$

### A.3.4 More Properties of the Kronecker-Product

The following important properties are listed also without proof (see [4]):

$$(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C}), \quad (\text{A.63})$$

$$(\mathbf{A} \otimes \mathbf{B})^\top = \mathbf{A}^\top \otimes \mathbf{B}^\top, \quad (\text{A.64})$$

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}. \quad (\text{A.65})$$

## A.4 Derivatives of Vectors/Matrices with Respect to Vectors/Matrices

### A.4.1 Definitions

**Definition:** The derivative of a matrix  $\mathbf{A} \in \mathbb{R}^{n \times m}$  with respect to a matrix  $\mathbf{M} \in \mathbb{R}^{r \times s}$  is defined as follows:

$$\frac{\partial \mathbf{A}}{\partial \mathbf{M}} \stackrel{\text{def}}{=} \begin{pmatrix} \frac{\partial \mathbf{A}}{\partial M_{11}} & \frac{\partial \mathbf{A}}{\partial M_{12}} & \cdots & \frac{\partial \mathbf{A}}{\partial M_{1s}} \\ \frac{\partial \mathbf{A}}{\partial M_{21}} & \frac{\partial \mathbf{A}}{\partial M_{22}} & \cdots & \frac{\partial \mathbf{A}}{\partial M_{2s}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \mathbf{A}}{\partial M_{r1}} & \frac{\partial \mathbf{A}}{\partial M_{r2}} & \cdots & \frac{\partial \mathbf{A}}{\partial M_{rs}} \end{pmatrix} \in \mathbb{R}^{nr \times ms}. \quad (\text{A.66})$$

With the  $r \times s$ -operator

$$\frac{\partial}{\partial \mathbf{M}} \stackrel{\text{def}}{=} \begin{pmatrix} \frac{\partial}{\partial M_{11}} & \frac{\partial}{\partial M_{12}} & \cdots & \frac{\partial}{\partial M_{1s}} \\ \frac{\partial}{\partial M_{21}} & \frac{\partial}{\partial M_{22}} & \cdots & \frac{\partial}{\partial M_{2s}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial M_{r1}} & \frac{\partial}{\partial M_{r2}} & \cdots & \frac{\partial}{\partial M_{rs}} \end{pmatrix} \quad (\text{A.67})$$

the definition of the derivative (A.66) is also written as

$$\frac{\partial \mathbf{A}}{\partial \mathbf{M}} \stackrel{\text{def}}{=} \frac{\partial}{\partial \mathbf{M}} \otimes \mathbf{A}. \quad (\text{A.68})$$



Thus one can show that

$$\left(\frac{\partial \mathbf{A}}{\partial \mathbf{M}}\right)^\top = \left(\frac{\partial}{\partial \mathbf{M}} \otimes \mathbf{A}\right)^\top = \left(\left(\frac{\partial}{\partial \mathbf{M}}\right)^\top \otimes \mathbf{A}^\top\right) = \frac{\partial \mathbf{A}^\top}{\partial \mathbf{M}^\top}. \quad (\text{A.69})$$

For the derivatives of vectors with respect to vectors one has:

$$\frac{\partial \mathbf{f}^\top}{\partial \mathbf{p}} \stackrel{\text{def}}{=} \frac{\partial}{\partial \mathbf{p}} \otimes \mathbf{f}^\top = \begin{pmatrix} \frac{\partial \mathbf{f}^\top}{\partial p_1} \\ \frac{\partial \mathbf{f}^\top}{\partial p_2} \\ \vdots \\ \frac{\partial \mathbf{f}^\top}{\partial p_r} \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial p_1} & \frac{\partial f_2}{\partial p_1} & \cdots & \frac{\partial f_n}{\partial p_1} \\ \frac{\partial f_1}{\partial p_2} & \frac{\partial f_2}{\partial p_2} & \cdots & \frac{\partial f_n}{\partial p_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_1}{\partial p_r} & \frac{\partial f_2}{\partial p_r} & \cdots & \frac{\partial f_n}{\partial p_r} \end{pmatrix} \in \mathbb{R}^{r \times n} \quad (\text{A.70})$$

and

$$\frac{\partial \mathbf{f}}{\partial \mathbf{p}^\top} \stackrel{\text{def}}{=} \frac{\partial}{\partial \mathbf{p}^\top} \otimes \mathbf{f} = \left(\frac{\partial}{\partial \mathbf{p}} \otimes \mathbf{f}^\top\right)^\top = \left(\frac{\partial \mathbf{f}^\top}{\partial \mathbf{p}}\right)^\top \in \mathbb{R}^{n \times r}. \quad (\text{A.71})$$

### A.4.2 Product Rule

Let  $\mathbf{A} = \mathbf{A}(\alpha)$  and  $\mathbf{B} = \mathbf{B}(\alpha)$ . Then we obviously have

$$\frac{\partial(\mathbf{A}\mathbf{B})}{\partial \alpha} = \frac{\partial \mathbf{A}}{\partial \alpha} \mathbf{B} + \mathbf{A} \frac{\partial \mathbf{B}}{\partial \alpha}. \quad (\text{A.72})$$

In addition, (A.66) can be rewritten as:

$$\frac{\partial \mathbf{A}}{\partial \mathbf{M}} = \sum_{i,k} \mathbf{E}_{ik}^{s \times t} \otimes \frac{\partial \mathbf{A}}{\partial m_{ik}}, \quad \mathbf{M} \in \mathbb{R}^{s \times t}. \quad (\text{A.73})$$

Using (A.72) and (A.73), this product rule can be derived:

$$\begin{aligned} \underline{\underline{\frac{\partial(\mathbf{A}\mathbf{B})}{\partial \mathbf{M}}}} &= \sum_{i,k} \mathbf{E}_{ik}^{s \times t} \otimes \frac{\partial(\mathbf{A}\mathbf{B})}{\partial m_{ik}} = \sum_{i,k} \mathbf{E}_{ik}^{s \times t} \otimes \left(\frac{\partial \mathbf{A}}{\partial m_{ik}} \mathbf{B} + \mathbf{A} \frac{\partial \mathbf{B}}{\partial m_{ik}}\right) \\ &= \left(\frac{\partial}{\partial \mathbf{M}} \otimes \mathbf{A}\right) (\mathbf{I}_t \otimes \mathbf{B}) + (\mathbf{I}_s \otimes \mathbf{A}) \left(\frac{\partial}{\partial \mathbf{M}} \otimes \mathbf{B}\right) \\ &= \underline{\underline{\frac{\partial \mathbf{A}}{\partial \mathbf{M}} (\mathbf{I}_t \otimes \mathbf{B}) + (\mathbf{I}_s \otimes \mathbf{A}) \frac{\partial \mathbf{B}}{\partial \mathbf{M}}}}. \end{aligned} \quad (\text{A.74})$$

### A.4.3 Chain Rule

When a matrix  $\mathbf{A} \in \mathbb{R}^{n \times m}$  is a function of a matrix  $\mathbf{B} \in \mathbb{R}^{k \times \ell}$  which is again a function of a matrix  $\mathbf{M} \in \mathbb{R}^{r \times s}$ , then the chain rule is valid [4]:

$$\begin{aligned} \frac{\partial}{\partial \mathbf{M}} \mathbf{A}(\mathbf{B}(\mathbf{M})) &= \left( \mathbf{I}_r \otimes \frac{\partial \mathbf{A}}{\partial (\mathbf{vec}(\mathbf{B}^\top))^\top} \right) \left( \frac{\partial \mathbf{vec}(\mathbf{B}^\top)}{\partial \mathbf{M}} \otimes \mathbf{I}_m \right) \\ &= \left( \frac{\partial (\mathbf{vec}(\mathbf{B}))^\top}{\partial \mathbf{M}} \otimes \mathbf{I}_n \right) \left( \mathbf{I}_s \otimes \frac{\partial \mathbf{A}}{\partial \mathbf{vec}(\mathbf{B})} \right). \end{aligned} \quad (\text{A.75})$$

A special case of this is

$$\frac{d\mathbf{A}(\mathbf{x}(t))}{dt} = \frac{\partial \mathbf{A}}{\partial \mathbf{x}^\top} \left( \frac{d\mathbf{x}}{dt} \otimes \mathbf{I}_m \right) = \left( \frac{d\mathbf{x}^\top}{dt} \otimes \mathbf{I}_n \right) \frac{\partial \mathbf{A}}{\partial \mathbf{x}} \in \mathbb{R}^{n \times m}. \quad (\text{A.76})$$

## A.5 Differentiation with Respect to Time

### A.5.1 Differentiation of a Function with Respect to Time

Suppose a function  $a$  that depends on the three space variables  $x_1$ ,  $x_2$  and  $x_3$  is given. The local space variables themselves are, in turn, dependent on the time parameter  $t$ . So it is

$$a = a(\mathbf{x}(t)), \quad (\text{A.77})$$

if the three space variables are summarized in the vector  $\mathbf{x}$ .

We want to find the velocity

$$\dot{a} = \frac{da}{dt}. \quad (\text{A.78})$$

To determine this velocity, we first define the total difference

$$\Delta a \stackrel{\text{def}}{=} \frac{\partial a}{\partial x_1} \Delta x_1 + \frac{\partial a}{\partial x_2} \Delta x_2 + \frac{\partial a}{\partial x_3} \Delta x_3. \quad (\text{A.79})$$

After division by  $\Delta t$ , in the limit  $\Delta t \rightarrow 0$ , one has

$$\dot{a} = \frac{da}{dt} = \lim_{\Delta t \rightarrow 0} \frac{\Delta a}{\Delta t} = \frac{\partial a}{\partial x_1} \dot{x}_1 + \frac{\partial a}{\partial x_2} \dot{x}_2 + \frac{\partial a}{\partial x_3} \dot{x}_3. \quad (\text{A.80})$$

The right-hand side of this equation can be presented by the scalar product of the column vectors  $\dot{\mathbf{x}}$  and  $\frac{\partial a}{\partial \mathbf{x}}$  in two ways, namely

$$\dot{a} = \dot{\mathbf{x}}^\top \frac{\partial a}{\partial \mathbf{x}} = \frac{\partial a}{\partial \mathbf{x}^\top} \dot{\mathbf{x}}. \quad (\text{A.81})$$

### A.5.2 Differentiation of a Vector with Respect to Time

If two functions  $a_1$  and  $a_2$  are given, they have the same dependence on time  $t$  as  $a(t)$  in (A.77), and are summarized in the column vector

$$\mathbf{a} \stackrel{\text{def}}{=} \begin{pmatrix} a_1(\mathbf{x}(t)) \\ a_2(\mathbf{x}(t)) \end{pmatrix}, \quad (\text{A.82})$$

we obtain initially for the derivative with respect to time using (A.81)

$$\dot{\mathbf{a}} = \begin{pmatrix} \dot{a}_1 \\ \dot{a}_2 \end{pmatrix} = \begin{pmatrix} \dot{\mathbf{x}}^\top \frac{\partial a_1}{\partial \mathbf{x}} \\ \dot{\mathbf{x}}^\top \frac{\partial a_2}{\partial \mathbf{x}} \end{pmatrix} = \begin{pmatrix} \frac{\partial a_1}{\partial \mathbf{x}^\top} \dot{\mathbf{x}} \\ \frac{\partial a_2}{\partial \mathbf{x}^\top} \dot{\mathbf{x}} \end{pmatrix}. \quad (\text{A.83})$$

The next-to-last vector in (A.83) can be decomposed as follows:

$$\dot{\mathbf{a}} = \begin{pmatrix} \dot{\mathbf{x}}^\top \frac{\partial a_1}{\partial \mathbf{x}} \\ \dot{\mathbf{x}}^\top \frac{\partial a_2}{\partial \mathbf{x}} \end{pmatrix} = \begin{pmatrix} \dot{\mathbf{x}}^\top & \mathbf{o}_3^\top \\ \mathbf{o}_3^\top & \dot{\mathbf{x}}^\top \end{pmatrix} \begin{pmatrix} \frac{\partial a_1}{\partial \mathbf{x}} \\ \frac{\partial a_2}{\partial \mathbf{x}} \end{pmatrix} = (\mathbf{I}_2 \otimes \dot{\mathbf{x}}^\top) \left( \mathbf{a} \otimes \frac{\partial}{\partial \mathbf{x}} \right). \quad (\text{A.84})$$

Computing the last Kronecker-product, we would formally get  $a_i \frac{\partial}{\partial \mathbf{x}}$  which, of course, should be understood as  $\frac{\partial a_i}{\partial \mathbf{x}}$ . With the help of the permutation matrix  $\mathbf{U}_{\alpha \times \beta}$  and the exchange rule (A.62)

$$(\mathbf{A} \otimes \mathbf{B}) = \mathbf{U}_{s \times p} (\mathbf{B} \otimes \mathbf{A}) \mathbf{U}_{q \times t} \quad \text{if } \mathbf{A} \in \mathbb{R}^{p \times q} \text{ and } \mathbf{B} \in \mathbb{R}^{s \times t},$$

in the appendix, the last product in (A.84) can be written

$$\dot{\underline{\mathbf{a}}} = \underbrace{\left[ \mathbf{U}_{1 \times 2} (\dot{\mathbf{x}}^\top \otimes \mathbf{I}_2) \mathbf{U}_{2 \times r} \right]}_{\mathbf{I}_{2r}} \underbrace{\left[ \mathbf{U}_{r \times 2} \left( \frac{\partial}{\partial \mathbf{x}} \otimes \mathbf{a} \right) \mathbf{U}_{1 \times 1} \right]}_{\frac{\partial \mathbf{a}}{\partial \mathbf{x}}} = \underline{\underline{(\dot{\mathbf{x}}^\top \otimes \mathbf{I}_2) \frac{\partial \mathbf{a}}{\partial \mathbf{x}}}}. \quad (\text{A.85})$$

