Lecture Notes in Mathematics 1808

Springer
Berlin
Heidelberg
New York
Hong Kong
London
Milan
Paris
Tokyo

Werner Schindler

Measures with Symmetry Properties

Springer

Author

Werner Schindler

Bundesamt für Sicherheit in
der Informationstechnik (BSI)
Godesberger Allee 185-189
53175 Bonn, Germany

E-mail: Werner.Schindler@bsi.bund.de

Cataloging-in-Publication Data applied for

A catalog record for this book is available from the Library of Congress.

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie;
detailed bibliographic data is available in the Internet at http://dnb.ddb.de

Mathematics Subject Classification (2000):
28C10, 65C10, 62B05, 22E99

ISSN 0075-8434
ISBN 3-540-00235-9 Springer-Verlag Berlin Heidelberg New York

Springer-Verlag Berlin Heidelberg New York a member of BertelsmannSpringer
Science + Business Media GmbH

http://www.springer.de

Printed in Germany

Typesetting: Camera-ready TEX output by the author

SPIN: 10825787 41/3142/du-543210 - Printed on acid-free paper

In memory of my father

Preface

Symmetry has played an important role in art and architecture since early mankind. Symmetries occur in botany, biology, chemistry and physics. It is therefore not surprising that in various branches of mathematics symmetries and invariance principles have emerged and were put to good use as research tools. A prominent example is measure and integration theory where a rich theory has been developed for measures with remarkable symmetry properties such as Haar measures on locally compact groups, invariant measures on homogeneous spaces, or particular measures on $\mathbb{R}^n$ with special symmetries. In connection with invariant statistics research has focussed on specific situations in which measures that do not have 'full' symmetry occur.

This book also deals with measures having symmetry properties. However, our assumptions on the symmetry properties and on the general framework in which we are working are rather mild. This has two somewhat contradictory consequences. On the one hand, the weakness of the hypotheses we impose and the ensuing generality cause the mathematics with which we have to cope to be deep and hard. For instance, certain techniques which have been very efficient in differential geometry are not available to us. On the other hand, the generality of our approach allows us to apply it to an amazingly wide range of problems. This new scope of applications is the positive side of the generality of our attack.

The main results of this book belong to measure and integration theory and thus to pure mathematics. However, in the second part of the book many concrete applications are discussed in great detail. We consider integration problems, stochastic simulations and statistics. The examples span from computational geometry, applied mathematics, computer aided graphical design to coding theory. This book therefore does not only address pure and applied mathematicians, but also computer scientists and engineers. As a consequence, I have attempted to make the discourse accessible to readers from a great variety of different backgrounds and to keep the prerequisites to a minimum. Only in Chapter 2 where the core theorems are derived did it appear impossible to adhere strictly to this goal. However, I have arranged the presentation in such a fashion that a reader who wishes to skip this chapter completely, can read the remaining parts of the book without loss of under-

standing. As far as they are relevant for the applications, the main results of Chapter 2 are summarized at the beginning of Chapter 4.

This book arose from my postdoctoral thesis (Habilitationsschrift, [71]). I would like to thank Karl Heinrich Hofmann, Jürgen Lehn and Gunter Ritter again for acting as referees and for their valuable advice, Siegfried Graf for providing me with an interesting reference, and Hans Tzschach for the encouragement he has given me to even consider the writing of this thesis. Relative to the postdoctoral thesis, in the present book I have added some new results and a number of illustrative examples. Of course there are editorial adjustments, and the text has been made more 'self-contained'; this should enhance its readability without forcing the reader to consult too many references. I would like to thank Karl Heinrich Hofmann for his advice and support. Finally, I want to thank my wife Susanne for her patience with me when I was writing my postdoctoral thesis and this book.

Sinzig, *Werner Schindler*
October 2002

Table of Contents

1 Introduction

Symmetries and invariance principles play an important role in various branches in mathematics. In measure theory and functional analysis Haar measures on locally compact groups and invariant measures on homogeneous spaces are of particular interest as they have a maximum degree of symmetry. In particular, up to a multiplicative constant they are unique.

This book treats measures which are invariant under group actions. Loosely speaking, we say that these measures have symmetry properties. In general, these measures have a lower degree of symmetry than Haar measures or invariant measures on homogeneous spaces, and usually they are not unique.

This work falls into two large blocks. Chapter 2 may be viewed as its mathematical core as the main results are derived there. The main results and their proofs mainly belong to different areas of measure and integration theory. The importance of their own, their applicability and usefulness for solving specific high-dimensional integrals, for stochastic simulations and in statistics are explained and discussed in detail. In Chapter 4 these insights are used to simplify calculations and to save computing time and memory. Particular problems become practically feasible only with these improvements. We treat examples that range from Grassmannian manifolds and compact connected Lie groups to various matrix groups and coding theory. Under suitable conditions a priori bounds for the error propagation in numerical algorithms for the computation of the polar decomposition of $(n \times n)$-matrices can be derived from a posteriori bounds. From this the necessary machine precision can be determined before the first calculation rather than retrospectively after the final one. Similarly, by exploiting symmetries a whole class of random rotations in $\mathbb{R}^n$ can be simulated efficiently. Before we go into details we illustrate and motivate the essential idea of our concept by a well-known example.

For many computations it turns out to be profitable to use polar, spherical or cylinder coordinates. In particular,

$$\int_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\mathrm{d}y = \int_0^\infty \int_0^{2\pi} r f\left(\varphi_p(\alpha,r)\right)\,\mathrm{d}\alpha\mathrm{d}r \tag{1.1}$$

with $\varphi_p(\alpha,r) := (r\cos\alpha, r\sin\alpha)$. If f is radially symmetric, i.e. if the integrand f does only depend on the norm of its argument, the right-hand

W. Schindler: LNM 1808, pp. 1–9, 2003.

side simplifies to $\int_0^\infty 2\pi r f(r,0)\,\mathrm{d}r$. Moreover, this observation can be used to carry out stochastic simulations (see, e.g. Chapter 3 or [1, 18, 46] etc.) of radially symmetric Lebesgue densities on $\mathbb{R}^2$ in three steps.

Algorithm (Radially symmetric densities in $\mathbb{R}^2$)

1. Generate an equidistributed pseudorandom number $\widetilde{\alpha}$ (pseudorandom angle) on $[0, 2\pi)$
2. Generate a pseudorandom radius $\widetilde{R}$ on $[0, \infty)$ with respect to the Lebesgue density $g(r) := 2\pi r f(r, 0)$.
3. $\widetilde{Z} := \varphi_p(\widetilde{\alpha}, \widetilde{R})$.

The advantages of this approach are obvious. First, a two-dimensional simulation problem is decomposed into two independent one-dimensional simulation problems. Usually, this simplifies the algorithms and increases their efficiency (cf. Chapters 3 and 4). Moreover, it leads to a partial unification of a whole class of simulation problems as the first and the third step are the same for all two-dimensional distributions with radially symmetric Lebesgue density.

Suppose that a statistician has observed a sample $\mathbf{z}_1, \ldots, \mathbf{z}_m \in \mathbb{R}^2$ when repeating a random experiment m times. The statistician interprets the vectors $\mathbf{z}_1, \ldots, \mathbf{z}_m$ as realizations of independent and identically distributed (iid) random vectors $Z_1, \ldots, Z_m$ with unknown distribution ν. Depending on the random experiment it may be reasonable to assume that ν has a radially symmetric Lebesgue density f. In this case the angles of $\mathbf{z}_1, \ldots, \mathbf{z}_m$ with the x-axis do not deliver any useful information on f. (This is immediately obvious, and in Chapter 2 a formal proof will be given in a more general context.) The random vectors $Z_1, \ldots, Z_m$ induce independent random angles which are equidistributed on $[0, 2\pi)$ for *all* distributions with radially symmetric Lebesgue density (more generally, even for all radially symmetric distributions). As a consequence, $\mathbf{z}_1, \ldots, \mathbf{z}_m$ do not provide more information about the unknown distribution than their lengths $\|\mathbf{z}_1\|, \ldots, \|\mathbf{z}_m\|$. Usually, it is much easier to determine a powerful test, resp. to determine its distribution under the null hypothesis or even under all admissible hypotheses, for one-dimensional random variables than for two-dimensional random variables.