For the second form in (A.83), one obtains

$$\dot{\underline{\mathbf{a}}} = \begin{pmatrix} \frac{\partial a_1}{\partial \mathbf{x}^\top} \dot{\mathbf{x}} \\ \frac{\partial a_2}{\partial \mathbf{x}^\top} \dot{\mathbf{x}} \end{pmatrix} = \left( \mathbf{a} \otimes \frac{\partial}{\partial \mathbf{x}} \right) \dot{\mathbf{x}} = \left[ \underbrace{\mathbf{U}_{1 \times 2}}_{\mathbf{I}_2} \left( \frac{\partial}{\partial \mathbf{x}^\top} \otimes \mathbf{a} \right) \underbrace{\mathbf{U}_{1 \times r}}_{\mathbf{I}_r} \right] \dot{\mathbf{x}} = \underline{\underline{\frac{\partial \mathbf{a}}{\partial \mathbf{x}^\top} \dot{\mathbf{x}}}}, \quad (\text{A.86})$$

so that combined, these two possible representations are written as

$$\dot{\mathbf{a}} = (\dot{\mathbf{x}}^\top \otimes \mathbf{I}_2) \frac{\partial \mathbf{a}}{\partial \mathbf{x}} = \frac{\partial \mathbf{a}}{\partial \mathbf{x}^\top} \dot{\mathbf{x}}. \quad (\text{A.87})$$

### A.5.3 Differentiation of a $2 \times 3$ -Matrix with Respect to Time

For the derivative of a  $2 \times 3$ -matrix with respect to time, with the above results, one obtains

$$\begin{aligned}
 \underline{\dot{A}} &= \begin{pmatrix} \dot{a}_{11} & \dot{a}_{12} & \dot{a}_{13} \\ \dot{a}_{21} & \dot{a}_{22} & \dot{a}_{23} \end{pmatrix} = \begin{pmatrix} \dot{\mathbf{x}}^\top \frac{\partial a_{11}}{\partial \mathbf{x}} & \dot{\mathbf{x}}^\top \frac{\partial a_{12}}{\partial \mathbf{x}} & \dot{\mathbf{x}}^\top \frac{\partial a_{13}}{\partial \mathbf{x}} \\ \dot{\mathbf{x}}^\top \frac{\partial a_{21}}{\partial \mathbf{x}} & \dot{\mathbf{x}}^\top \frac{\partial a_{22}}{\partial \mathbf{x}} & \dot{\mathbf{x}}^\top \frac{\partial a_{23}}{\partial \mathbf{x}} \end{pmatrix} \\
 &= \begin{pmatrix} \dot{\mathbf{x}}^\top & \mathbf{o}_3^\top \\ \mathbf{o}_3^\top & \dot{\mathbf{x}}^\top \end{pmatrix} \left( A \otimes \frac{\partial}{\partial \mathbf{x}} \right) = (I_2 \otimes \dot{\mathbf{x}}^\top) \left( A \otimes \frac{\partial}{\partial \mathbf{x}} \right) \\
 &= \underbrace{\left[ U_{1 \times 2} (\dot{\mathbf{x}}^\top \otimes I_2) U_{2 \times r} \right]}_{I_{2r}} \underbrace{\left[ U_{r \times 2} \left( \frac{\partial}{\partial \mathbf{x}} \otimes A \right) U_{3 \times 1} \right]}_{\frac{\partial A}{\partial \mathbf{x}} I_3} = \underline{\underline{(\dot{\mathbf{x}}^\top \otimes I_2) \frac{\partial A}{\partial \mathbf{x}}}}, \quad (\text{A.88})
 \end{aligned}$$

or, with the second representation in (A.81) for the  $\dot{a}_{ij}$ ,

$$\begin{aligned}
 \underline{\dot{A}} &= \begin{pmatrix} \frac{\partial a_{11}}{\partial \mathbf{x}^\top} \dot{\mathbf{x}} & \frac{\partial a_{12}}{\partial \mathbf{x}^\top} \dot{\mathbf{x}} & \frac{\partial a_{13}}{\partial \mathbf{x}^\top} \dot{\mathbf{x}} \\ \frac{\partial a_{21}}{\partial \mathbf{x}^\top} \dot{\mathbf{x}} & \frac{\partial a_{22}}{\partial \mathbf{x}^\top} \dot{\mathbf{x}} & \frac{\partial a_{23}}{\partial \mathbf{x}^\top} \dot{\mathbf{x}} \end{pmatrix} = \left( A \otimes \frac{\partial}{\partial \mathbf{x}^\top} \right) \begin{pmatrix} \dot{\mathbf{x}} & \mathbf{o} & \mathbf{o} \\ \mathbf{o} & \dot{\mathbf{x}} & \mathbf{o} \\ \mathbf{o} & \mathbf{o} & \dot{\mathbf{x}} \end{pmatrix} \\
 &= \left[ U_{1 \times 2} \left( \frac{\partial}{\partial \mathbf{x}^\top} \otimes A \right) U_{3 \times r} \right] \underbrace{\left[ U_{r \times 3} (\dot{\mathbf{x}} \otimes I_3) U_{3 \times 1} \right]}_{I_3} = \underline{\underline{\frac{\partial A}{\partial \mathbf{x}^\top} (\dot{\mathbf{x}} \otimes I_3)}}. \quad (\text{A.89})
 \end{aligned}$$

Here the Kronecker-product is also present.

### A.5.4 Differentiation of an $n \times m$ -Matrix with Respect to Time

In general, one gets for a matrix  $A \in \mathbb{R}^{n \times m}$  and a vector  $\mathbf{x} \in \mathbb{R}^r$

$$\boxed{\dot{A} = (\dot{\mathbf{x}}^\top \otimes I_n) \frac{\partial A}{\partial \mathbf{x}} \text{ and } \frac{\partial A}{\partial \mathbf{x}^\top} (\dot{\mathbf{x}} \otimes I_m) \in \mathbb{R}^{n \times m}.} \quad (\text{A.90})$$

The derivation is given below without any comment.

$$\begin{aligned}
 \dot{A} &= \begin{pmatrix} \dot{\mathbf{x}}^\top & \dots & O \\ \vdots & \ddots & \vdots \\ O & \dots & \dot{\mathbf{x}}^\top \end{pmatrix} \left( A \otimes \frac{\partial}{\partial \mathbf{x}} \right) = (I_n \otimes \dot{\mathbf{x}}^\top) \left( A \otimes \frac{\partial}{\partial \mathbf{x}} \right) \\
 &= \underbrace{\left[ U_{1 \times n} (\dot{\mathbf{x}}^\top \otimes I_n) U_{n \times r} \right]}_{I_{nr}} \underbrace{\left[ U_{r \times n} \left( \frac{\partial}{\partial \mathbf{x}} \otimes A \right) U_{m \times 1} \right]}_{\frac{\partial A}{\partial \mathbf{x}} I_m} = \underline{\underline{(\dot{\mathbf{x}}^\top \otimes I_n) \frac{\partial A}{\partial \mathbf{x}}}}.
 \end{aligned}$$

$$\begin{aligned} \dot{\mathbf{A}} &= \left( \mathbf{A} \otimes \frac{\partial}{\partial \mathbf{x}^\top} \right) \begin{pmatrix} \dot{\mathbf{x}} & \dots & \mathbf{O} \\ \vdots & \ddots & \vdots \\ \mathbf{O} & \dots & \dot{\mathbf{x}} \end{pmatrix} = \left( \mathbf{A} \otimes \frac{\partial}{\partial \mathbf{x}^\top} \right) (\mathbf{I}_n \otimes \dot{\mathbf{x}}) \\ &= \underbrace{\left[ \mathbf{U}_{1 \times n} \left( \frac{\partial}{\partial \mathbf{x}^\top} \otimes \mathbf{A} \right) \mathbf{U}_{m \times r} \right]}_{\frac{\partial \mathbf{A}}{\partial \mathbf{x}^\top}} \underbrace{\left[ \mathbf{U}_{r \times m} (\dot{\mathbf{x}} \otimes \mathbf{I}_m) \mathbf{U}_{m \times 1} \right]}_{\mathbf{I}_{mr}} = \underline{\underline{\frac{\partial \mathbf{A}}{\partial \mathbf{x}^\top} (\dot{\mathbf{x}} \otimes \mathbf{I}_n)}}. \end{aligned}$$

## A.6 Supplements to Differentiation with Respect to a Matrix

For the derivative of a  $4 \times 4$  matrix with respect to itself, one has

$$\frac{\partial \mathbf{M}}{\partial \mathbf{M}} = \bar{\mathbf{U}}_{4 \times 4}, \quad (\text{A.91})$$

where  $\bar{\mathbf{U}}_{4 \times 4}$  is defined by

$$\begin{aligned} \bar{\mathbf{U}}_{4 \times 4} &\stackrel{\text{def}}{=} \sum_i^4 \sum_k^4 \mathbf{E}_{ik} \otimes \mathbf{E}_{ik} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}. \quad (\text{A.92}) \end{aligned}$$

This fact can be easily made clear just by the definition of the differentiation of a matrix with respect to a matrix. The result is more complex when the matrix  $\mathbf{M} = \mathbf{M}^\top$



## Appendix B

### Some Differential Geometry

From a sheet of letter paper, one can form a cylinder or a cone, but it is impossible to obtain a surface element of a sphere without folding, stretching or cutting. The reason lies in the geometry of the spherical surface: No part of such a surface can be isometrically mapped onto the plane.

#### B.1 Curvature of a Curved Line in Three Dimensions

In a plane, the tangent vector remains constant when one moves on it: the plane has no curvature. The same is true for a straight line when the tangent vector coincides with the line. If a line in a neighbourhood of one of its points is not a straight line, it is called a *curved* line. The same is valid for a curved surface. We consider curves in the three-dimensional space with the position vector  $\mathbf{x}(q)$  of points parametrized by the length  $q$ . The direction of a curve  $\mathcal{C}$  at the point  $\mathbf{x}(q)$  is given by the normalized tangent vector

$$\mathbf{t}(q) \stackrel{\text{def}}{=} \frac{\mathbf{x}'(q)}{\|\mathbf{x}'(q)\|},$$

with  $\mathbf{x}'(q) \stackrel{\text{def}}{=} \frac{\partial \mathbf{x}(q)}{\partial q}$ . Passing through a curve from a starting point  $\mathbf{x}(q_0)$  to an end-point  $\mathbf{x}(q)$  does not change the tangent vector  $\mathbf{t}$  in a straight line, the tip of the tangent vector does not move, thus describing a curve of length 0. If the curve is curved, then the tip of the tangent vector describes an arc of length not equal to zero. As *arc length* we call the integral over a curved arc with endpoints  $q_0$  and  $q$  ( $q > q_0$ ):

$$\int_{q_0}^q \sqrt{x_1'^2 + x_2'^2 + x_3'^2} dq = \int_{q_0}^q \sqrt{\mathbf{x}'^T \mathbf{x}'} dq = \int_{q_0}^q \|\mathbf{x}'\| dq. \quad (\text{B.1})$$

This formula arises as follows: Suppose the interval  $[q_0, q]$  is divided by the points  $q_0 < q_1 < \dots < q_n = q$ , then the length  $\sigma_n$  of the inscribed polygon in the curve

arch is

$$\sigma_n = \sum_{k=0}^n \sqrt{[x_1(q_k) - x_1(q_{k+1})]^2 + [x_2(q_k) - x_2(q_{k+1})]^2 + [x_3(q_k) - x_3(q_{k+1})]^2}. \quad (\text{B.2})$$

By the mean value theorem of differential calculus, for every smooth curve between  $q_k$  and  $q_{k+1}$  there exists a point  $q_k^{(i)}$  such that

$$x_i(q_k) - x_i(q_{k+1}) = x_i'(q_k^{(i)})(q_{k+1} - q_k) \quad (\text{B.3})$$

for  $i = 1, 2$  and  $3$ . Inserting (B.3) into (B.2) yields

$$\sigma_n = \sum_{k=0}^n [q_{k+1} - q_k] \sqrt{[x_1'(q_k^{(1)})]^2 + [x_2'(q_k^{(2)})]^2 + [x_3'(q_k^{(3)})]^2}. \quad (\text{B.4})$$

From (B.4) one gets, as  $q_{k+1} - q_k \rightarrow 0$ , (B.1).

A measure of the curvature of a curve is the rate of change of the direction. The curvature is larger when the change of direction of the tangent vector  $\mathbf{t}$  is greater. Generally, we therefore define as a *curvature of a curve C* at a point  $\mathbf{x}(q_0)$

$$\kappa(q_0) \stackrel{\text{def}}{=} \lim_{q \rightarrow q_0} \frac{\text{length of } \mathbf{t}}{\text{length of } \mathbf{x}} = \frac{\|\mathbf{t}'(q_0)\|}{\|\mathbf{x}'(q_0)\|}, \quad (\text{B.5})$$

where

$$\text{length of } \mathbf{t} \stackrel{\text{def}}{=} \int_{q_0}^q \|\mathbf{t}'\| \, dq$$

and

$$\text{length of } \mathbf{x} \stackrel{\text{def}}{=} \int_{q_0}^q \|\mathbf{x}'\| \, dq.$$

A straight line has zero curvature. In the case of a circle, the curvature is constant; the curvature is greater, the smaller the radius. The reciprocal value  $1/\kappa$  of the curvature is called the *curvature radius*.

## B.2 Curvature of a Surface in Three Dimensions

### B.2.1 Vectors in the Tangent Plane

Already in the nineteenth century, Gauss investigated how from the measurements on a surface one can make conclusions about its spatial form. He then came to his main result, the *Theorema Egregium*, which states that the Gaussian curvature of a



surface depends only on the internal variables  $g_{ij}$  and their derivatives. This result is deduced in the following.

Suppose an area is defined as a function  $\mathbf{x}(q_1, q_2) \in \mathbb{R}^3$  of the two coordinates  $q_1$  and  $q_2$ . At a point  $P$  of the surface, the tangent plane is, for example, spanned by the two tangent vectors  $\mathbf{x}_1 \stackrel{\text{def}}{=} \frac{\partial \mathbf{x}}{\partial q_1}$  and  $\mathbf{x}_2 \stackrel{\text{def}}{=} \frac{\partial \mathbf{x}}{\partial q_2}$ . If the two tangent vectors  $\mathbf{x}_1$  and  $\mathbf{x}_2$  are linearly independent, any vector in the tangent plane can be decomposed in a linear combination of the two vectors, e.g. as

$$v^1 \mathbf{x}_1 + v^2 \mathbf{x}_2.$$

The scalar product of two vectors from the tangent plane is then defined as

$$\begin{aligned} \underline{\underline{(\mathbf{v} \cdot \mathbf{w})}} &\stackrel{\text{def}}{=} (v^1 \mathbf{x}_1^\top + v^2 \mathbf{x}_2^\top)(w^1 \mathbf{x}_1 + w^2 \mathbf{x}_2) \\ &= \mathbf{v}^\top \begin{pmatrix} \mathbf{x}_1^\top \\ \mathbf{x}_2^\top \end{pmatrix} [\mathbf{x}_1, \mathbf{x}_2] \mathbf{w} = \mathbf{v}^\top \begin{pmatrix} \mathbf{x}_1^\top \mathbf{x}_1 & \mathbf{x}_1^\top \mathbf{x}_2 \\ \mathbf{x}_2^\top \mathbf{x}_1 & \mathbf{x}_2^\top \mathbf{x}_2 \end{pmatrix} \mathbf{w} = \underline{\underline{\mathbf{v}^\top \mathbf{G} \mathbf{w}}}. \end{aligned}$$

As  $\mathbf{x}_1^\top \mathbf{x}_2 = \mathbf{x}_2^\top \mathbf{x}_1$ , the matrix  $\mathbf{G}^\top = \mathbf{G}$  is symmetric. In addition, one has

$$\|\mathbf{v}\| = \sqrt{(\mathbf{v} \cdot \mathbf{v})} = \sqrt{\mathbf{v}^\top \mathbf{G} \mathbf{v}}.$$

Now we want to define the curvature by oriented parallelograms. Let  $\mathbf{v} \wedge \mathbf{w}$  be the oriented parallelogram defined by the vectors  $\mathbf{v}$  and  $\mathbf{w}$  in *that order*.  $\mathbf{w} \wedge \mathbf{v} = -\mathbf{v} \wedge \mathbf{w}$  is then the parallelogram with the opposite orientation,  $\text{area}(\mathbf{w} \wedge \mathbf{v}) = -\text{area}(\mathbf{v} \wedge \mathbf{w})$ . The determinant of the matrix  $\mathbf{G}$  is then obtained as

$$\begin{aligned} \underline{\underline{g}} &\stackrel{\text{def}}{=} \det \mathbf{G} = \|\mathbf{x}_1\|^2 \cdot \|\mathbf{x}_2\|^2 - (\mathbf{x}_1 \cdot \mathbf{x}_2)^2 \\ &= \|\mathbf{x}_1\|^2 \cdot \|\mathbf{x}_2\|^2 - \|\mathbf{x}_1\|^2 \cdot \|\mathbf{x}_2\|^2 \cos^2 \Theta \\ &= \|\mathbf{x}_1\|^2 \cdot \|\mathbf{x}_2\|^2 (1 - \cos^2 \Theta) = \underline{\underline{\|\mathbf{x}_1\|^2 \cdot \|\mathbf{x}_2\|^2 \sin^2 \Theta}}. \end{aligned}$$

On the other hand, one has  $\|\mathbf{x}_1 \times \mathbf{x}_2\| = \|\mathbf{x}_1\| \cdot \|\mathbf{x}_2\| \cdot \sin \Theta$ , and so

$$\underline{\underline{\|\mathbf{x}_1 \times \mathbf{x}_2\|}} = \sqrt{g} = \text{area}(\mathbf{x}_1 \wedge \mathbf{x}_2). \quad (\text{B.6})$$

For the vector product of two vectors  $\mathbf{v}$  and  $\mathbf{w}$  from the tangent plane, one gets, on the other hand,

$$\begin{aligned} &(v^1 \mathbf{x}_1 + v^2 \mathbf{x}_2) \times (w^1 \mathbf{x}_1 + w^2 \mathbf{x}_2) \\ &= v^1 w^1 \underbrace{(\mathbf{x}_1 \times \mathbf{x}_1)}_0 + v^1 w^2 (\mathbf{x}_1 \times \mathbf{x}_2) + v^2 w^1 \underbrace{(\mathbf{x}_2 \times \mathbf{x}_1)}_{-(\mathbf{x}_1 \times \mathbf{x}_2)} + v^2 w^2 \underbrace{(\mathbf{x}_2 \times \mathbf{x}_2)}_0 \\ &= \underbrace{(v^1 w^2 - v^2 w^1)}_{\stackrel{\text{def}}{=} \det \mathbf{R}} (\mathbf{x}_1 \times \mathbf{x}_2), \end{aligned}$$

or

$$\text{area}(\mathbf{v} \wedge \mathbf{w}) = \det \mathbf{R} \cdot \text{area}(\mathbf{x}_1 \wedge \mathbf{x}_2),$$

so

$$\underline{\underline{\text{area}(\mathbf{v} \wedge \mathbf{w}) = \det \mathbf{R} \sqrt{g}}}. \quad (\text{B.7})$$

The area of the parallelogram spanned by two vectors in the tangential surface is thus determined by the vector components and the surface defining the matrix  $\mathbf{G}$ .

## B.2.2 Curvature and Normal Vectors

For two-dimensional surfaces, one should define the curvature at a point without using the tangent vectors directly because there are infinitely many of them in the tangent plane. However, any smooth surface in  $\mathbb{R}^3$  has at each point a unique normal direction, which is one-dimensional, so it can be described by the unit normal vector. A *normal vector* at  $\mathbf{x}(q)$  is defined as the normalized vector perpendicular to the tangent plane:

$$\mathbf{n}(q) \stackrel{\text{def}}{=} \frac{\mathbf{x}_1 \times \mathbf{x}_2}{\|\mathbf{x}_1 \times \mathbf{x}_2\|}.$$

If the surface is curved, the normal vector changes with displacement according to

$$\mathbf{n}_i(q) \stackrel{\text{def}}{=} \frac{\partial \mathbf{n}(q)}{\partial q^i}.$$

These change vectors  $\mathbf{n}_i$  lie in the tangent plane because

$$\frac{\partial(\mathbf{n} \cdot \mathbf{n})}{\partial q^i} = 0 = \left( \frac{\partial \mathbf{n}}{\partial q^i} \cdot \mathbf{n} \right) + \left( \mathbf{n} \cdot \frac{\partial \mathbf{n}}{\partial q^i} \right) = 2(\mathbf{n}_i \cdot \mathbf{n}).$$

The bigger the area spanned by the two change vectors  $\mathbf{n}_1$  and  $\mathbf{n}_2$  lying in the tangent plane, the bigger the curvature at the considered point. If  $\Omega$  is an area of the tangent surface which contains the point under consideration, then the following curvature definition is obvious:

$$\begin{aligned} \kappa(q) &\stackrel{\text{def}}{=} \lim_{\Omega \rightarrow q} \frac{\text{area of } \mathbf{n}(\Omega)}{\text{area of } \Omega} = \lim_{\Omega \rightarrow q} \frac{\iint_{\Omega} \|\mathbf{n}_1(\tilde{q}) \times \mathbf{n}_2(\tilde{q})\| d\tilde{q}^1 d\tilde{q}^2}{\iint_{\Omega} \|\mathbf{x}_1(\tilde{q}) \times \mathbf{x}_2(\tilde{q})\| d\tilde{q}^1 d\tilde{q}^2} \\ &= \frac{\|\mathbf{n}_1(q) \times \mathbf{n}_2(q)\|}{\|\mathbf{x}_1(q) \times \mathbf{x}_2(q)\|} = \frac{\text{area of } \mathbf{n}_1(q) \wedge \mathbf{n}_2(q)}{\text{area of } \mathbf{x}_1(q) \wedge \mathbf{x}_2(q)}. \end{aligned} \quad (\text{B.8})$$

Since both  $\mathbf{n}_1$  and  $\mathbf{n}_2$  lie in the tangent plane, they can be displayed as linear combinations of the vectors  $\mathbf{x}_1$  and  $\mathbf{x}_2$ :

$$\mathbf{n}_1 = -b_1^1 \mathbf{x}_1 - b_1^2 \mathbf{x}_2 \quad \text{and} \quad \mathbf{n}_2 = -b_2^1 \mathbf{x}_1 - b_2^2 \mathbf{x}_2. \quad (\text{B.9})$$

Combining together the coefficients  $-b_i^j$  in the matrix  $\overline{\mathbf{B}}$ , this corresponds to the matrix  $\mathbf{R}$  in (B.7), and (B.6) then yields the Gauss-curvature

$$\underline{\underline{\kappa(\mathbf{q}) = \det \overline{\mathbf{B}}.}} \quad (\text{B.10})$$

To confirm the Theorema Egregium of Gauss, one must now shown that  $\overline{\mathbf{B}}$  depends only on the inner values  $g_{ij}$  and their derivatives!

### B.2.3 Theorema Egregium and the Inner Values $g_{ij}$

First, we examine the changes of the tangent vectors  $\mathbf{x}_k$  by looking at their derivatives

$$\mathbf{x}_{jk} \stackrel{\text{def}}{=} \frac{\partial \mathbf{x}_k}{\partial q^j} = \frac{\partial^2 \mathbf{x}}{\partial q^j \partial q^k}, \quad (\text{B.11})$$

which implies that

$$\mathbf{x}_{jk} = \mathbf{x}_{kj}. \quad (\text{B.12})$$

Since the two vectors  $\mathbf{x}_1$  and  $\mathbf{x}_2$  are a basis for the tangent plane and the vector  $\mathbf{n}$  is orthonormal to this plane, any vector in  $\mathbb{R}^3$ , including the vector  $\mathbf{x}_{jk}$ , can assembled as a linear combination of these three vectors:

$$\underline{\underline{\mathbf{x}_{jk} = \Gamma_{jk}^1 \mathbf{x}_1 + \Gamma_{jk}^2 \mathbf{x}_2 + b_{jk} \mathbf{n}.}} \quad (\text{B.13})$$

The vectors  $\mathbf{x}_{jk}$  can be summarized in a  $4 \times 2$ -matrix as follows:

$$\frac{\partial^2 \mathbf{x}}{\partial \mathbf{q} \partial \mathbf{q}^\top} = \begin{pmatrix} \mathbf{x}_{11} & \mathbf{x}_{12} \\ \mathbf{x}_{21} & \mathbf{x}_{22} \end{pmatrix} = \mathbf{\Gamma}_1 \otimes \mathbf{x}_1 + \mathbf{\Gamma}_2 \otimes \mathbf{x}_2 + \mathbf{B} \otimes \mathbf{n}, \quad (\text{B.14})$$

where

$$\mathbf{\Gamma}_i \stackrel{\text{def}}{=} \begin{pmatrix} \Gamma_{11}^i & \Gamma_{12}^i \\ \Gamma_{21}^i & \Gamma_{22}^i \end{pmatrix}$$

and

$$\mathbf{B} \stackrel{\text{def}}{=} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