These superficial considerations underline the usefulness of polar coordinates and show how radial symmetry in $\mathbb{R}^2$ can be exploited to simplify computations. The generalization of this approach to radially symmetric functions in $\mathbb{R}^n$ and to distributions with radially symmetric densities is obvious. More generally, *radially symmetric distributions* in $\mathbb{R}^n$ are usually characterized by the property that they are invariant under left multiplication of its argument with orthogonal matrices (cf. [18], p. 225). In this context it should be noted that radially symmetric distributions in $\mathbb{R}^n$ represent the most important subclass of *elliptically contoured distributions*, which have

been investigated intensively for the last 20 years. As their name indicates the hypersurfaces of constant mass distribution are ellipsoids. Research activities have considered the characterization of elliptically contoured distributions, possible stochastic representations and the distribution of particular test statistics and parameter estimation problems, for example. However, as elliptically contoured distributions will not be considered in the following we refer the interested reader to the relevant literature (cf. [25, 26, 31, 32], for example).

Returning to our introductory example, what are the underlying phenomena that enable this increase of efficiency? The answer is given by the transformation theorem if it is interpreted suitably. In fact, each radially symmetric Lebesgue density $f\colon \mathbb{R}^2 \to [0,\infty)$ fulfils

$$f \cdot \lambda_2 = \left(\frac{1}{2\pi} \lambda_{[0,2\pi)} \otimes 2\pi r \bar{f} \cdot \lambda_{[0,\infty)} \right)^{\varphi_p} \tag{1.2}$$

with $\bar{f}(r) := f(r,0)$. That is, $f \cdot \lambda_2$ can be expressed as an image of a product measure whose left-hand factor does not depend on f. The individuality of f, i.e. the individuality of the measure $\nu = f \cdot \lambda_2$ is exclusively grounded on the second factor. (The symbols λ_2, $\lambda_{[0,2\pi)}$ and $\lambda_{[0,\infty)}$ stand for the Lebesgue measures on $\mathbb{R}^2, [0,2\pi)$ and $[0,\infty)$, resp.)

Clearly, it was desirable to transfer the symmetry concept to more general settings to obtain similar results there. For this, we have to formalize the notion of symmetry (cf. Definitions 2.9 and 4.5). A topological group G *acts* on a topological space M if $(g,m) \in G \times M$ is mapped continuously onto an element $gm \in M$, if the equality $(g_1 g_2)m = g_1(g_2 m)$ holds for all $g_1, g_2 \in G$, and if the identity element $e_G \in G$ fulfils $e_G m = m$ for all $m \in M$. We say that a function $f\colon M \to \mathbb{R}$ is *G-invariant* if $f(gm) = f(m)$ for all $(g,m) \in G \times M$. A measure ν on a topological space M (or, to be more precise: on its Borel σ-algebra $\mathcal{B}(M)$) is called *G-invariant* if $\nu(gB) = \nu(B)$ for all $g \in G$ and $B \in \mathcal{B}(M)$.

Before we formulate the central problem of this monograph in its most general form we once again return to the introductory example where we use the terminology just introduced. In particular, this enables a straightforward generalization of this specific setting to arbitrary group actions. We first note that the $\mathrm{SO}(2)$ acts on $\mathbb{R}^2$ via $(\boldsymbol{T}(\alpha), \mathbf{x}) \mapsto \boldsymbol{T}(\alpha)\mathbf{x}$ where $\boldsymbol{T}(\alpha)$ denotes that (2×2)-matrix which induces a rotation by angle α. We first point out that the radially symmetric functions are in particular $\mathrm{SO}(2)$-invariant. Further, $(\boldsymbol{T}(\alpha), z) \mapsto (\alpha + z) \,(\mathrm{mod}\, 2\pi)$ defines an $\mathrm{SO}(2)$-action on the interval $[0,2\pi)$. This action is *transitive*, i.e. any $z_1 \in [0,2\pi)$ can be mapped onto a given $z_2 \in [0,2\pi)$ with a suitable matrix $\boldsymbol{T} \in \mathrm{SO}(2)$ (more precisely, $(\boldsymbol{T}(z_2 - z_1), z_1) \mapsto z_2$). Consequently, $(\boldsymbol{T}(\alpha), (z,r)) \mapsto \boldsymbol{T}(\alpha).(z,r) := (\alpha + z (\mathrm{mod}\, 2\pi), r)$ defines an $\mathrm{SO}(2)$-action on the product space $[0,2\pi) \times [0,\infty)$. The mapping $\varphi_p\colon \mathbb{R}^2 \to [0,\infty) \times [0,2\pi)$ is *equivariant* with respect to these $\mathrm{SO}(2)$ actions, i.e. φ_p commutes with the group actions. This assertion can be easily verified using

the addition theorems for sine and cosine: $\varphi_p\left(\boldsymbol{T}(\alpha).(z,r)\right)^t = \varphi_p\left(\alpha+z,r\right)^t = \left(r\cos(\alpha+z), r\sin(\alpha+z)\right)^t = \boldsymbol{T}(\alpha)\varphi_p(z,r)^t$. The equivariance property may be surprising at first sight. In fact, it will turn out to be crucial for the following.

A large number of applications of practical importance are of the following type:

(V) S, T and H denote topological spaces, $\varphi\colon S\times T\to H$ is a surjective mapping and G a compact group which acts on S and H. The G-action on the compact space S is transitive and φ is G-equivariant with respect to the G-actions $g(s,t)\mapsto(gs,t)$ and $(g,h)\mapsto gh$ on the spaces $S\times T$ and on H, resp., i.e. $g\varphi(s,t)=\varphi(gs,t)$ for all $(g,s,t)\in G\times S\times T$.

This monograph investigates the situation characterized by (V). To be precise, in Chapter 2 condition (V) will be augmented by some topological conditions which yield Property $(*)$ (cf. Section 2.2 or Section 4.1). These additional assumptions, however, are not particularly restrictive and met by nearly all applications of practical interest (see Chapter 4). Our interest lies in the G-invariant probability measures or, more generally, in the G-invariant Borel measures on H. In many cases the space S and the mapping φ are essentially ancillary tools. Unlike on H there exists exactly one G-invariant probability measure $\mu_{(\mathrm{S})}$ on S.

Identifying 0 with 2π the interval $[0,2\pi)$ is homeomorphic to the unit circle $\mathrm{S}^1\cong\mathbb{R}/2\pi\mathbb{Z}$ and hence compact. It can be easily checked that the introductory example fulfils (V) with $G=\mathrm{SO}(2)$, $S=[0,2\pi)$, $T=[0,\infty)$, $H=\mathbb{R}^2$ and $\varphi=\varphi_p$. Similarly, cylinder coordinates meet (V) with $G\cong\mathrm{SO}(2)$, $S=[0,2\pi)$, $T=[0,\infty)\times\mathbb{R}$, $H=\mathbb{R}^3$ and $\varphi(\alpha,r,z):=(r\cos\alpha, r\sin\alpha, z)$. For spherical coordinates in $\mathbb{R}^n$ we have $G=\mathrm{SO}(n)$ (or $G=\mathrm{O}(n)$) and $S=\mathrm{S}^{n-1}$ where the latter denotes the surface of the n-dimensional unit ball, and $\varphi\colon\mathrm{S}^{n-1}\times[0,\infty)\to\mathbb{R}^n$ is given by $\varphi(\mathbf{s},r):=r\mathbf{s}$. (Recall that using spherical coordinates means integrating with respect to radius and angles. Up to a constant the latter corresponds to an integration on S^{n-1} with respect to the Lebesgue surface measure (= equidistribution on S^{n-1})). Condition (V) is also met by the QR- and the polar decomposition of invertible real $(n\times n)$-matrices (cf. Sections 4.7 and 4.8). In both cases $G=S=\mathrm{O}(n)$ and $H=\mathrm{GL}(n)$. The orthogonal group $\mathrm{O}(n)$ acts on S and H by left multiplication while T is given by the group of all upper triangular $(n\times n)$-matrices with positive diagonal elements or the set of all symmetric positive definite $(n\times n)$-matrices, respectively. The $\mathrm{O}(n)$-invariant measures on $\mathrm{GL}(n)$ of most practical importance are the Lebesgue measure $\lambda_{\mathrm{GL}(n)}$ and the Haar measure $\mu_{\mathrm{GL}(n)}$.