Multiplying (B.14) from the left by the matrix  $(\mathbf{I}_2 \otimes \mathbf{n}^\top)$ , due to  $\mathbf{n}^\top \mathbf{x}_i = 0$  and  $(\mathbf{I}_2 \otimes \mathbf{n}^\top)(\mathbf{B} \otimes \mathbf{n}) = \mathbf{B} \otimes (\mathbf{n}^\top \mathbf{n}) = \mathbf{B} \otimes \mathbf{1} = \mathbf{B}$ , one finally obtains

$$\underline{\underline{\mathbf{B} = (\mathbf{I}_2 \otimes \mathbf{n}^\top) \frac{\partial^2 \mathbf{x}}{\partial \mathbf{q} \partial \mathbf{q}^\top} = \begin{pmatrix} \mathbf{n}^\top \mathbf{x}_{11} & \mathbf{n}^\top \mathbf{x}_{12} \\ \mathbf{n}^\top \mathbf{x}_{21} & \mathbf{n}^\top \mathbf{x}_{22} \end{pmatrix}.}} \quad (\text{B.15})$$

i.e. the  $2 \times 2$ -matrix  $\overline{\mathbf{B}}$  is symmetric. This matrix is called the *Second Fundamental Form* of the surface. The *First Quadratic Fundamental Form* is given by the connection

$$\underline{\underline{\mathbf{G}}} = \underline{\underline{\frac{\partial \mathbf{x}^\top}{\partial \mathbf{q}} \cdot \frac{\partial \mathbf{x}}{\partial \mathbf{q}^\top}}}, \quad (\text{B.16})$$

or more precisely, by the right-hand side of the equation for the squared line element  $ds$  of the surface:

$$ds^2 = d\mathbf{q}^\top \mathbf{G} d\mathbf{q}.$$

Gauss designates, as is still in elementary geometry today, the elements of the matrix  $\mathbf{G}$  in this way:

$$\mathbf{G} = \begin{pmatrix} E & F \\ F & G \end{pmatrix}.$$

While  $\mathbf{G}$  therefore plays a critical role in determining the length of a curve in an area,  $\mathbf{B}$  is, as we will see later, decisively involved in the determination of the curvature of a surface. A further representation of the matrix  $\mathbf{B}$  is obtained from the derivative of the scalar product of the mutually orthogonal vectors  $\mathbf{n}$  and  $\mathbf{x}_j$ :

$$\mathbf{n}^\top \frac{\partial \mathbf{x}}{\partial \mathbf{q}^\top} = \mathbf{0}^\top;$$

because this derivative is, according to (A.73),

$$\frac{\partial}{\partial \mathbf{q}} \left( \mathbf{n}^\top \frac{\partial \mathbf{x}}{\partial \mathbf{q}^\top} \right) = \frac{\partial \mathbf{n}^\top}{\partial \mathbf{q}} \cdot \frac{\partial \mathbf{x}}{\partial \mathbf{q}^\top} + (\mathbf{I}_2 \otimes \mathbf{n}^\top) \frac{\partial^2 \mathbf{x}}{\partial \mathbf{q} \partial \mathbf{q}^\top} = \mathbf{0}_{2 \times 2}.$$

With (B.15) we obtain

$$\underline{\underline{\mathbf{B}}} = - \underline{\underline{\frac{\partial \mathbf{n}^\top}{\partial \mathbf{q}} \cdot \frac{\partial \mathbf{x}}{\partial \mathbf{q}^\top}}}, \quad (\text{B.17})$$

i.e. a further interesting form for the matrix  $\mathbf{B}$

$$\underline{\underline{\mathbf{B}}} = - \underline{\underline{\begin{pmatrix} \mathbf{n}_1^\top \mathbf{x}_1 & \mathbf{n}_1^\top \mathbf{x}_2 \\ \mathbf{n}_2^\top \mathbf{x}_1 & \mathbf{n}_2^\top \mathbf{x}_2 \end{pmatrix}}}. \quad (\text{B.18})$$

The two equations (B.9) can be summarized to a matrix equation as follows:

$$[\mathbf{n}_1 | \mathbf{n}_2] = \underline{\underline{\frac{\partial \mathbf{n}}{\partial \mathbf{q}^\top}}} = \frac{\partial \mathbf{x}}{\partial \mathbf{q}^\top} \begin{pmatrix} -b_1^1 & -b_2^1 \\ -b_1^2 & -b_2^2 \end{pmatrix} = \underline{\underline{\frac{\partial \mathbf{x}}{\partial \mathbf{q}^\top}}} \cdot \underline{\underline{\mathbf{B}}}. \quad (\text{B.19})$$

This used in (B.18) together with (B.16) yields

$$\mathbf{B} = -\overline{\mathbf{B}}^\top \cdot \frac{\partial \mathbf{x}^\top}{\partial \mathbf{q}} \cdot \frac{\partial \mathbf{x}}{\partial \mathbf{q}^\top} = -\overline{\mathbf{B}}^\top \mathbf{G},$$

or transposed

$$\underline{\underline{\mathbf{B}}} = -\underline{\underline{\mathbf{G}\overline{\mathbf{B}}}}, \quad (\text{B.20})$$

since  $\mathbf{B}$  and  $\mathbf{G}$  are both symmetric matrices. Our goal remains to show that the Gaussian curvature depends only on the  $g_{ij}$ 's and their derivatives with respect to  $q_k$ , i.e., according to (B.10), one must show that for the matrix  $\overline{\mathbf{B}}$

$$\kappa(\mathbf{q}) = \det \overline{\mathbf{B}}$$

is valid. We first examine the matrix  $\overline{\mathbf{G}}$ . For this purpose, its elements are differentiated:

$$\frac{\partial g_{ij}}{\partial q_k} = \mathbf{x}_{ik}^\top \mathbf{x}_j + \mathbf{x}_{jk}^\top \mathbf{x}_i.$$

With (B.13) we obtain

$$\begin{aligned} \mathbf{x}_j^\top \mathbf{x}_{ik} &= \Gamma_{ik}^1 \mathbf{x}_j^\top \mathbf{x}_1 + \Gamma_{ik}^2 \mathbf{x}_j^\top \mathbf{x}_2 + b_{ik} \mathbf{x}_j^\top \mathbf{n} \\ &= \Gamma_{ik}^1 g_{j1} + \Gamma_{ik}^2 g_{j2}. \end{aligned}$$

If we define

$$\check{\Gamma}_{ik}^j \stackrel{\text{def}}{=} \Gamma_{ik}^1 g_{j1} + \Gamma_{ik}^2 g_{j2}$$

and assemble all four components into a matrix  $\check{\Gamma}_j$ , we obtain

$$\begin{pmatrix} \check{\Gamma}_{j1}^1 & \check{\Gamma}_{j1}^2 \\ \check{\Gamma}_{j2}^1 & \check{\Gamma}_{j2}^2 \end{pmatrix} = \begin{pmatrix} \Gamma_{j1}^1 & \Gamma_{j1}^2 \\ \Gamma_{j2}^1 & \Gamma_{j2}^2 \end{pmatrix} \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} = \mathbf{\Gamma}_j \mathbf{G}^\top \quad (\text{B.21})$$

or, because of  $\mathbf{G}^\top = \mathbf{G}$ ,

$$\check{\Gamma}_j = \mathbf{\Gamma}_j \mathbf{G}. \quad (\text{B.22})$$

It is therefore true that

$$\frac{\partial g_{ij}}{\partial q_k} = \check{\Gamma}_{ik}^j + \check{\Gamma}_{jk}^i. \quad (\text{B.23})$$

With the following expression of three different derivatives, one obtains

$$\frac{\partial g_{ij}}{\partial q_k} + \frac{\partial g_{ik}}{\partial q_j} - \frac{\partial g_{jk}}{\partial q_i} = \check{\Gamma}_{ik}^j + \check{\Gamma}_{jk}^i + \check{\Gamma}_{ij}^k + \check{\Gamma}_{jk}^i - \check{\Gamma}_{ik}^j - \check{\Gamma}_{ji}^k, \quad (\text{B.24})$$

so

$$\underline{\underline{\check{\Gamma}_{jk}^i}} = \frac{1}{2} \left( \frac{\partial g_{ij}}{\partial q_k} + \frac{\partial g_{ik}}{\partial q_j} - \frac{\partial g_{jk}}{\partial q_i} \right). \quad (\text{B.25})$$

Multiplying (B.22) from the right with

$$\mathbf{G}^{-1} \stackrel{\text{def}}{=} \begin{pmatrix} g_{11}^{[-1]} & g_{12}^{[-1]} \\ g_{21}^{[-1]} & g_{22}^{[-1]} \end{pmatrix},$$

one gets the relation

$$\Gamma_j = \check{\Gamma}_j \mathbf{G}^{-1}, \quad (\text{B.26})$$

i.e. element by element

$$\Gamma_{jk}^\ell = \frac{1}{2} \sum_i g_{i\ell}^{[-1]} \check{\Gamma}_{jk}^i, \quad (\text{B.27})$$

so with (B.25)

$$\Gamma_{jk}^\ell = \frac{1}{2} \sum_i g_{i\ell}^{[-1]} \left( \frac{\partial g_{ij}}{\partial q_k} + \frac{\partial g_{ik}}{\partial q_j} - \frac{\partial g_{jk}}{\partial q_i} \right). \quad (\text{B.28})$$

This clarifies the relationship of the Christoffel-symbols  $\Gamma_{jk}^\ell$  with the  $g_{ij}$  and their derivatives. Now the direct relationship of these variables with the Gaussian curvature  $\kappa$  has to be made. One gets this finally by repeated differentiation of  $\mathbf{x}_{jk}$  with respect to  $q^\ell$ :

$$\begin{aligned} \mathbf{x}_{jkl} &\stackrel{\text{def}}{=} \frac{\partial \mathbf{x}_{jk}}{\partial q^\ell} = \sum_i \frac{\partial \Gamma_{jk}^i}{\partial q^\ell} \mathbf{x}_i + \sum_i \Gamma_{jk}^i \mathbf{x}_{i\ell} + \frac{\partial b_{jk}}{\partial q^\ell} \mathbf{n} + b_{jk} \mathbf{n}_\ell \\ &= \sum_i \left( \frac{\partial \Gamma_{jk}^i}{\partial q^\ell} + \Gamma_{j\ell}^i \Gamma_{pk}^i - b_{jk} b_\ell^i \right) \mathbf{x}_i + \left( \frac{\partial b_{jk}}{\partial q^\ell} + \sum_p \Gamma_{jk}^p b_{p\ell} \right) \mathbf{n}. \end{aligned} \quad (\text{B.29})$$

Interchanging in (B.29)  $k$  and  $\ell$ , we obtain

$$\mathbf{x}_{j\ell k} = \sum_i \left( \frac{\partial \Gamma_{j\ell}^i}{\partial q^k} + \Gamma_{jk}^i \Gamma_{p\ell}^i - b_{j\ell} b_k^i \right) \mathbf{x}_i + \left( \frac{\partial b_{j\ell}}{\partial q^k} + \sum_p \Gamma_{j\ell}^p b_{pk} \right) \mathbf{n}. \quad (\text{B.30})$$

Subtracting the two third-order derivatives, we obtain

$$\mathbf{0} = \mathbf{x}_{j\ell k} - \mathbf{x}_{jkl} = \sum_i [R_{jk}^{i\ell} - (b_{j\ell} b_k^i - b_{jk} b_\ell^i)] \mathbf{x}_i + (\dots) \mathbf{n}, \quad (\text{B.31})$$

with

$$R_{jk}^{i\ell} \stackrel{\text{def}}{=} \frac{\partial \Gamma_{j\ell}^i}{\partial q^k} - \frac{\partial \Gamma_{jk}^i}{\partial q^\ell} + \sum_p \Gamma_{j\ell}^p \Gamma_{pk}^i - \sum_p \Gamma_{jk}^p \Gamma_{p\ell}^i. \quad (\text{B.32})$$

Since the vectors  $\mathbf{x}_1, \mathbf{x}_2$  and  $\mathbf{n}$  are linearly independent, the square bracket in (B.31) must be zero, which implies

$$R_{jk}^{i\ell} = b_{j\ell}b_k^i - b_{jk}b_\ell^i. \tag{B.33}$$

Defining

$$\check{R}_{jk}^{i\ell} = \sum_i g_{ih}R_{jk}^{i\ell}, \tag{B.34}$$

one gets

$$\check{R}_{jk}^{i\ell} = g_{1h}b_{j\ell}b_k^1 - g_{1h}b_{jk}b_\ell^1 + g_{2h}b_{j\ell}b_k^2 - g_{2h}b_{jk}b_\ell^2 = b_{j\ell}b_{kh} - b_{jk}b_{\ell h}. \tag{B.35}$$

In particular,

$$\underline{\underline{\check{R}_{12}^{12} = b_{22}b_{11} - b_{21}b_{21} = \det \mathbf{B}}}. \tag{B.36}$$

It is therefore true that

$$\kappa(\mathbf{q}) = \det \bar{\mathbf{B}} = \frac{\det \mathbf{B}}{\det \mathbf{G}} = \frac{\check{R}_{12}^{12}}{g},$$

$\kappa(\mathbf{q}) = \frac{\check{R}_{12}^{12}}{g},$	$\tag{B.37}$
-------------------------------------------------------	--------------

which has finally proved the Theorema Egregium because, due to (B.34),  $\check{R}_{jk}^{i\ell}$  depends on  $R_{jk}^{i\ell}$ ; according to (B.32),  $R_{jk}^{i\ell}$  only depends on  $\Gamma_{ij}^k$  and their derivatives and, in accordance with (B.25), the  $\Gamma_{ij}^k$ 's depend only on the  $g_{ik}$ 's and their derivatives. In the form

$\kappa(\mathbf{q}) = \frac{\det \mathbf{B}}{\det \mathbf{G}},$	$\tag{B.38}$
-----------------------------------------------------------------	--------------

the paramount importance of the two fundamental forms is expressed.