The outstanding property of polar, cylinder and spherical coordinates, namely their invertibility outside a Lebesgue zero set is not demanded in (V). Resigning on such a bijectivity condition complicates the situation considerably. Pleasant properties get lost and proofs become more complicated.

On the positive side, it enlarges the field of applications considerably (cf. Chapter 4).

Measures with weaker symmetry properties than Haar measures on locally compact groups, invariant measures on homogeneous spaces or radially symmetric measures in $\mathbb{R}^n$ have been under-represented in research and literature. Most of this research work has been motivated by statistical applications (cf. Section 2.3 and in particular the extensive Remark 2.46). As already pointed out the usefulness of symmetry properties is not restricted to statistical applications but can also used for the evaluation of high-dimensional integrals and stochastic simulations. Anyway, interesting questions and problems arise which have to be solved.

If the 5-Tupel (G, S, T, H, φ) fulfils (V) under weak additional assumptions on G, S, T, H and φ the following statements are valid:

(A) The image measure $(\mu_{(S)} \otimes \tau)^{\varphi}$ is G-invariant for each Borel measure τ on T. Vice versa, to each G-invariant Borel measure ν on H there exists a (usually not unique) Borel measure τ_ν on T with $\left(\mu_{(S)} \otimes \tau_\nu\right)^{\varphi} = \nu$.

A *Borel measure* is a measure on a Borel σ-algebra which has finite mass on each compact subset. Clearly, probability measures and, more general, finite measures on a Borel σ-algebra are Borel measures. In particular, the class of Borel measures should contain all measures of practical relevance. Note that the first assertion of (A) can be extended to non-Borelian measures on $\mathcal{B}(T)$. Actually, the additional assumptions from $(*)$ (compared with (V)) are met by nearly all applications with practical relevance. For concrete examples these assumptions usually are easy to verify. Counterexamples show that these conditions may not be dropped. In the introductory example $\mu_{(S)} = \lambda_{[0,2\pi)}/2\pi$, and for $h(r) := 2\pi r$ the image measure $\left(\mu_{(S)} \otimes h \cdot \lambda_{[0,\infty)}\right)^{\varphi_p}$ equals the Lebesgue measure in $\mathbb{R}^2$. Note that (A) is not only valid for measures with radially symmetric Lebesgue densities but for all radially symmetric Borel measures on $\mathbb{R}^2$ (which in particular are SO(2)-invariant), e.g. also for those whose total mass is concentrated on countably many circles.

For more complex situations than the repeatedly mentioned polar, cylinder and spherical coordinates enormous efforts may be required to determine τ_ν explicitly, in particular if the transformation theorem cannot be applied. For this, the following remark will turn out to be very useful. If τ_ν is known a whole class of pre-images $\tau_{f\cdot\nu}$ can be determined without additional computations:

(B) Assume that ν is a G-invariant Borel measure on H and $f\colon H \to [0,\infty]$ a G-invariant ν-density. Then the measure $f \cdot \nu$ is also G-invariant. Let $f_T(t) := f\left(\varphi(s_0, t)\right)$ for any fixed $s_0 \in S$. Then $\left(\mu_{(S)} \otimes \tau_\nu\right)^{\varphi} = \nu$ implies $\left(\mu_{(S)} \otimes f_T \cdot \tau_\nu\right)^{\varphi} = f \cdot \nu$.

Both statements from (A) have enormous consequences which are briefly sketched in the remainder of this section. Further aspects concerning the

relevance of (A) and (B) are discussed and deeper considerations are made in Chapters 3 and 4.

a) Let ν be a G-invariant Borel measure H. Then (A) implies

$$\int_H f(h)\,\nu(\mathrm{d}h) = \int_T \int_S f\left(\varphi(s,t)\right)\,\mu_{(\mathrm{S})}(ds)\tau_\nu(\mathrm{d}t) \tag{1.3}$$

for all ν-integrable functions $f\colon H \to \mathbb{R}$. If the integrand f is also G-invariant the right-hand side simplifies to $\int_T f(\varphi(s_0,t))\,\tau_\nu(\mathrm{d}t)$ for any $s_0 \in S$. If a Borel measure ν on $\mathrm{GL}(n)$ is invariant under the left multiplication with orthogonal matrices then both the QR-decomposition and the polar decomposition of real-valued $(n \times n)$-matrices reduce the dimension of the domain of integration from n^2 to $n(n+1)/2$. If the integrand f merely depends on the absolute value of the determinant of its argument, for example, the QR-decomposition additionally simplifies the integrand considerably. On $\mathrm{GL}(n)$ the determinant function is given by a homogeneous polynomial in the matrix components which has a large number of summands. For upper triangular matrices the determinant function equals the product of the diagonal elements. Note that no QR-decomposition of any matrix has to be carried out explicitly. If, additionally, ν and f are also invariant under the right multiplication with orthogonal matrices (then ν and f are called 'biinvariant'), the integral $\int_{\mathrm{GL}(n)} f(\boldsymbol{M})\,\nu(\mathrm{d}\boldsymbol{M})$ simplifies to $\int_{\mathrm{D}_{+\geq}(n)} f(\boldsymbol{D})\,\tau_\nu(\mathrm{d}\boldsymbol{D})$ where $\mathrm{D}_{+\geq}(n)$ denotes the set of all diagonal $(n \times n)$-matrices with positive, monotonously decreasing diagonal elements. The dimension of the domain of integration even reduces to n in this case. Many Borel measures and functions on $\mathrm{GL}(n)$ of practical relevance are biinvariant, for example the Lebesgue measure, the Haar measure, normal distributions and functions which merely depend on the spectral norm, the Frobenius norm or the absolute value of the determinant of the argument or the inverse of the argument, resp. To be precise, a function $f\colon \mathrm{GL}(n) \to \mathbb{R}$ is biinvariant if it can be expressed as a function of the singular values of its argument. In fact, this surprising property will be used for various purposes. The repeatedly mentioned polar decomposition of real-valued matrices also plays an important role in natural and social sciences. Upper bounds for the relative error of the output matrix are known which depend on the machine precision ε and the singular values of the input matrix. Unfortunately, the singular values are not known until the polar decomposition of the respective input matrix has been computed (a posteriori bound). If the input matrices may be viewed as realizations of (i.e., as values assumed by) biinvariantly distributed random variables it is possible to determine the machine-precision *before* the first numerical calculation so that the relative error does not exceed a given (tolerable) bound with a probability of at least $1-\beta$. This saves time-consuming computations with needlessly high machine precision as well as frequent restarts with increased machine precision because the previously chosen machine precision has turned out to be insufficient after the calculations have been carried out. In principle, using

a symmetrization technique described in Chapter 2 this result could be extended to any (not necessarily biinvariant) distribution of the input matrices. However, depending on the concrete distribution this additional step may be very costly.

b) As in the introductory example assertion (A) can generally be used for an efficient generation of G-invariantly distributed pseudorandom elements $\widetilde{Z}_1, \widetilde{Z}_2, \ldots, \widetilde{Z}_N$ on H. More precisely, from $\mu_{(\mathrm{S})}$-distributed pseudorandom elements $\widetilde{X}_1, \widetilde{X}_2, \ldots, \widetilde{X}_N$ on S and τ_ν-distributed pseudorandom elements $\widetilde{Y}_1, \widetilde{Y}_2, \ldots, \widetilde{Y}_N$ on T one computes $\widetilde{Z}_1 := \varphi(\widetilde{X}_1, \widetilde{Y}_1), \widetilde{Z}_2 := \varphi(\widetilde{X}_2, \widetilde{Y}_2), \ldots, \widetilde{Z}_N := \varphi(\widetilde{X}_N, \widetilde{Y}_N)$. This approach has various advantages. First of all, the simulation problem on H is decomposed into two simulations problems which can be treated independently. In many applications H, S and T are manifolds or subsets of vector spaces where at least the dimension of T is considerably smaller than that of H. Due to its 'total' symmetry there usually exist efficient algorithms to simulate $\mu_{(\mathrm{S})}$. Moreover, this approach yields a partial unification of the simulation algorithms for a whole class of distributions. Additionally, it usually reduces the average number of standard random numbers needed for the generation of a pseudorandom element on H, and it often saves time-consuming computations of unwieldy chart mappings (see Chapter 3 for a more detailed discussion). An important example constitute the probability measures ν on $\mathrm{SO}(3)$ which are invariant under conjugation ('conjugation-invariant' probability measures), i.e. probability measures with $\nu\left(\boldsymbol{T}B\boldsymbol{T}^{-1}\right) = \nu(B)$ for all $\boldsymbol{T} \in \mathrm{SO}(3)$ and all measurable $B \subseteq \mathrm{SO}(3)$. In this case $S = \mathrm{S}^2$ while $T \subseteq \mathrm{SO}(3)$ equals the subgroup of all rotations around the z-axis which can be identified with the interval $[0, 2\pi)$ equipped with the addition modulo 2π. If the axis of the rotation and the angle may be interpreted as realizations of independent random variables, and if the standardized oriented random axis is equidistributed on S^2 the induced random rotation is conjugation-invariantly distributed, and vice versa, each conjugation-invariantly distributed random variable on $\mathrm{SO}(3)$ can be represented in this way. This property underlines the significance of conjugation-invariant measures, e.g. for applications in the field of computer graphics. Compared with 'ordinary' simulation algorithms the algorithms based on our symmetry concept require considerably less computing time. Depending on ν the saving may be more than 75 per cent.

In this book we mainly consider stochastic simulations on continuous spaces. However, under suitable conditions the symmetry calculus can also be fruitfully applied to simulation problems on large finite sets. To simulate a distribution ν on a finite set M which is not the equidistribution one usually needs a table which contains the probabilities $\nu(m)$ for all $m \in M$. If $|M|$ is large this table permanently requires a large area of memory. Its initialization and each table search take up computing time. If ν is invariant under the action of a finite group G, i.e. if ν is equidistributed on each G-orbit, we can apply the symmetry calculus with $S = G$ and $H = M$. The group G acts on

S by left multiplication, and $T \subseteq M$ is a set of representatives of the G-orbits on M. The mapping $\varphi\colon G \times T \to M$ is given by the group action on M, i.e. by $\varphi(g,t) = gt$, and $\tau_\nu(t)$ equals $\nu(G\text{-orbit of } t)$. Instead of simulating ν 'directly' one simulates the equidistribution on G and a distribution τ_ν on a (hopefully) small set T. For example, let $G = M = \mathrm{GL}(6;\mathrm{GF}(2))$, the group of invertible (6×6)-matrices over $\mathrm{GF}(2)$, on which G acts on M by conjugation and let ν be equidistributed on each orbit. Then $G = S = M$ and $|T| = 60$ while $|M| = 20\,158\,709\,760$. If ν is not the equidistribution a 'direct' simulation of ν on $\mathrm{GL}\,(6;\mathrm{GF}(2))$ is hardly feasible in practice. Exploiting the symmetry it essentially suffices to simulate the equidistribution on $\{0,1,\ldots,2^{36}-1\}$ and a distribution τ_ν on a set with 60 elements. In Section 4.11 we further present an application from coding theory.

Occasionally, an unknown distribution ν on a large finite set M or at least particular properties of ν have to be estimated. If it can be shown that the unknown distribution ν is invariant under a particular group action it obviously suffices to estimate the probabilities of the respective orbits instead of single elements. The empirical distribution of the orbits converges much faster to the exact distribution than the empirical distribution of the individual elements.

c) Under weak additional assumptions there exists a measurable mapping $\psi\colon H \to T$ for which $\varphi(s,\psi(h))$ lies in the G-orbit of h for each $(s,h) \in S \times H$. In particular $\left(\mu_{(\mathrm{S})} \otimes \nu^\psi\right)^\varphi = \nu$, i.e. the image measure ν^ψ is a concrete candidate for τ_ν. The existence of such a mapping ψ opens up applications in the field of statistics. The statistician typically observes a sample $h_1,\ldots,h_m \in H$ which he interprets as realizations of independent random variables $Z_1,\ldots,Z_m$ with unknown distributions $\nu_1,\ldots,\nu_m$. Depending on the random experiment it may be reasonable to assume that the distributions $\nu_1,\ldots,\nu_m$ are G-invariant. Then the mapping $\psi^m\,(h_1,\ldots,h_m) := (\psi(h_1),\ldots,\psi(h_m))$ is sufficient, i.e. the statistician does not lose any information in considering the images $\psi(h_1),\ldots,\psi(h_m) \in T$ instead of the observed values $h_1,\ldots,h_m \in H$ themselves. For the introductory example the mapping $\psi\colon \mathbb{R}^2 \to [0,\infty)$, $\psi(\mathbf{z}) := \|\mathbf{z}\|$ has the desired property. The conjugation-invariant probability measures on $\mathrm{SO}(n)$ are another important example. The compact group $\mathrm{SO}(n)$ is an $n(n-1)/2$-dimensional manifold. For $n \in \{2k, 2k+1\}$ the space $T \subseteq \mathrm{SO}(n)$ equals a commutative subgroup which can be identified with the k-dimensional unit cube $[0,1)^k$ (equipped with the componentwise addition modulo 1). The advantages are obvious. First, it suffices to store $N \cdot k$ real numbers instead of $N \cdot n^2$ many. Moreover, concrete analytical or numerical calculations (e.g. the determination of the rejection area) can usually be performed much easier on a cube than on a high-dimensional manifold. For $n = 3$ both the phase of a non-trivial eigenvalue and the trace function yield sufficient statistics. We mention that our results are also relevant for invariant test problems.

d) In Chapter 2 we present the main theorems of this book. As a 'by-product' we obtain various results from the field of measure extensions.

In Chapter 2 the theoretical foundations of the symmetry concept are laid and the main theorems are proved. Chapter 2 is the mathematical core of this book. Its results have an importance of their own and they provide the basis for the applications which are investigated in detail in Chapter 4. By the time the readers have reached Chapter 3 they have probably a pretty good intuition for the applications of the theory because we pointed these out in a cursory fashion before. Now we shall deepen this aspect of the applications of our theoretical results. In Chapter 4 a number of examples will be investigated which underline the usefulness of the symmetry concept in various domains in pure and applied mathematics, information theory and applied sciences. The examples we discuss are important quite independently, but in addition, they should put the readers in a position to apply the symmetry concept to problems of their own.

At the beginning of Chapter 4 the main results from Chapter 2 are collected; of course we do not repeat their proofs or the preparatory lemmas at this point. We hope that this summary or requisite material will put the reader in the position to understand Chapter 4 without going through Chapter 2 first. This should facilitate the understanding of the applications, especially for non-mathematicians.

2 Main Theorems

This chapter contains the main theorems of this book. The applications which we shall discuss in Chapter 4 below will be grounded upon the material provided in this chapter. The main results themselves and the mathematical tools and techniques which are used to derive and prove them mainly belong to the domain of measure and integration theory. Moreover, we shall not hesitate to employ techniques from topology and transformation groups. For the sake of a better understanding of significant definitions and facts we shall illustrate them by simple examples. Counterexamples underline the necessity of particular assumptions and hopefully improve the insight into the mathematical background.

2.1 Definitions and Preparatory Lemmata

Section 2.1 begins with a number of definitions and lemmata from topology and measure theory. Although some of these results are important of their own Section 2.1 essentially provides preparatory work for Sections 2.2 and 2.3. A number of elementary examples and counterexamples and explaining remarks shall facilitate the lead-in for readers who are not familiar with topology, measure theory and the calculus of group actions.