**Remarks**

1. Euclidean geometry of is based on a number of Axioms that require no proofs. One is the parallel postulate stating that to every line one can draw through a point not belonging to it one and only one other line which lies in the same plane and does not intersect the former line. This axiom is replaced in the *hyperbolic* geometry in that it admits infinitely many parallels. An example is the surface of a hyperboloid. In the *elliptical* geometry, for example, on the surface of an ellipsoid and, as a special case, on a spherical surface, there are absolutely no parallels because all great circles, which are here the “straight lines”, meet in

two points. In Euclidean geometry, the distance between two points with the Cartesian coordinates  $x_1, x_2, x_3$  and  $x_1 + dx_1, x_2 + dx_2, x_3 + dx_3$  is simply

$$ds = \sqrt{dx_1^2 + dx_2^2 + dx_3^2},$$

and in the other two geometries this formula is replaced by

$$ds^2 = a_1 dx_1^2 + a_2 dx_2^2 + a_3 dx_3^2,$$

where the coefficients  $a_i$  are certain simple functions of  $x_i$ , in the hyperbolic case, of course, different than in the elliptic case. A convenient analytical representation of curved surfaces is the above used Gaussian parameter representation  $\mathbf{x} = \mathbf{x}(q_1, q_2)$ , where Gauss attaches as curved element:

$$ds^2 = E dq_1^2 + 2F dq_1 dq_2 + G dq_2^2.$$

As an example, we introduce the Gauss-specific parameter representation of the unit sphere, with  $\theta = q_1$  and  $\varphi = q_2$ :

$$x_1 = \sin \theta \cos \varphi, \quad x_2 = \sin \theta \sin \varphi, \quad x_3 = \cos \theta.$$

For the arc element of the unit sphere, we obtain

$$ds^2 = d\theta^2 + \sin^2 \theta (d\varphi)^2.$$

Riemann generalized the Gaussian theory of surfaces, which is valid for two-dimensional surfaces in three-dimensional spaces, to  $p$ -dimensional hypersurfaces in  $n$ -dimensional spaces, i.e. where

$$\mathbf{x} = \mathbf{x}(q_1, \dots, q_p) \in \mathbb{R}^n$$

is a point on the hypersurface. He made in addition the fundamentally important step, to set up a homogeneous quadratic function of  $dq_i$  with arbitrary functions of the  $q_i$  as coefficients, as the square of the line elements (quadratic form)

$$ds^2 = \sum_{ik} g_{ik} dq_i dq_k = d\mathbf{q}^T \mathbf{G} d\mathbf{q}.$$

2. The above-occurring  $R_{jk}^{i\ell}$  can be used as matrix elements of the  $4 \times 4$ -matrix  $\mathbf{R}$ , the Riemannian *Curvature Matrix*, to be constructed as a block matrix as follows:

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}^{11} & \mathbf{R}^{12} \\ \mathbf{R}^{21} & \mathbf{R}^{22} \end{pmatrix},$$

where the  $2 \times 2$  sub-matrices have the form:

$$\mathbf{R}^{i\ell} = \begin{pmatrix} R_{11}^{i\ell} & R_{12}^{i\ell} \\ R_{21}^{i\ell} & R_{22}^{i\ell} \end{pmatrix}.$$



In particular,  $\check{R}_{12}^{12}$  is the element in the top right corner of the matrix  $\check{\mathbf{R}} = \mathbf{G}\mathbf{R}$ .

3. Expanding the representation of  $\mathbf{x}(\mathbf{q} + \Delta\mathbf{q})$  in a Taylor series, one obtains

$$\mathbf{x}(\mathbf{q} + \Delta\mathbf{q}) = \mathbf{x}(\mathbf{q}) + \sum_i \mathbf{x}_i \Delta q_i + \frac{1}{2} \sum_{i,k} \mathbf{x}_{ik} \Delta q_i \Delta q_k + \sigma(3).$$

Subtracting  $\mathbf{x}(\mathbf{q})$  on both sides of this equation and multiplying the result from the left with the transposed normal vector  $\mathbf{n}^\top$ , we obtain

$$\begin{aligned} \mathbf{n}^\top [\mathbf{x}(\mathbf{q} + \Delta\mathbf{q}) - \mathbf{x}(\mathbf{q})] &= \sum_i \underbrace{\mathbf{n}^\top \mathbf{x}_i}_0 \Delta q_i + \frac{1}{2} \sum_{i,k} \underbrace{\mathbf{n}^\top \mathbf{x}_{ik}}_{b_{ik}} \Delta q_i \Delta q_k + \sigma(3) \\ &= \mathbf{n}^\top \Delta \mathbf{x}(\mathbf{q}) \stackrel{\text{def}}{=} \Delta \ell. \end{aligned}$$

Thus

$$d\ell \approx \frac{1}{2} \sum_{i,k} b_{ik} dq_i dq_k.$$

The coefficients of the second fundamental form, i.e. the elements of the matrix  $\mathbf{B}$ , Gauss denotes by  $L, M$  and  $N$ . Then the distance  $d\ell$  of the point  $\mathbf{x}(q_1 + dq_1, q_2 + dq_2)$  to the tangent surface at the point  $\mathbf{x}(q_1, q_2)$  is

$$d\ell \approx \frac{1}{2} (L dq_1^2 + 2M dq_1 dq_2 + N dq_2^2).$$

The *normal curvature*  $\kappa$  of a surface at a given point  $P$  and in a given direction  $\mathbf{q}$  is defined as

$$\kappa \stackrel{\text{def}}{=} \frac{L dq_1^2 + 2M dq_1 dq_2 + N dq_2^2}{E dq_1^2 + 2F dq_1 dq_2 + G dq_2^2}. \tag{B.39}$$

The so-defined normal curvature depends, in general, on the chosen direction  $d\mathbf{q}$ . Those directions, in which the normal curvatures at a given point assume an extreme value, are named the *main* directions of the surface at this point. As long as we examine real surfaces, the quadratic differential form  $E dq_1^2 + 2F dq_1 dq_2 + G dq_2^2$  is positive definite, i.e. it is always positive for  $d\mathbf{q} \neq \mathbf{0}$ . Thus the sign of the curvature depends only on the quadratic differential form  $L dq_1^2 + 2M dq_1 dq_2 + N dq_2^2$  in the numerator of (B.39). There are three cases:

- (a)  $LN - M^2 > 0$ , i.e.  $\mathbf{B}$  is positive definite, and the numerator retains the same sign, in each direction one is looking. Such a point is called an *elliptical* point. An example is any point on an ellipsoid, in particular, of course, on a sphere.
- (b)  $LN - M^2 = 0$ , i.e.  $\mathbf{B}$  is semi-definite. The surface behaves at this point as at an elliptical point except in one direction where is  $\kappa = 0$ . This point is called *parabolic*. An example is any point on a cylinder.

- (c)  $LN - M^2 < 0$ , i.e.  $\mathbf{B}$  is indefinite. The numerator does not keep the same sign for all directions. Such a point is called *hyperbolic*, or a *saddle point*. An example is a point on a hyperbolic paraboloid.

Dividing the numerator and the denominator in (B.39) by  $dq_2$  and introducing  $dq_1/dq_2 \stackrel{\text{def}}{=} \lambda$ , we obtain

$$\kappa(\lambda) = \frac{L + 2M\lambda + N\lambda^2}{E + 2F\lambda + G\lambda^2} \quad (\text{B.40})$$

and from this the extreme values from

$$\frac{d\kappa}{d\lambda} = 0$$

as those satisfying

$$(E + 2F\lambda + G\lambda^2)(M + N\lambda) - (L + 2M\lambda + N\lambda^2)(F + G\lambda) = 0. \quad (\text{B.41})$$

In this case, the resulting expression for  $\kappa$  is

$$\kappa = \frac{L + 2M\lambda + N\lambda^2}{E + 2F\lambda + G\lambda^2} = \frac{M + N\lambda}{F + G\lambda}. \quad (\text{B.42})$$

Since furthermore

$$E + 2F\lambda + G\lambda^2 = (E + F\lambda) + \lambda(F + G\lambda)$$

and

$$L + 2M\lambda + N\lambda^2 = (L + M\lambda) + \lambda(M + N\lambda),$$

(B.40) can be transformed into the simpler form

$$\kappa = \frac{L + M\lambda}{E + F\lambda}. \quad (\text{B.43})$$

From this the two equations for  $\kappa$  follow:

$$(\kappa E - L) + (\kappa F - M)\lambda = 0,$$

$$(\kappa F - M) + (\kappa G - N)\lambda = 0.$$

These equations are simultaneously satisfied if and only if

$$\det \begin{pmatrix} \kappa E - L & \kappa F - M \\ \kappa F - M & \kappa G - N \end{pmatrix} = 0. \quad (\text{B.44})$$

This can also be written as

$$\det(\kappa \mathbf{G} - \mathbf{B}) = 0. \quad (\text{B.45})$$

This is the solvability condition for the eigenvalue equation

$$\kappa \mathbf{G} - \mathbf{B} = \mathbf{0},$$

which can be transformed into

$$\kappa \mathbf{I} - \mathbf{G}^{-1} \mathbf{B} = \mathbf{0}. \tag{B.46}$$

This results in a quadratic equation for  $\kappa$ . The two solutions are called the *principal curvatures* and are denoted as  $\kappa_1$  and  $\kappa_2$ . The Gaussian *curvature*  $\kappa$  of a surface at a given point is the product of the principal curvatures  $\kappa_1$  and  $\kappa_2$  of the surface in this point. According to Vieta's root theorem, the product of the solutions is equal to the determinant of the matrix  $\mathbf{G}^{-1} \mathbf{B}$ , so finally,

$$\kappa = \kappa_1 \kappa_2 = \det(\mathbf{G}^{-1} \mathbf{B}) = \frac{\det \mathbf{B}}{\det \mathbf{G}} = \frac{LN - M^2}{EG - F^2}.$$

## Appendix C

# Geodesic Deviation

Geodesics are the lines of general manifolds along which, for example, free particles move. In a flat space the relative velocity of each pair of particles is constant, so that their relative acceleration is always equal to zero. Generally, due to the curvature of space, the relative acceleration is not equal to zero.

The curvature of a surface can be illustrated as follows [21]. Suppose there are two ants on an apple which leave a starting line at the same time and follow with the same speed geodesics which are initially perpendicular to the start line. Initially, their paths are parallel, but, due to the curvature of the apple, they are approaching each other from the beginning. Their distance  $\xi$  from one another is not constant, i.e., in general, the relative acceleration of the ants moving on geodesics with constant velocity is not equal zero if the area over which they move is curved. So the curvature can be indirectly perceived through the so-called *geodesic deviation*  $\xi$ .

The two neighboring geodesics  $\mathbf{x}(u)$  and  $\check{\mathbf{x}}(u)$  have the distance

$$\xi(u) \stackrel{\text{def}}{=} \check{\mathbf{x}}(u) - \mathbf{x}(u), \quad (\text{C.1})$$

where  $u$  is the proper time or distance.

The mathematical descriptions of these geodesics are

$$\check{\ddot{\mathbf{x}}} + (\mathbf{I}_4 \otimes \dot{\check{\mathbf{x}}}^\top) \check{\Gamma} \check{\dot{\mathbf{x}}} = \mathbf{0}, \quad (\text{C.2})$$

$$\ddot{\mathbf{x}} + (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \Gamma \dot{\mathbf{x}} = \mathbf{0}. \quad (\text{C.3})$$

The Christoffel-matrix  $\check{\Gamma}$  is approximated by

$$\check{\Gamma} \approx \Gamma + \frac{\partial \Gamma}{\partial \mathbf{x}^\top} (\xi \otimes \mathbf{I}_4). \quad (\text{C.4})$$

Subtracting (C.3) from (C.2) and considering (C.1) and (C.4), one obtains

$$\check{\ddot{\xi}} + (\mathbf{I}_4 \otimes \dot{\check{\mathbf{x}}}^\top) \Gamma \check{\dot{\mathbf{x}}} - (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \Gamma \dot{\mathbf{x}} + (\mathbf{I}_4 \otimes \dot{\check{\mathbf{x}}}^\top) \frac{\partial \Gamma}{\partial \mathbf{x}^\top} (\xi \otimes \mathbf{I}_4) \check{\dot{\mathbf{x}}} = \mathbf{0}. \quad (\text{C.5})$$

With  $\dot{\mathbf{x}} = \dot{\boldsymbol{\xi}} + \dot{\mathbf{x}}$  and neglecting quadratic and higher powers of  $\boldsymbol{\xi}$  and  $\dot{\boldsymbol{\xi}}$ , one obtains from (C.5)

$$\ddot{\boldsymbol{\xi}} + (\mathbf{I}_4 \otimes \dot{\boldsymbol{\xi}}^\top) \boldsymbol{\Gamma} \dot{\mathbf{x}} + (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \boldsymbol{\Gamma} \dot{\boldsymbol{\xi}} + (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \frac{\partial \boldsymbol{\Gamma}}{\partial \mathbf{x}^\top} (\boldsymbol{\xi} \otimes \mathbf{I}_4) \dot{\mathbf{x}} = \mathbf{0}. \quad (\text{C.6})$$

Hence

$$\frac{D\boldsymbol{\xi}}{du} = \dot{\boldsymbol{\xi}} + (\mathbf{I}_4 \otimes \boldsymbol{\xi}^\top) \boldsymbol{\Gamma} \dot{\mathbf{x}} \quad (\text{C.7})$$

and

$$\begin{aligned} \frac{D^2\boldsymbol{\xi}}{du^2} &= \frac{D}{du} (\dot{\boldsymbol{\xi}} + (\mathbf{I}_4 \otimes \boldsymbol{\xi}^\top) \boldsymbol{\Gamma} \dot{\mathbf{x}}) \\ &= \ddot{\boldsymbol{\xi}} + \frac{d}{du} \{ (\mathbf{I}_4 \otimes \boldsymbol{\xi}^\top) \boldsymbol{\Gamma} \dot{\mathbf{x}} \} + (\mathbf{I}_4 \otimes [\dot{\boldsymbol{\xi}} + (\mathbf{I}_4 \otimes \boldsymbol{\xi}^\top) \boldsymbol{\Gamma} \dot{\mathbf{x}}]^\top) \boldsymbol{\Gamma} \dot{\mathbf{x}} \\ &= \ddot{\boldsymbol{\xi}} + \frac{d}{du} \{ (\mathbf{I}_4 \otimes \boldsymbol{\xi}^\top) \boldsymbol{\Gamma} \dot{\mathbf{x}} \} + (\mathbf{I}_4 \otimes \dot{\boldsymbol{\xi}}^\top) \boldsymbol{\Gamma} \dot{\mathbf{x}} + (\mathbf{I}_4 \otimes [(\mathbf{I}_4 \otimes \boldsymbol{\xi}^\top) \boldsymbol{\Gamma} \dot{\mathbf{x}}]^\top) \boldsymbol{\Gamma} \dot{\mathbf{x}}. \end{aligned} \quad (\text{C.8})$$