Definition 2.1. *The restriction of a mapping* $\chi\colon M_1 \to M_2$ *to* $E_1 \subseteq M_1$ *is denoted with* $\chi_{|E_1}$*, the pre-image of* $F \subseteq M_2$*, i.e. the set* $\{m \in M_1 \mid \chi(m) \in F\}$*, with* $\chi^{-1}(F)$*. If* χ *is invertible* χ^{-1} *also stands for the inverse of* χ*. The union of* disjoint *subsets* A_j *is denoted with* $A_1 + \cdots + A_k$ *or* $\sum_j A_j$*, resp.*

Definition 2.2. *Let* M *be a topological space. A collection of open subsets of* M *is called a* topological base *if each open subset* $O \subseteq M$ *can be represented as a union of elements of this collection. The space* M *is said to be* second countable *if there exists a countable base. A subset* $U \subseteq M$ *is called a* neighbourhood *of* $m \in M$ *if there is an open subset* $U' \subseteq M$ *with* $\{m\} \subseteq U' \subseteq U$*. We call a subset* $F \subseteq M$ quasi compact *if each open cover* $\bigcup_{\iota \in I} U_\iota \supseteq F$ *has a finite subcover. If* M *itself has this property then* M *is called a* quasi compact space*. A topological space* M *is called* Hausdorff space *if for each* $m \neq m' \in M$ *there exist disjoint open subsets* U *and* U' *with* $m \in U$ *and*

W. Schindler: LNM 1808, pp. 11–54, 2003.

$m' \in U'$. *A quasi compact subset of a Hausdorff space is called* compact. *Similarly, we call a quasi compact Hausdorff space* compact. *A subset* $F \subseteq M$ *is said to be* relatively compact *if its closure* $\overline{F}$ *(i.e. the smallest closed superset of* F*) is compact. A Hausdorff space* M *is called* locally compact *if each* $m \in M$ *has a compact neighbourhood. The topological space* M *is called* σ-compact *if it can be represented as a countable union of compact subsets. For* $F \subseteq M$ *the open sets in the* induced topology *(or synonymously:* relative topology*) on* F *are given by* $\{O \cap F \mid O \text{ is open in } M\}$. *In the* discrete topology *each subset* $F \subseteq M$ *is an open subset of* M.

Let M_1 *and* M_2 *be topological spaces. The* product topology *on* $M_1 \times M_2$ *is generated by* $\{O_1 \times O_2 \mid O_i \text{ is open in } M_i\}$. *A group* G *is said to be a* topological group *if* G *is a topological space and if the group operation* $(g_1 g_2) \mapsto g_1 g_2$ *and the inversion* $g \mapsto g^{-1}$ *are continuous where* $G \times G$ *is equipped with the product topology. We call the group* G compact *or* locally compact*, resp., if the underlying space* G *is compact or locally compact, resp.*

Readers who are completely unfamiliar with the foundations of topology are referred to introductory books (e.g. to [59, 23, 43]). Unless otherwise stated in the following all countable sets are equipped with the discrete topology while the $\mathbb{R}^n$ and the matrix groups are equipped with the usual Euclidean topology.

Remark 2.3. (i) Each compact space is locally compact.
(ii) The definition of local compactness is equivalent to the condition that each $m \in M$ has a neighbourhood base of compact subsets ([59], p. 88).
(iii) By demanding the Hausdorff property in the definition of compactness we follow the common convention. We point out that some topology works (e.g. [43]) do not demand the Hausdorff property, i.e. they use 'compact' in the sense of 'quasi compact'. Consequently, one has to be careful when combining lemmata and theorems with compactness assumptions from different topology articles or books.

Lemma 2.4. *(i) Let* M *be a topological space with a countable topological base and let* $F \subseteq M$. *Then* F *is second countable in the induced topology.*
(ii) Let M *be locally compact. If* $F \subseteq M$ *can be represented as an intersection of an open and a closed subset then* F *is a locally compact space in the induced topology.*
(iii) For a locally compact space are equivalent:

(a) M *is second countable.*
(b) M *is metrizable and* σ*-compact.*

(iv) Let M *be second countable and let* $\bigcup_{j \in J} U_j \supseteq F$ *be an open cover of* $F \subseteq M$. *Then there exists a countable subcover of* F.

Proof. Assertion (i) follows immediately from the definition of the induced topology. Assertions (ii) and (iii) are shown in [59], pp. 88 and 110. To prove (iv) we fix a countable topological base. Let $V_1, V_2, \ldots$ denote those elements of this base which are a subset of at least one U_j. There exists a sequence $U_1', U_2', \ldots \in \{U_j \mid j \in J\}$ with $V_k \subseteq U_k'$. Hence $\bigcup_{j\in J} U_j = \bigcup_{k\in N} V_k \subseteq \bigcup_{k\in N} U_k'$ which completes the proof of (iv). □

If M is locally compact a countable topological base implies its σ-compactness. In particular, it is an immediate consequence from (i) and (ii) that (iii)(a) (and hence also the equivalent condition (iii)(b)) are induced on subsets $F \subseteq M$ which can be expressed as an intersection of finitely many open and arbitrarily many closed subsets. Example 2.5 underlines that many topological spaces which are relevant for applications (cf. Chapter 4) are second countable and locally compact.

Example 2.5. The space $\mathbb{R}^n$ is locally compact. Let $U_\varepsilon(x) := \{y \in \mathbb{R}^n \mid \|x - y\| < \varepsilon\} \subseteq \mathbb{R}^n$ denote the open ε-ball around $x \in \mathbb{R}^n$. Then $\{U_{1/n}(x) \mid n \in \mathbb{N}; x \in \mathbb{Q}^n\}$ is a countable topological base of the Euclidean topology on $\mathbb{R}^n$. The vector space of all real $(n \times n)$-matrices is canonically isomorphic to $\mathbb{R}^{n^2}$ and hence also second countable and locally compact. Lemma 2.4 implies that the latter is also true for $\mathrm{GL}(n)$, $\mathrm{O}(n)$, $\mathrm{SO}(n)$ and $\mathrm{SL}(n) := \{\mathbf{M} \in \mathrm{GL}(n) \mid \det \mathbf{M} = 1\}$ as well as for open or closed subsets of these subgroups. Also $\mathrm{U}(n)$, the group of all unitary matrices, is compact and also second countable. Further examples of locally and σ-compact spaces are countable discrete spaces, i.e. countable sets which are equipped with the discrete topology.

Definition 2.6. *The* power set *of Ω is denoted with $\mathcal{P}(\Omega)$. A* σ-algebra *(or synonymously: a* σ-field*) $\mathcal{A}$ on Ω is a subset of $\mathcal{P}(\Omega)$ with the following properties: $\emptyset \in \mathcal{A}$, $A \in \mathcal{A}$ implies $A^c \in \mathcal{A}$, and $A_1, A_2, \ldots \in \mathcal{A}$ implies $\bigcup_{j\in\mathbb{N}} A_j \in \mathcal{A}$. The pair $(\Omega, \mathcal{A})$ is a* measurable space*. A subset $F \subseteq \Omega$ is* measurable *if $F \in \mathcal{A}$. Let $(\Omega_1, \mathcal{A}_1)$ and $(\Omega_2, \mathcal{A}_2)$ be measurable spaces. A mapping $\varphi\colon \Omega_1 \to \Omega_2$ is called* $(\mathcal{A}_1, \mathcal{A}_2)$-measurable *if $\varphi^{-1}(A_2) \in \mathcal{A}_1$ for each $A_2 \in \mathcal{A}_2$. We briefly call φ measurable if there is no ambiguity about the σ-algebras $\mathcal{A}_1$ and $\mathcal{A}_2$. The* product σ-algebra *of $\mathcal{A}_1$ and $\mathcal{A}_2$ is denoted with $\mathcal{A}_1 \otimes \mathcal{A}_2$. A function $\Omega \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$ is called* numerical*.*

A measure ν on $\mathcal{A}$ is non-negative *if $\nu(F) \in [0, \infty]$ for all $F \in \mathcal{A}$. If $\nu(\Omega) = 1$ then ν is a* probability measure*. The set of all probability measures on $\mathcal{A}$, resp. the set of all non-negative measures on $\mathcal{A}$, are denoted with $\mathcal{M}^1(\Omega, \mathcal{A})$, resp. with $\mathcal{M}^+(\Omega, \mathcal{A})$. Let $\tau_1, \tau_2 \in \mathcal{M}^+(\Omega, \mathcal{A})$. The measure τ_2 is said to be* absolutely continuous *with respect to τ_1, abbreviated with $\tau_2 \ll \tau_1$, if $\tau_1(N) = 0$ implies $\tau_2(N) = 0$. If τ_2 has the τ_1-density f we write $\tau_2 = f \cdot \tau_1$. Let $\varphi\colon \Omega_1 \to \Omega_2$ be a measurable mapping between two measurable spaces Ω_1 and Ω_2, and let $\eta \in \mathcal{M}^+(\Omega_1, \mathcal{A}_1)$. The term η^φ stands for the* image measure *of η under φ, i.e. $\eta^\varphi(A_2) := \eta(\varphi^{-1}(A_2))$ for each $A_2 \in \mathcal{A}_2$. A measure $\tau_1 \in \mathcal{M}^+(\Omega, \mathcal{A})$ is called* σ-finite *if Ω can be represented as the limit of a*

sequence of non-decreasing measurable subsets $(A_n)_{n\in\mathbb{N}}$ *with* $\tau_1(A_n) < \infty$ *for all* $n \in \mathbb{N}$. *The set of all* σ*-finite measures on* $\mathcal{A}$ *is denoted with* $\mathcal{M}^\sigma(\Omega, \mathcal{A})$. *The* product measure *of* $\eta_1 \in \mathcal{M}^\sigma(\Omega_1, \mathcal{A}_1)$ *and* $\eta_2 \in \mathcal{M}^\sigma(\Omega_2, \mathcal{A}_2)$ *is denoted with* $\eta_1 \otimes \eta_2$.

Let $\mathcal{A}_0 \subseteq \mathcal{A} \subseteq \mathcal{A}_1$ *denote* σ*-algebras over* Ω. *The* restriction *of* $\eta \in \mathcal{M}^+(\Omega, \mathcal{A})$ *to the sub-*σ*-algebra* $\mathcal{A}_0$ *is denoted by* $\eta_{|\mathcal{A}_0}$. *A measure* $\eta_1 \in \mathcal{A}_1$ *is called an* extension *of* η *if its restriction to* $\mathcal{A}$ *coincides with* η, *i.e. if* $\eta_{1|\mathcal{A}} = \eta$.

The Borel σ-algebra $\mathcal{B}(M)$ *over a topological space* M *is generated by the open subsets of* M. *If it is unambiguous we briefly write* $\mathcal{M}^1(M)$, $\mathcal{M}^\sigma(M)$ *or* $\mathcal{M}^+(M)$, *resp., instead of* $\mathcal{M}^1(M, \mathcal{B}(M))$, $\mathcal{M}^\sigma(M, \mathcal{B}(M))$ *and* $\mathcal{M}^+(M, \mathcal{B}(M))$, *resp., and we call* $\nu \in \mathcal{M}^+(M)$ *a measure on* M. *If* M *is locally compact then* $\nu \in \mathcal{M}^+(M)$ *is said to be a* Borel measure *if* $\nu(K) < \infty$ *for each compact subset* $K \subseteq M$. *The set of all Borel measures is denoted with* $\mathcal{M}(M)$.

Readers who are completely unfamiliar with measure theory are referred to relevant works ([5, 6, 28, 33] etc.) For all applications considered in Chapter 4 the respective spaces will be locally compact. Probability measures and, more generally, Borel measures are of particular interest. However, occasionally it will turn out to be reasonable to consider $\mathcal{M}^+(M)$ or $\mathcal{M}^\sigma(M)$ as this saves complicated pre-verifications of well-definedness and the distinction of cases. The most familiar Borel measure on $\mathbb{R}$ surely is the Lebesgue measure. Note that $B \in \mathcal{B}(\overline{\mathbb{R}})$ iff $(B \cap \mathbb{R}) \in \mathcal{B}(\mathbb{R})$.

Example 2.7. If M is a discrete space then $\mathcal{B}(M) = \mathcal{P}(M)$.

Lemma 2.8. *(i) Let* $\tau_1 \in \mathcal{M}^\sigma(M, \mathcal{A})$. *Then the following conditions are equivalent:*

(α*)* $\tau_2 \ll \tau_1$.
(β*)* τ_2 *has a* τ_1*-density* f, *i.e.* $\tau_2 = f \cdot \tau_1$ *for a measurable* $f\colon M \to [0, \infty]$.

(ii) Let M *be a topological space and* $A \in \mathcal{B}(M)$. *If* A *is equipped with the induced topology then* $\mathcal{B}(A) = \mathcal{B}(M) \cap A$.
(iii) Let M_1 *and* M_2 *be topological spaces with countable topological bases. Then* $\mathcal{B}(M_1 \times M_2) = \mathcal{B}(M_1) \otimes \mathcal{B}(M_2)$.
(iv) If M *is a second countable locally compact space* $\mathcal{M}(M) \subseteq \mathcal{M}^\sigma(M)$.

Proof. For proofs of (i) to (iii) see [28], pp. 41 (Satz 1.7.12), 17 and 24f. (Satz 1.3.12). If M is second countable 2.4(iii) guarantees the existence of a sequence of non-decreasing compact subsets $K_1, K_2, \ldots \subseteq M$ with $M = (K_n)_{n\in\mathbb{N}}$. For any $\eta \in \mathcal{M}(M)$ it is $\eta(K_j) < \infty$ and hence $\eta \in \mathcal{M}^\sigma(M)$. □

Definition 2.9. *Let* G *denote a compact group,* e_G *its identity element and* M *a topological space. A continuous mapping* $\Theta\colon G \times M \to M$ *is said to be a* group action, *or more precisely, a* G-action *if* $\Theta(e_G, m) = m$ *and*

$\Theta(g_2, \Theta(g_1, m)) = \Theta(g_2 g_1, m)$ *for all* $(g_1, g_2, m) \in G \times G \times M$. *We say that* G acts *or, synonymously,* operates *on* M *and call* M *a* G-space. *If it is unambiguous we briefly write* gm *instead of* $\Theta(g, m)$. *The* G*-action is said to be* trivial *if* $gm = m$ *for all* $(g, m) \in G \times M$. *For* $m \in M$ *we call* $G_m := \{g \in G \mid gm = m\}$ *the* isotropy group. *The union* $Gm := \bigcup_{g \in G} \{gm\}$ *is called the* orbit *of* m. *The* G*-action is said to be* transitive *if for any* $m_1, m_2 \in M$ *there exists a* $g \in G$ *with* $gm_1 = m_2$. *If* G *acts transitively on* M *and if the mapping* $g \mapsto gm$ *is open for all* $m \in M$ *then* M *is called a* homogeneous space. *A mapping* $\pi\colon M_1 \to M_2$ *between two* G*-spaces is said to be* G-equivariant *(or short:* equivariant*) if it commutes with the* G*-actions on* M_1 *and* M_2, *i.e. if* $\pi(gm) = g\pi(m)$ *for all* $(g, m) \in G \times M_1$. *A non-negative measure* ν *on* M *is called* G-invariant *if* $\nu(gB) := \nu(\{gx \mid x \in B\}) = \nu(B)$ *for all* $(g, B) \in G \times \mathcal{B}(M)$. *The set of all* G*-invariant probability measures (resp.* G*-invariant Borel measures, resp.* G*-invariant* σ*-finite measures, resp.* G*-invariant non-negative measures) on* M *are denoted with* $\mathcal{M}_G^1(M)$ *(resp.* $\mathcal{M}_G(M)$, *resp.* $\mathcal{M}_G^\sigma(M)$, *resp.* $\mathcal{M}_G^+(M)$*).*

Example 2.10. (Group action) $G = \mathrm{SO}(3)$, $M = \mathbb{R}^3$ and $\Theta\colon \mathrm{SO}(3) \times \mathbb{R}^3 \to \mathbb{R}^3$, $\Theta(\mathbf{T}, \mathbf{x}) := \mathbf{T}\mathbf{x}$. □