For the second term, by (C.3), one gets

$$\begin{aligned} \frac{d}{du} \{ (\mathbf{I}_4 \otimes \boldsymbol{\xi}^\top) \boldsymbol{\Gamma} \dot{\mathbf{x}} \} &= (\mathbf{I}_4 \otimes \dot{\boldsymbol{\xi}}^\top) \boldsymbol{\Gamma} \dot{\mathbf{x}} + (\mathbf{I}_4 \otimes \boldsymbol{\xi}^\top) \frac{\partial \boldsymbol{\Gamma}}{\partial \mathbf{x}^\top} (\dot{\mathbf{x}} \otimes \mathbf{I}_4) \dot{\mathbf{x}} + (\mathbf{I}_4 \otimes \boldsymbol{\xi}^\top) \boldsymbol{\Gamma} \ddot{\mathbf{x}} \\ &= (\mathbf{I}_4 \otimes \dot{\boldsymbol{\xi}}^\top) \boldsymbol{\Gamma} \dot{\mathbf{x}} + (\mathbf{I}_4 \otimes \boldsymbol{\xi}^\top) \frac{\partial \boldsymbol{\Gamma}}{\partial \mathbf{x}^\top} (\dot{\mathbf{x}} \otimes \mathbf{I}_4) \dot{\mathbf{x}} \\ &\quad - (\mathbf{I}_4 \otimes \boldsymbol{\xi}^\top) \boldsymbol{\Gamma} (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \boldsymbol{\Gamma} \dot{\mathbf{x}}. \end{aligned} \quad (\text{C.9})$$

Equation (C.9) used in (C.8) yields

$$\begin{aligned} \frac{D^2\boldsymbol{\xi}}{du^2} &= \ddot{\boldsymbol{\xi}} + (\mathbf{I}_4 \otimes \dot{\boldsymbol{\xi}}^\top) \boldsymbol{\Gamma} \dot{\mathbf{x}} + (\mathbf{I}_4 \otimes \boldsymbol{\xi}^\top) \frac{\partial \boldsymbol{\Gamma}}{\partial \mathbf{x}^\top} (\dot{\mathbf{x}} \otimes \mathbf{I}_4) \dot{\mathbf{x}} - (\mathbf{I}_4 \otimes \boldsymbol{\xi}^\top) \boldsymbol{\Gamma} (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \boldsymbol{\Gamma} \dot{\mathbf{x}} \\ &\quad + (\mathbf{I}_4 \otimes \dot{\boldsymbol{\xi}}^\top) \boldsymbol{\Gamma} \dot{\mathbf{x}} + (\mathbf{I}_4 \otimes [(\mathbf{I}_4 \otimes \boldsymbol{\xi}^\top) \boldsymbol{\Gamma} \dot{\mathbf{x}}]^\top) \boldsymbol{\Gamma} \dot{\mathbf{x}}. \end{aligned} \quad (\text{C.10})$$

*Remark* Since the sub-matrices  $\boldsymbol{\Gamma}_i$  are symmetric, it is generally true that

$$(\mathbf{I}_4 \otimes \mathbf{a}^\top) \boldsymbol{\Gamma} \mathbf{b} = (\mathbf{I}_4 \otimes \mathbf{b}^\top) \boldsymbol{\Gamma} \mathbf{a}. \quad (\text{C.11})$$

In addition, one has  $(\mathbf{I}_4 \otimes \mathbf{a}^\top) \boldsymbol{\Gamma} \mathbf{b} = \overline{\boldsymbol{\Gamma}} (\mathbf{I}_4 \otimes \mathbf{a}) \mathbf{b} = \overline{\boldsymbol{\Gamma}} (\mathbf{b} \otimes \mathbf{a})$  and  $(\mathbf{I}_4 \otimes \mathbf{b}^\top) \boldsymbol{\Gamma} \mathbf{a} = \overline{\boldsymbol{\Gamma}} (\mathbf{I}_4 \otimes \mathbf{b}) \mathbf{a} = \overline{\boldsymbol{\Gamma}} (\mathbf{a} \otimes \mathbf{b})$ , thus, due to (C.11),

$$\overline{\boldsymbol{\Gamma}} (\mathbf{b} \otimes \mathbf{a}) = \overline{\boldsymbol{\Gamma}} (\mathbf{a} \otimes \mathbf{b}). \quad (\text{C.12})$$

With (C.11), one has from (C.10)

$$\ddot{\boldsymbol{\xi}} + (\mathbf{I}_4 \otimes \dot{\boldsymbol{\xi}}^\top) \boldsymbol{\Gamma} \dot{\mathbf{x}} + (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \boldsymbol{\Gamma} \dot{\boldsymbol{\xi}}$$

$$\begin{aligned}
&= \frac{D^2 \xi}{du^2} - (\mathbf{I}_4 \otimes \xi^\top) \frac{\partial \Gamma}{\partial \mathbf{x}^\top} (\dot{\mathbf{x}} \otimes \mathbf{I}_4) \dot{\mathbf{x}} \\
&\quad + (\mathbf{I}_4 \otimes \xi^\top) \Gamma (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \Gamma \dot{\mathbf{x}} - (\mathbf{I}_4 \otimes [(\mathbf{I}_4 \otimes \xi^\top) \Gamma \dot{\mathbf{x}}]^\top) \Gamma \dot{\mathbf{x}}. \tag{C.13}
\end{aligned}$$

For  $(\mathbf{I}_4 \otimes \xi^\top) \Gamma (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \Gamma \dot{\mathbf{x}}$  one can write

$$(\mathbf{I}_4 \otimes \xi^\top) \Gamma (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \Gamma \dot{\mathbf{x}} = \underline{\underline{(\mathbf{I}_4 \otimes \xi^\top) \Gamma \bar{\Gamma} (\mathbf{I}_4 \otimes \dot{\mathbf{x}}) \dot{\mathbf{x}}}}, \tag{C.14}$$

and the expression  $(\mathbf{I}_4 \otimes [(\mathbf{I}_4 \otimes \xi^\top) \Gamma \dot{\mathbf{x}}]^\top) \Gamma \dot{\mathbf{x}}$  can be rewritten as

$$\begin{aligned}
&(\mathbf{I}_4 \otimes [(\mathbf{I}_4 \otimes \xi^\top) \Gamma \dot{\mathbf{x}}]^\top) \Gamma \dot{\mathbf{x}} \\
&= \bar{\Gamma} (\mathbf{I}_4 \otimes (\mathbf{I}_4 \otimes \xi^\top) \Gamma \dot{\mathbf{x}}) \dot{\mathbf{x}} \\
&= \bar{\Gamma} (\mathbf{I}_{16} \otimes \xi^\top) (\mathbf{I}_4 \otimes \Gamma \dot{\mathbf{x}}) \dot{\mathbf{x}} = \underline{\underline{(\mathbf{I}_4 \otimes \xi^\top) (\bar{\Gamma} \otimes \mathbf{I}_4) (\mathbf{I}_4 \otimes \Gamma) (\mathbf{I}_4 \otimes \dot{\mathbf{x}}) \dot{\mathbf{x}}}}. \tag{C.15}
\end{aligned}$$

With (C.14) (in somewhat modified form) and (C.15) one obtains for (C.13)

$$\begin{aligned}
&\ddot{\xi} + (\mathbf{I}_4 \otimes \dot{\xi}^\top) \Gamma \dot{\mathbf{x}} + (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \Gamma \dot{\xi} \\
&= \frac{D^2 \xi}{du^2} - (\mathbf{I}_4 \otimes \xi^\top) \frac{\partial \Gamma}{\partial \mathbf{x}^\top} (\dot{\mathbf{x}} \otimes \mathbf{I}_4) \dot{\mathbf{x}} \\
&\quad + (\mathbf{I}_4 \otimes \xi^\top) [\Gamma \bar{\Gamma} - (\bar{\Gamma} \otimes \mathbf{I}_4) (\mathbf{I}_4 \otimes \Gamma)] (\dot{\mathbf{x}} \otimes \dot{\mathbf{x}}). \tag{C.16}
\end{aligned}$$

Equation (C.16) used in (C.6) provides

$$\begin{aligned}
\frac{D^2 \xi}{du^2} &= -(\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \frac{\partial \Gamma}{\partial \mathbf{x}^\top} (\xi \otimes \mathbf{I}_4) \dot{\mathbf{x}} + (\mathbf{I}_4 \otimes \xi^\top) \frac{\partial \Gamma}{\partial \mathbf{x}^\top} (\dot{\mathbf{x}} \otimes \mathbf{I}_4) \dot{\mathbf{x}} \\
&\quad + (\mathbf{I}_4 \otimes \xi^\top) [\Gamma \bar{\Gamma} - (\bar{\Gamma} \otimes \mathbf{I}_4) (\mathbf{I}_4 \otimes \Gamma)] (\dot{\mathbf{x}} \otimes \dot{\mathbf{x}}). \tag{C.17}
\end{aligned}$$

As the  $16 \times 16$ -matrix  $\frac{\partial \Gamma}{\partial \mathbf{x}^\top}$  is symmetric, the first term of the right-hand side can be transformed as follows:

$$\begin{aligned}
(\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \frac{\partial \Gamma}{\partial \mathbf{x}^\top} (\xi \otimes \mathbf{I}_4) \dot{\mathbf{x}} &= (\mathbf{I}_4 \otimes \dot{\mathbf{x}}^\top) \frac{\partial \Gamma}{\partial \mathbf{x}^\top} \mathbf{U}_{4 \times 4} (\mathbf{I}_4 \otimes \xi) \dot{\mathbf{x}} \\
&= (\mathbf{I}_4 \otimes \xi^\top) \frac{\partial \Gamma}{\partial \mathbf{x}^\top} \mathbf{U}_{4 \times 4} (\mathbf{I}_4 \otimes \dot{\mathbf{x}}) \dot{\mathbf{x}}.
\end{aligned}$$

This in (C.17) provides

$$\begin{aligned}
\frac{D^2 \xi}{du^2} &= (\mathbf{I}_4 \otimes \xi^\top) \left[ \frac{\partial \Gamma}{\partial \mathbf{x}^\top} - \frac{\partial \Gamma}{\partial \mathbf{x}^\top} \mathbf{U}_{4 \times 4} \right] (\dot{\mathbf{x}} \otimes \mathbf{I}_4) \dot{\mathbf{x}} \\
&\quad + (\mathbf{I}_4 \otimes \xi^\top) [\Gamma \bar{\Gamma} - (\bar{\Gamma} \otimes \mathbf{I}_4) (\mathbf{I}_4 \otimes \Gamma)] (\dot{\mathbf{x}} \otimes \dot{\mathbf{x}}), \tag{C.18}
\end{aligned}$$

and finally,

$$\frac{D^2 \xi}{du^2} = (\mathbf{I}_4 \otimes \xi^\top) \underbrace{\left[ \frac{\partial \Gamma}{\partial \mathbf{x}^\top} (\mathbf{I}_{16} - \mathbf{U}_{4 \times 4}) + (\Gamma \bar{\Gamma} - (\bar{\Gamma} \otimes \mathbf{I}_4)(\mathbf{I}_4 \otimes \Gamma)) \right]}_{-\mathbf{R}} (\dot{\mathbf{x}} \otimes \dot{\mathbf{x}}). \quad (\text{C.19})$$

Using a slightly modified Riemannian curvature matrix  $\mathbf{R}$ , we finally obtain for the dynamic behavior of the geodesic deviation

$$\boxed{\frac{D^2 \xi}{du^2} + (\mathbf{I}_4 \otimes \xi^\top) \mathbf{R} (\dot{\mathbf{x}} \otimes \dot{\mathbf{x}}) = \mathbf{0}.} \quad (\text{C.20})$$

In a flat manifold, i.e. in a gravity-free space one has  $\mathbf{R} \equiv \mathbf{0}$  and in Cartesian coordinates  $D/du = d/du$  so that (C.20) reduces to the equation  $d^2 \xi / du^2 = \mathbf{0}$  whose solution is the linear relationship  $\xi(u) = \dot{\xi}_0 \cdot u + \xi_0$ . If  $\mathbf{R} \neq \mathbf{0}$ , gravity exists and the solution of (C.20) is nonlinear, curved.

# Appendix D

## Another Ricci-Matrix

The Ricci-matrix  $R_{\text{Ric}}$  is now defined as the sum of the sub-matrices on the main diagonal of  $R$

$$R_{\text{Ric}} \stackrel{\text{def}}{=} \sum_{\nu=0}^3 R^{\nu\nu}. \tag{D.1}$$

Analogously, we define

$$\check{R}_{\text{Ric}} \stackrel{\text{def}}{=} \sum_{\nu=0}^3 \check{R}^{\nu\nu}. \tag{D.2}$$

From (2.194) it can immediately be read that the Ricci-matrix  $\check{R}_{\text{Ric}}$  is *symmetric* because  $\check{R}_{\alpha\beta}^{\gamma\gamma} = \check{R}_{\beta\alpha}^{\gamma\gamma}$ . It is also true that

$$R = (G^{-1} \otimes I_4) \check{R},$$

so

$$R^{\gamma\delta} = (g_{\gamma}^{-T} \otimes I_4) \check{R}^{\delta} = \sum_{\nu=0}^3 g_{\gamma\nu}^{[-1]} \check{R}^{\nu\delta}, \tag{D.3}$$

where  $g_{\gamma}^{-T}$  is the  $\gamma$ th row of  $G^{-1}$  and  $\check{R}^{\delta}$  is the matrix consisting of the sub-matrices in the  $\delta$ th block column of  $\check{R}$ , i.e. the matrix elements are

$$R_{\alpha\beta}^{\gamma\delta} = \sum_{\nu=0}^3 g_{\gamma\nu}^{[-1]} \check{R}_{\alpha\beta}^{\nu\delta}. \tag{D.4}$$

With the help of (D.3), the Ricci-matrix is obtained as

$$R_{\text{Ric}} = \sum_{\gamma} R^{\gamma\gamma} = \sum_{\gamma} \sum_{\nu} g_{\gamma\nu}^{[-1]} \check{R}^{\nu\nu}, \tag{D.5}$$



i.e. for the components one has

$$R_{\text{Ric},\alpha\beta} = \sum_{\gamma} \sum_{\nu} g_{\gamma\nu}^{[-1]} \check{R}_{\alpha\beta}^{\nu\gamma}, \quad (\text{D.6})$$

or with (2.173)

$$R_{\text{Ric},\alpha\beta} = \sum_{\gamma} \sum_{\nu} g_{\gamma\nu}^{[-1]} \check{R}_{\nu\gamma}^{\alpha\beta}. \quad (\text{D.7})$$

The *curvature scalar*  $R$  is obtained from the Ricci-matrix by taking the trace

$$R \stackrel{\text{def}}{=} \sum_{\alpha} R_{\text{Ric},\alpha\alpha} = \sum_{\alpha} \sum_{\gamma} \sum_{\nu} g_{\gamma\nu}^{[-1]} \check{R}_{\nu\gamma}^{\alpha\alpha} = \sum_{\gamma} \sum_{\nu} g_{\gamma\nu}^{[-1]} \check{R}_{\text{Ric},\nu\gamma}. \quad (\text{D.8})$$

Conversely, we obtain a corresponding relationship

$$\check{R}_{\alpha\beta}^{\gamma\delta} = \sum_{\nu=0}^3 g_{\gamma\nu} R_{\alpha\beta}^{\nu\delta}. \quad (\text{D.9})$$

From (2.165) it directly follows that

$$R_{\text{Ric},\alpha\beta} = \sum_{\gamma=0}^3 \left( \frac{\partial}{\partial x_{\beta}} \Gamma_{\alpha\gamma}^{\gamma} - \frac{\partial}{\partial x_{\gamma}} \Gamma_{\alpha\beta}^{\gamma} + \sum_{\nu=0}^3 \Gamma_{\beta\nu}^{\gamma} \Gamma_{\gamma\alpha}^{\nu} - \sum_{\nu=0}^3 \Gamma_{\gamma\nu}^{\gamma} \Gamma_{\alpha\beta}^{\nu} \right) \quad (\text{D.10})$$

and from (2.168)

$$\check{R}_{\text{Ric},\alpha\beta} = \sum_{\gamma=0}^3 \left( \frac{\partial}{\partial x_{\beta}} \check{\Gamma}_{\alpha\gamma}^{\gamma} - \frac{\partial}{\partial x_{\gamma}} \check{\Gamma}_{\alpha\beta}^{\gamma} + \sum_{\nu=0}^3 \Gamma_{\alpha\beta}^{\nu} \check{\Gamma}_{\gamma\gamma}^{\nu} - \sum_{\nu=0}^3 \Gamma_{\alpha\gamma}^{\nu} \check{\Gamma}_{\gamma\beta}^{\nu} \right). \quad (\text{D.11})$$

**Symmetry of the Ricci-Matrix  $R_{\text{Ric}}$**  Even if  $R$  itself is not symmetric, the from  $R$  derived Ricci-matrix  $R_{\text{Ric}}$  is symmetric; this will be shown in the following. The symmetry will follow from the components equation (D.10) of the Ricci-matrix. One sees immediately that the second and fourth summands are symmetric in  $\alpha$  and  $\beta$ .

The symmetry of the term  $\sum_{\gamma=0}^3 \frac{\partial}{\partial x_{\beta}} \Gamma_{\alpha\gamma}^{\gamma}$  in  $\alpha$  and  $\beta$  is not seen directly. This can be checked using the Laplace-expansion theorem for determinants.<sup>1</sup> Developing the determinant of  $G$  along the  $\gamma$ th row yields

$$g \stackrel{\text{def}}{=} \det(G) = g_{\gamma 1} A_{\gamma 1} + \cdots + g_{\gamma \beta} A_{\gamma \beta} + \cdots + g_{\gamma n} A_{\gamma n},$$

<sup>1</sup>The sum of the products of all elements of a row (or column) with their adjuncts is equal to the determinant's value.

where  $A_{\gamma\beta}$  is the element in the  $\gamma$ th row and  $\beta$ th column of the adjoint of  $\mathbf{G}$ . If  $g_{\beta\gamma}^{[-1]}$  is the  $(\beta\gamma)$ th element of the inverse of  $\mathbf{G}$ , then  $g_{\beta\gamma}^{[-1]} = \frac{1}{g} A_{\gamma\beta}$ , so  $A_{\gamma\beta} = g g_{\beta\gamma}^{[-1]}$ . Thus we obtain

$$\frac{\partial g}{\partial g_{\gamma\beta}} = A_{\gamma\beta} = g g_{\beta\gamma}^{[-1]},$$

or

$$\delta g = g g_{\beta\gamma}^{[-1]} \delta g_{\gamma\beta},$$

or

$$\frac{\partial g}{\partial x_\alpha} = g g_{\beta\gamma}^{[-1]} \frac{\partial g_{\gamma\beta}}{\partial x_\alpha},$$

i.e.

$$\frac{1}{g} \frac{\partial g}{\partial x_\alpha} = g_{\beta\gamma}^{[-1]} \frac{\partial g_{\gamma\beta}}{\partial x_\alpha}. \tag{D.12}$$

Using (2.62), on the other hand, one has

$$\sum_{\gamma=0}^3 \Gamma_{\alpha\gamma}^\gamma = \sum_{\gamma=0}^3 \sum_{\beta=0}^3 \frac{g_{\beta\gamma}^{[-1]}}{2} \left( \frac{\partial g_{\gamma\beta}}{\partial x_\alpha} + \frac{\partial g_{\alpha\beta}}{\partial x_\gamma} - \frac{\partial g_{\alpha\gamma}}{\partial x_\beta} \right),$$

i.e. the last two summands cancel out and it remains to deal with

$$\sum_{\gamma=0}^3 \Gamma_{\alpha\gamma}^\gamma = \sum_{\gamma=0}^3 \sum_{\beta=0}^3 \frac{1}{2} g_{\beta\gamma}^{[-1]} \frac{\partial g_{\gamma\beta}}{\partial x_\alpha}.$$

It follows from (D.12) that

$$\sum_{\gamma=0}^3 \frac{\partial}{\partial x_\beta} \Gamma_{\alpha\gamma}^\gamma = \sum_{\gamma=0}^3 \sum_{\beta=0}^3 \frac{1}{\sqrt{|g|}} \frac{\partial^2 \sqrt{|g|}}{\partial x_\alpha \partial x_\beta}. \tag{D.13}$$

But this form is immediately seen symmetric in  $\alpha$  and  $\beta$ .

Now it remains to shown that the third term in (D.10) is symmetric. Its expression is

$$\sum_{\gamma=0}^3 \sum_{v=0}^3 \Gamma_{\beta v}^\gamma \Gamma_{\gamma\alpha}^v.$$

Now one can see that this term is symmetric because

$$\sum_{\gamma,v=0}^3 \Gamma_{\beta v}^\gamma \Gamma_{\gamma\alpha}^v = \sum_{\gamma,v=0}^3 \Gamma_{v\beta}^\gamma \Gamma_{\alpha\gamma}^v = \sum_{v,\gamma=0}^3 \Gamma_{\gamma\beta}^v \Gamma_{\alpha v}^\gamma.$$

All this shows that the Ricci-matrix  $\mathbf{R}_{\text{Ric}}$  is symmetric.

**Divergence of the Ricci-Matrix  $R_{\text{Ric}}$**  Multiplying the Bianchi-identities (2.199) in the form of

$$\frac{\partial}{\partial x_\kappa} R_{\alpha\beta}^{\nu\delta} + \frac{\partial}{\partial x_\beta} R_{\alpha\delta}^{\nu\kappa} + \frac{\partial}{\partial x_\delta} R_{\alpha\kappa}^{\nu\beta} = 0$$

with  $g_{\gamma\nu}$  and summing over  $\nu$ , we obtain in  $\mathcal{P}$ , since there  $\frac{\partial G}{\partial x} = \mathbf{0}$ ,

$$\frac{\partial}{\partial x_\kappa} \sum_{\nu=0}^3 g_{\gamma\nu} R_{\alpha\beta}^{\nu\delta} + \frac{\partial}{\partial x_\beta} \sum_{\nu=0}^3 g_{\gamma\nu} R_{\alpha\delta}^{\nu\kappa} + \frac{\partial}{\partial x_\delta} \sum_{\nu=0}^3 g_{\gamma\nu} R_{\alpha\kappa}^{\nu\beta} = 0.$$

Combined with (D.9) this becomes

$$\frac{\partial}{\partial x_\kappa} \check{R}_{\alpha\beta}^{\gamma\delta} + \frac{\partial}{\partial x_\beta} \check{R}_{\alpha\delta}^{\gamma\kappa} + \frac{\partial}{\partial x_\delta} \check{R}_{\alpha\kappa}^{\gamma\beta} = 0. \quad (\text{D.14})$$

The second term, according to (2.172), may be written as

$$-\frac{\partial}{\partial x_\beta} \check{R}_{\alpha\kappa}^{\gamma\delta}.$$

Substituting now  $\gamma = \delta$  and summing over  $\gamma$ , one obtains

$$\frac{\partial}{\partial x_\kappa} \check{R}_{\text{Ric},\alpha\beta} - \frac{\partial}{\partial x_\beta} \check{R}_{\text{Ric},\alpha\kappa} + \sum_{\gamma=0}^3 \frac{\partial}{\partial x_\gamma} \check{R}_{\alpha\kappa}^{\gamma\beta} = 0. \quad (\text{D.15})$$

In the third summand, one can, according to (2.171), replace  $\check{R}_{\alpha\kappa}^{\gamma\beta}$  by  $-\check{R}_{\gamma\kappa}^{\alpha\beta}$ . If we set  $\alpha = \beta$  and sum over  $\alpha$ , we obtain for (D.15) with the trace  $\check{R} \stackrel{\text{def}}{=} \sum_{\alpha=0}^3 \check{R}_{\text{Ric},\alpha\alpha}$  of the Riccati-matrix  $\check{R}_{\text{Ric}}$

$$\frac{\partial}{\partial x_\kappa} \check{R} - \sum_{\alpha=0}^3 \frac{\partial}{\partial x_\alpha} \check{R}_{\text{Ric},\alpha\kappa} - \sum_{\gamma=0}^3 \frac{\partial}{\partial x_\gamma} \check{R}_{\text{Ric},\gamma\kappa} = 0. \quad (\text{D.16})$$

If in the last sum the summation index  $\gamma$  is replaced by  $\alpha$ , we can finally summarize:

$$\frac{\partial}{\partial x_\kappa} \check{R} - 2 \sum_{\alpha=0}^3 \frac{\partial}{\partial x_\alpha} \check{R}_{\text{Ric},\alpha\kappa} = 0. \quad (\text{D.17})$$

One would get the same result, if one were to start with the equation:

$$\frac{\partial}{\partial x_\kappa} \check{R}_{\alpha\beta}^{\gamma\delta} - 2 \frac{\partial}{\partial x_\beta} \check{R}_{\alpha\kappa}^{\gamma\delta} = 0. \quad (\text{D.18})$$

Indeed, if we set  $\delta = \gamma$  and sum over  $\gamma$ , we get first

$$\frac{\partial}{\partial x_\kappa} \check{R}_{\text{Ric},\alpha\beta} - 2 \frac{\partial}{\partial x_\beta} \check{R}_{\text{Ric},\alpha\kappa} = 0.$$

If we now set  $\alpha = \beta$  and sum over  $\alpha$ , we will again arrive at (D.17).

A different result is obtained when starting from (D.18) (with  $\nu$  instead of  $\gamma$ ) first, multiplying this equation by  $g_{\gamma\nu}^{[-1]}$  to get

$$\frac{\partial}{\partial x_\kappa} g_{\gamma\nu}^{[-1]} \check{R}_{\alpha\beta}^{\nu\delta} - 2 \frac{\partial}{\partial x_\beta} g_{\gamma\nu}^{[-1]} \check{R}_{\alpha\kappa}^{\nu\delta} = 0,$$

then again setting  $\gamma = \delta$  and summing over  $\gamma$  and  $\nu$  and noting (D.6):

$$\begin{aligned} & \sum_\gamma \sum_\nu \frac{\partial}{\partial x_\kappa} g_{\gamma\nu}^{[-1]} \check{R}_{\alpha\beta}^{\nu\gamma} - 2 \sum_\gamma \sum_\nu \frac{\partial}{\partial x_\beta} g_{\gamma\nu}^{[-1]} \check{R}_{\alpha\kappa}^{\nu\gamma} \\ &= \frac{\partial}{\partial x_\kappa} R_{\text{Ric},\alpha\beta} - 2 \frac{\partial}{\partial x_\beta} R_{\text{Ric},\alpha\kappa} = 0. \end{aligned}$$

If we now take  $\alpha = \eta$  and sum over  $\alpha$ , we finally obtain the important relationship

$$\frac{\partial}{\partial x_\kappa} R - 2 \sum_\alpha \frac{\partial}{\partial x_\alpha} R_{\text{Ric},\alpha\kappa} = 0. \quad (\text{D.19})$$

These are the four equations for the four spacetime coordinates  $x_0, \dots, x_3$ .