Lemma 2.11. *Suppose that a compact group* G *acts on a topological space* M. *Then the following statements are valid:*
(i) $(B \in \mathcal{B}(M)) \Rightarrow (gB \in \mathcal{B}(M)$ *for all* $g \in G)$.
(ii) The intersection, the union, the complements and the differences of G*-invariant subsets of* M *are themselves* G*-invariant.*
(iii) $\mathcal{B}_G(M) := \{B \in \mathcal{B}(M) \mid gB = B$ *for all* $g \in G\}$ *is a sub-*σ*-algebra of* $\mathcal{B}(M)$.
(iv) Let M *be a locally compact space which is* σ*-compact or second countable. Then* M *can be represented as a countable union of disjoint* G*-invariant relatively compact subsets* $A_1, A_2, \ldots \in \mathcal{B}_G(M)$, *i.e.* $M = \sum_{j=1}^{\infty} A_j$.
(v) Assume that M *is a second countable locally compact space. Then the group action* $\Theta\colon G \times M \to M$ *is* $(\mathcal{B}(G) \otimes B(M), \mathcal{B}(M))$*-measurable.*

Proof. Within this proof let $\mathcal{E} := \{E \in \mathcal{B}(M) \mid gE \in \mathcal{B}(M)$ for all $g \in G\}$. In a straight-forward manner one first verifies that $\mathcal{E}$ is a sub-σ-algebra of $\mathcal{B}(M)$. The mapping $\Theta_g\colon M \to M$, $\Theta_g(m) := gm$ is a homeomorphism for each $g \in G$. In particular, it maps open subsets of M onto open subsets. Hence $\mathcal{E}$ contains the generators of $\mathcal{B}(M)$. Thus $\mathcal{E} = \mathcal{B}(M)$ which proves assertion (i). Let the subset $A \subseteq M$ be G-invariant and $m \in M$. From $gm \in A$ we obtain $m = g^{-1}(gm) \in A$. Consequently, $m \in A^c$ implies $gm \in A^c$ which shows that the complement of A is also G-invariant. The remaining assertions of (ii) can be verified analogously. As $\emptyset \in \mathcal{B}_G(M)$ assertion (iii) is an immediate consequence of (ii). Lemma 2.4(iii) assures that the locally compact space M is σ-compact under both assumptions. That is, there exist countably many compact subsets K_j of M with $M = \bigcup_{j=1}^{\infty} K_j$. The set $K_j' := GK_j = \Theta(G, K_j)$ is a continuous image of a compact set and hence itself

compact. Defining $A_1 := K_1'$ and inductively $A_{n+1} := \bigcup_{j=1}^{n+1} K_j' \setminus \bigcup_{j=1}^{n} K_j'$ we obtain disjoint G-invariant relatively compact subsets $A_1, A_2, \ldots$ with $M = \sum_{j=1}^{\infty} A_j$. In particular, $A_1, A_2, \ldots$ are contained in $\mathcal{B}_G(M)$ which completes the proof of (iv). To prove (v) we equip the semigroup $C(M,M) := \{F\colon M \to M \mid F \text{ is continuous}\}$ with the compact-open topology which is generated by the subsets $W(K,U) := \{F' \in C(M,M) \mid F'(K) \subseteq U\}$ where K and U range through the compact and the open subsets of M, resp. (cf. [59], p. 166). We claim that $C(M,M)$ is second countable. For a proof of this claim it suffices to exhibit a countable subbasis $\mathcal{S}$ of $C(M,M)$, i.e. a countable collection of subsets of $C(M,M)$ such that the finite intersections of elements of $\mathcal{S}$ are a topological base of $C(M,M)$. Assume for the moment that $\mathcal{U}$ is a second countable base of M. Applying Satz 8.22(b) in [59], p. 89, we may assume that all $U \in \mathcal{U}$ are relatively compact. We claim that the countable family of all $W(\overline{U}, V)$ with U and V ranging through $\mathcal{U}$ is a subbasis. Let $f \in W(K,U)$ where K and U denote a compact and an open subset of M, resp. We claim that there are $U_1, \ldots, U_m$ and $V_1, \ldots, V_m$ in $\mathcal{U}$ such that $f \in \bigcap_{j=1}^{m} W(\overline{U_j}, V_j) \subseteq W(K,U)$. For each $x \in K$ there is a $V_x \in \mathcal{U}$ such that $f(x) \in V_x \subseteq U$. As f is continuous and M is locally compact there is a $U_x \in \mathcal{U}$ with $x \in U_x$ and $f(\overline{U_x}) \subseteq V_x$. As K is compact there are elements $x_1, \ldots x_n$ such that $K \subseteq \bigcup_{j=1}^{m} U_{x_j}$. In particular, $f(\overline{U_{x_j}}) \subseteq V_{x_j}$ for all $j \leq m$, i.e. f is contained in the intersection of the $W(\overline{U}_j, V_j)$ with $U_j := U_{x_j}$ and $V_j := V_{x_j}$. It remains to show that any $g \in \bigcap_{j=1}^{m} W(U_j, V_j)$ is also contained in $W(K,U)$, i.e. that $g(K) \subseteq U$. Let $x \in K$. Then $x \in U_j$ for some j and thus $g(x) \in g(\overline{U_j}) \subseteq V_j \subseteq U$. Hence $g(K) \subseteq U$ which finishes the proof of the claim that the compact-open topology on $C(M,M)$ is second countable. Let $\psi\colon G \to C(M,M)$ be defined by $\psi(g)(m) := \Theta(g,m)$. Clearly, ψ is an algebraic group homomorphism onto its image. Applying Satz 14.17 from [59], p. 167, with $X = G$ and $Y = Z = M$ we see that $\psi\colon G \to C(M,M)$ is continuous. Since M is locally compact the composition $(f,g) \mapsto f \circ g\colon C(M,M) \times C(M,M) \to C(M,M)$ is continuous ([12], par. 3, no. 4, Prop. 9), and the image $\psi(G)$ is a second countable compact semigroup (2.4(i)). Moreover, $\psi(G)$ is a compact group ([12], par. 3, no. 5, Corollaire) which is isomorphic to G/N where N denotes the kernel of ψ. The factor group G/N acts on M by $\Theta_N(gN,m) := \Theta(g,m)$, and by 2.8(iii) the G/N-action Θ_N is $(\mathcal{B}(G/N) \otimes \mathcal{B}(M), \mathcal{B}(M))$-measurable. The projection $\mathrm{pr}_N : G \to G/N$, $\mathrm{pr}_N(g) := gN$ is continuous and in particular $(\mathcal{B}(G), \mathcal{B}(G/N))$-measurable. Consequently, $\mathrm{pr}_N \times \mathrm{id}\colon G \times M \to G/N \times M$ is $(\mathcal{B}(G) \otimes \mathcal{B}(M), \mathcal{B}(G/N) \otimes \mathcal{B}(M))$-measurable. Altogether, $\Theta = \Theta_N \circ (\mathrm{pr}_N \times \mathrm{id})$ has been shown to be $(\mathcal{B}(G) \otimes \mathcal{B}(M), \mathcal{B}(M))$-measurable which completes the proof of (v). □

Lemma 2.11 collects useful properties which will be needed later. Assertion 2.11(v) is elementary if the group G is assumed to be second countable as it will be the case in the main theorems (cf. Section 2.2) and all applications considered in Chapter 4 since 2.8(iii) then implies $\mathcal{B}(G \times M) = \mathcal{B}(G) \otimes \mathcal{B}(M)$. The central idea of the proof of 2.11(v) was suggested to me by K.H. Hof-

mann. Without 2.11(v) we additionally had to demand in 2.18 that G is second countable. Although this generalization will not be relevant later it surely is interesting of its own.

Remark 2.12. For each compact group G there exists a unique probability measure $\mu_G \in \mathcal{M}^1(G)$ with $\mu_G(gB) = \mu_G(B)$ for all $(g, B) \in G \times \mathcal{B}(G)$. Further, also $\mu_G(Bg) = \mu_G(B)$ for each $g \in G$, i.e. μ_G is invariant under both, the left and the right multiplication of the argument with group elements. The probability measure μ_G is said to be *left-invariant* and *right-invariant*. It is called *Haar measure*. Because of its maximal symmetry μ_G is of outstanding importance. Moreover, $\mu_G(B) = \mu_G(B^{-1})$ for all $B \in \mathcal{B}(G)$ where $B^{-1} := \{m^{-1} \mid m \in B\}$. (In fact, as the inversion on G is a homeomorphism $\mu'(B) := \mu_G(B^{-1})$ defines a probability measure on G with $\mu'(gB) = \mu_G(B^{-1}g^{-1}) = \mu_G(B^{-1}) = \mu'(B)$ for all $g \in G$ and $B \in \mathcal{B}(G)$, i.e. μ' is a left-invariant probability measure and hence $\mu' = \mu_G$; see also [63], p. 123 (Theorem 5.14) for reference.) Note that left- and right-invariant Borel measures do also exist on locally compact spaces which are not compact. For a comprehensive treatment of Haar measures we refer the interested reader to [54] (cf. also Remark 4.15).