Finally, this overall result can be represented as

$$\underline{\underline{\vec{V}^\top \left( \mathbf{R}_{\text{Ric}} - \frac{1}{2} R \mathbf{I}_4 \right) = \mathbf{0}^\top.}} \quad (\text{D.20})$$

# References

1. V.I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer, Berlin, 1978)
2. G. Barton, *Introduction to the Relativity Principle* (Wiley, New York, 1999)
3. H. Bondi, *Relativity and Common Sense* (Dover Publications, Dover, 1986)
4. J.W. Brewer, Kronecker products and matrix calculus in system theory. *IEEE Trans. Circuits Syst.* 772–781 (1978)
5. J.J. Callahan, *The Geometry of Spacetime* (Springer, Berlin, 2000)
6. S. Carrol, *Spacetime and Geometry: An Introduction to General Relativity* (Addison Wesley, San Francisco, 2003)
7. S. Chandrasekhar, *The Mathematical Theory of Black Holes* (Clarendon Press, Oxford, 1983)
8. I. Ciufolini, J.A. Wheeler, *Gravitation and Inertia* (Princeton University Press, Princeton, 1995)
9. R. D’Inverno, *Introducing Einstein’s Relativity* (Clarendon Press, New York, 1992)
10. A. Einstein, *Relativity, The Special and the General Theory*, 2nd edn. (Edit Benei Noaj, 2007)
11. A. Föppl, Über einen Kreisversuch zur Messung der Umdrehungsgeschwindigkeit der Erde. *Sitzungsber. Bayer. Akad. Wiss.* 34, 5–28 (1904)
12. J. Foster, J.D. Nightingale, *A Short Course in General Relativity* (Springer, Berlin, 1995)
13. R.G. Gass, F.P. Esposito, L.C.R. Wijewardhansa, L. Witten, Detecting event horizons and stationary surfaces (1998). [arXiv:gr-qc/9808055v](https://arxiv.org/abs/gr-qc/9808055)
14. R. Geroch, *Relativity from A to B* (Chicago University Press, Chicago, 1981)
15. O. Gron, S. Hervik, *Einstein’s General Theory of Relativity* (Springer, New York, 2007)
16. O. Gron, A. Naess, *Einstein’s Theory* (Springer, New York, 2011)
17. J.B. Hartle, *Gravity: An Introduction to Einstein’s General Relativity* (Addison Wesley, San Francisco, 2003)
18. S.W. Hawking, G.F.R. Ellis, *The Large Scale Structure of Space-Time* (Cambridge University Press, Cambridge, 1973)
19. I.R. Kenyon, *General Relativity* (Oxford University Press, Oxford, 1990)
20. L.D. Landau, E.M. Lifschitz, *Classical Theory of Fields: 2. Course of Theoretical Physics Series* (Butterworth Heinemann, Stoneham, 1987)
21. C.W. Misner, K.S. Thorne, J.A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973)
22. R.A. Mould, *Basic Relativity* (Springer, Berlin, 2002)
23. J. Natario, *General Relativity Without Calculus* (Springer, Berlin, 2011)
24. W. Rindler, *Essential Relativity* (Springer, Berlin, 1977)
25. L. Sartori, *Understanding Relativity: A Simplified Approach to Einstein’s Theories* (University of California Press, Berkeley, 1996)
26. B.F. Schutz, *A First Course in General Relativity* (Cambridge University Press, Cambridge, 1985)
27. H. Stephani, *General Relativity* (Cambridge University Press, Cambridge, 1977)

28. J. Stewart, *Advanced General Relativity* (Cambridge University Press, Cambridge, 1991)
29. T. Takeuchi, *An Illustrated Guide to Relativity* (Cambridge University Press, Cambridge, 2010)
30. E.F. Taylor, J.A. Wheeler, *Spacetime Physics* (Freeman, New York, 1992)
31. E.F. Taylor, J.A. Wheeler, *Black Holes* (Addison Wesley, Reading, 2000)
32. E.F. Taylor, J.A. Wheeler, *Exploring Black Holes: Introduction to General Relativity* (Addison Wesley, San Francisco, 2000)
33. M. von Laue, *Die Relativitätstheorie. Erster Band: Die spezielle Relativitätstheorie*, 7th edn. (Vieweg, Wiesbaden, 1961)
34. R. Wald, *General Relativity* (Chicago University Press, Chicago, 1984)
35. R. Wald, *Space, Time and Gravity: Theory of the Big Bang and Black Holes* (University Press, Chicago, 1992)
36. S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972)
37. W.J. Wild, A matrix formulation of Einstein's vacuum field equations, [gr-qc/9812095](#), 31 Dec. 1998

# Index

## A

- Acceleration, 3, 26
  - vector
    - fourdimensional, 27
- Action functional
  - Hilbert's, 101
- Action-at-a-distance-equation, 98
- Aether, 7
- Affine mapping, 6

## B

- Bianchi, 186
- Black hole, 109, 116, 127, 129
  - rotating, 138
- Block-transposed matrix, 63
- Brown dwarf, 128

## C

- Calculus of variation, 57, 58
- Carbon, 128
- Chandrasekhar, 138
- Change
  - length, 117
- Charge, 6
- Circle
  - great, 173
- Conservation of mass, 44
- Constant
  - gravitational, 97
- Contraction, 13
  - time, 118
- Coordinate
  - Boyer–Lindquist, 139
  - Eddington's, 137, 142
  - spherical, 109
- Coordinate system, 51
- Covariant, 99

- Current, 32
- Curvature, 78
  - Gaussian, 166, 177
  - definition, 168
  - Gauss, 169
  - normal, 175
  - principal, 177
- Curvature matrix
  - Riemannian, 174
- Curvature scalar, 92
- Curve
  - geodesic, 57, 68
  - smooth, 166
- Curved space, 66
- Cylinder, 175

## D

- Definite
  - positive, 175
- Density
  - energy, 41
  - Lagrange, 102
  - scalar, 102
  - Schwarzschild's, 128, 142
- Derivation, 160
  - covariant, 69
  - partial, 69
- Derivative
  - absolute, 68
  - covariant, 75
- Deviation
  - geodesic, 179
- Direction, 165
  - main, 175
  - normal, 168
- Displacement
  - parallel, 79

- Dynamics
  - Newton's, 101
- E**
- Earth, 128
- Eddington, 124
- Effect
  - Lense–Thirring, 109, 139
- Einstein, 7, 93, 101
- Einsteins field equation, 98
- Electric field, 6, 32
- Ellipsoid, 173, 175
- Energy, 30
  - kinetic, 31
- Equation
  - Einstein's, 102
  - Einstein's field, 98
  - Euler/Lagrange, 103
  - Euler's, 44, 46
  - Maxwell's, 32
  - Poisson, 78, 94, 97, 101
- Equation of motion
  - relativistic, 29
- Equivalence principle, 99
- Event, 1
  - time-like, 64
- Experiment
  - Michelson–Morley, 7, 15
- F**
- Field
  - gravitomagnetic, 140
- Field equation
  - Einstein's, 101
- Field strength
  - electric, 38
- Fluid
  - perfect, 44
- Force
  - Lorentz-, 141
  - centrifugal, 66, 115
  - Coriolis, 66
  - electromagnetic, 7, 39
  - gravitational, 44
- Form
  - quadratic differential, 175
  - second fundamental, 175
- Four-vector, 4
- Fundamental law, 3
- Fusion
  - nuclear, 127
- G**
- Galilei, 3
- Galilei-transformation, 1
- Gauss, 78, 166, 174
- General relativity, 51
- General relativity principle, 7
- Geodesic, 78, 179
- Geometry
  - differential, 78, 165
  - elliptical, 173
  - Euclidean, 173, 174
  - hyperbolic, 173
- Gravitational field, 52, 101
- Gravitomagnetic field, 140
- Gravity, 51
- Gtransformation
  - coordinate, 51
- H**
- Hilbert, 101
- Hyperboloid, 173
- Hypersurface, 174
- I**
- Identity
  - Bianchi, 88, 92
- Induction, 32
  - magnetic, 38
- Inertial system, 3, 53
- Integral
  - action, 103
  - variated, 58
- Interstellar cloud, 127
- Invariance, 51, 67
- Invariant, 99
- Invariant
  - electromagnetic, 38
  - Kretschmann's, 132, 135
- K**
- Kronecker-product, 154, 160
- L**
- Light cone, 17, 18
- Light wave propagation, 7
- Line
  - curved, 165
  - geodesic, 68
  - straight, 57
- Line element, 51
- Liquid
  - neutron, 129
- Lorentz-transformation, 11
- M**
- Magnetic field, 6



**Mass**

- rotating, 138
- solar, 128
- spherical, 109

**Matrix**

- inversion-lemma, 153
- adjoint, 151
- block, 152
- Christoffel, 55, 66, 69, 71, 88, 100, 102
- curvature, 95
- diagonal, 147
- differentiation, 157
- Einstein's, 93, 98
- electromagnetic energy-momentum-, 42
- energy-momentum, 41, 97, 98
- Faraday's, 33
- identity, 147
- inverse, 151
- Jacobi, 52, 55, 65
- Maxwell's, 34
- mechanical energy-momentum-, 43
- Minkowski, 99
- permutation, 54, 156, 160
- product, 150
- Ricci, 89, 90, 92, 97, 98, 101, 113
- Ricci-, 184
- Riemannian curvature, 81, 85, 86, 89, 98, 102, 112
- rotation, 5, 6
- symmetric, 147
- total energy-momentum-, 47
- transposed, 147

**Maxwell, 6****Mechanic**

- classical, 2

**Metric**

- Kerr's, 139
- Schwarzschild's, 116

**Metric matrix, 52****Milky Way, 129****Minkowski, 19****Mirror, 8****N****Neutron star, 127****Newton, 1, 94, 98****Nuclear solar fuel, 128****O****Observer, 2, 8****Operator**

- Laplace, 78
- vec-, 63

**Oxygen, 128****P****Paraboloid**

- hyperbolic, 176

**Paradox**

- twin, 15

**Parallelogram, 168****Phenomena**

- electromagnetic, 7

**Photon, 19, 53****Plane**

- tangent, 166, 168

**Point**

- elliptical, 175
- hyperbolic, 176
- parabolic, 175
- saddle, 176

**Polygon**

- inscribed, 165

**Position**

- initial, 53

**Position vector, 1****Postulate**

- parallel, 173

**Potential**

- gravitational, 77, 97, 115

**Principle**

- covariance, 99
- relativity, 53

**Product**

- Kronecker-, 154

**Proper time, 20****Q****Quadratic form, 57****R****Radiation, 31****Radius**

- curvature, 166
- Schwarzschild's, 116, 128, 141

**Rate of change, 166****Rectilinear, 3****Redshift, 120****Reference frame, 1****Reference system, 2****Relativity principle**

- Galilei, 3

**Rest mass, 30****Riemann, 78, 174****Rotation, 6****S****Scalar**

- curvature, 90, 101, 107

Schwarzschild, 109  
 Shell  
   Schwarzschild's spherical, 131  
 Shift  
   color, 120  
 Ship  
   moving, 3  
 Simultaneity, 12  
 Singularity, 116, 131  
   coordinate, 131, 136  
 Solution  
   Schwarzschild's, 116, 128  
 Space  
   curved, 66  
   homogeneous, 5  
   isotropic, 5  
   n-dimensional, 78  
 Space-time vector, 22  
 Spaceship, 8  
 Spacetime, 19  
 Spacetime-interval, 51  
 Sphere, 175  
   unit, 174  
 Star, 127  
 Supernova, 128  
 Surface  
   curved, 165  
   first fundamental form, 170  
   hyper-, 78  
   second fundamental form, 170  
   spherical, 173  
 Symbol  
   Christoffel-, 172  
 Symmetric, 91  
 System  
   inertial, 3  
   rotating, 64, 97

**T**

Tangent vector, 68  
 Tension  
   Maxwell, 7

Theorem  
   Laplace expansion, 90, 104  
   Pythagorean, 8  
 Theorema Egregium, 166, 169  
 Theory  
   Lagrange/Hamilton, 103  
 Thought experiment, 8  
 Time  
   proper, 19, 53, 64  
 Time dilation, 15  
 Time shifting, 5  
 Total difference, 159  
 Train, 3  
 Transformation  
   coordinate, 72  
   Galilei, 1, 4, 5  
   general Galilei, 5  
   similarity, 75  
 Transport  
   parallel, 78

**V**

Variation  
   calculus of, 57  
 Variational principle, 101  
 Vec-operator, 64  
 Vector  
   Poynting-, 41  
   energy-momentum, 31  
   extended, 6  
   normal, 168  
   normalized tangent, 165  
   spacetime, 53  
 Velocity  
   Lorentz-transformation, 22  
   initial, 53  
 Velocity addition, 20  
 Velocity of light, 52  
 Virgo galaxy, 129

**W**

White dwarf, 127, 128  
 World-aether, 7