Example 2.13. (i) If G acts trivially on M then $\mathcal{B}_G(M) = \mathcal{B}(M)$, and each measure on M is G-invariant. In contrast, if G acts transitively then $\mathcal{B}_G(M) = \{\emptyset, M\}$.
(ii) Compact groups of great practical relevance are $\mathrm{O}(n)$, $\mathrm{SO}(n)$, $\mathrm{U}(n)$ and particular finite groups.
(iii) The factor group $\mathbb{R}/\mathbb{Z}$ is compact. Clearly, $(x+\mathbb{Z}) = (y+\mathbb{Z})$ iff $x-y \in Z$. This observation yields another common representation of $\mathbb{R}/\mathbb{Z}$ which is more suitable for concrete calculations, namely $([0,1), \oplus)$. In this context the symbol '$\oplus$' denotes the group operation 'addition modulo 1', i.e. $x \oplus y := x+y \pmod 1$. (That is $x \oplus y = x + y$ if $x + y < 1$ and $x + y - 1$ else.) Note that unlike the 'ordinary' interval $[0,1)$ the group $([0,1), \oplus)$ is compact as $\lim_{x \nearrow 1} x = 0$, i.e. as one identifies 0 and 1. To avoid confusion with the interval $[0,1)$ (where 0 and 1 are not idetified) we append the symbol '$\oplus$'. The Borel σ-algebra $\mathcal{B}([0,1), \oplus)$ and the Haar measure on $([0,1), \oplus)$, resp., are given by $\mathcal{B}([0,1))$ and the Lebesgue measure on the interval $[0,1)$, resp. Equipped with the complex multiplication the group S^1, the unit circle in the complex plane, is also isomorphic to the factor group $\mathbb{R}/\mathbb{Z}$. The mappings $x \mapsto \exp^{2\pi \mathrm{i} x}$, resp. $r \mapsto \exp^{2\pi \mathrm{i} r}$ define group isomorphisms between $([0,1), \oplus)$ and S^1, resp. between $\mathbb{R}/\mathbb{Z}$ and S^1.

Lemma 2.14. *Let G be a compact group which acts on the topological spaces M, M_1 and M_2. Assume further that the mapping $\pi\colon M_1 \to M_2$ is G-equivariant and measurable. Then the following statements are valid:*
(i) For each $\nu \in \mathcal{M}_G^+(M_1)$ we have $\nu^\pi \in \mathcal{M}_G^+(M_2)$. If $\nu \in \mathcal{M}_G^1(M_1)$ then $\nu^\pi \in \mathcal{M}_G^1(M_2)$.

(ii) Let M_1 and M_2 be locally compact. Then $\nu \in \mathcal{M}_G(M_1)$ implies $\nu^\pi \in \mathcal{M}_G(M_2)$ if one of the following conditions hold:

- *(α) The pre-image $\pi^{-1}(K)$ is relatively compact for each compact $K \subseteq M_2$.*
- *(β) The pre-image $\pi^{-1}(K)$ is relatively compact for each compact G-invariant $K \subseteq M_2$.*
- *(γ) Each $m \in M_2$ has a compact neighbourhood K_m with relatively compact pre-image $\pi^{-1}(K_m)$.*
- *(δ) Each $m \in M_2$ has a compact G-invariant neighbourhood K_m with relatively compact pre-image $\pi^{-1}(K_m)$.*

(iii) $\mathcal{M}_G^1(M) \neq \emptyset$. If M is finite then every G-invariant measure ν on M is equidistributed on each orbit, i.e. ν has equal mass on any two points lying on the same orbit.
(iv) Let $\nu \in \mathcal{M}_G^+(M)$ and $f \geq 0$ a measurable G-invariant numerical function. Then $\eta := f \cdot \nu \in \mathcal{M}_G^+(M)$. If M is locally compact, $\nu \in \mathcal{M}_G(M)$ and f locally ν-integrable (i.e. if for each $m \in M$ there exists a neighbourhood $U_m \subseteq M$ with $\int_{U_m} f(m)\,\nu(\mathrm{d}m) < \infty$) then $\eta \in \mathcal{M}_G(M)$.
(v) If G acts transitively on a Hausdorff space M then $|\mathcal{M}_G^1(M)| = 1$. In particular, M is a compact homogeneous G-space.

Proof. As π is equivariant we obtain the equivalences

$$\begin{aligned}\big(x \in \pi^{-1}(gB)\big) &\iff (\pi(x) \in gB) \iff \big(\pi(g^{-1}x) = g^{-1}\pi(x) \in B\big)\\ &\iff \big(g^{-1}x \in \pi^{-1}(B)\big) \iff \big(x \in g\pi^{-1}(B)\big)\,.\end{aligned}$$

for all $B \in \mathcal{B}(M_2)$ and $(g,x) \in G \times M_1$. This implies $\nu^\pi(gB) = \nu(\pi^{-1}(gB)) = \nu(g\pi^{-1}(B)) = \nu(\pi^{-1}(B)) = \nu^\pi(B)$ which proves (i). Let $K \subseteq M_2$ be a compact subset of M_2. Then condition (ii) (α) implies $\nu^\pi(K) = \nu(\pi^{-1}(K)) < \infty$. As K was arbitrary ν^π is a Borel measure. Since K is contained in the compact G-invariant subset $GK = \Theta(G,K)$ condition (β) implies (α) as a closed subset of a compact set is itself compact. As K is compact $K \subseteq \bigcup_{j=1}^N K_{m_j}$ for suitable $m_1, m_2, \ldots, m_N \in K$. Condition ($\gamma$) implies $\nu(\pi^{-1}(K)) \leq \nu\big(\pi^{-1}(\bigcup_{j=1}^N K_{m_j})\big) \leq \sum_{j=1}^N \nu(\pi^{-1}(K_{m_j})) < \infty$. As condition ($\delta$) implies ($\gamma$) this finishes the proof of assertion (ii).
For each fixed $m \in M$ the mapping $\Theta_m\colon G \to M$, $\Theta_m(g) := gm$ is continuous. The group G acts on itself by left multiplication. In particular, $\mu_G \in \mathcal{M}_G^1(G)$ and $g_2\Theta_m(g_1) = g_2(g_1 m) = (g_2 g_1)m = \Theta_m(g_2 g_1)$, i.e. $\Theta_m\colon G \to M$ is equivariant. The first assertion of (iii) hence follows immediately from (i) since $\mu_G^{\Theta_m} \in \mathcal{M}_G^1(M)$. The second assertion of (iii) is obvious. The G-invariance of ν implies $\eta(gB) = \int_{gB} f(m)\,\nu(\mathrm{d}m) = \int_{gB} f(m)\,\nu^{(m \mapsto gm)}(\mathrm{d}m) = \int_B f(gm)\,\nu(\mathrm{d}m) = \int_B f(m)\,\nu(\mathrm{d}m) = \eta(B)$, i.e. $\eta \in \mathcal{M}_G^+(M)$. Let f be locally ν-integrable. As in the proof of (ii) the finite-cover-property of compact sets implies $\eta(K) < \infty$ for all compact $K \subseteq M$. Hence $\eta \in \mathcal{M}_G(M)$ which proves (iv). Let G act transitively on the Hausdorff space M. Then $M = \Theta_m(G)$ is the continuous image of a compact set and hence itself compact. In particular,

$M = \Theta_m(G)$ is a Baire space ([59], p. 152). Trivially, the compact group G is locally compact and σ-compact. Consequently, M is a homogeneous G-space ([11], p. 97 (chap. 7, par. 7 App. I, Lemme 2). This completes the proof of (v). □

Note that 2.14(i) and (v) may be very useful for stochastic simulations, for instance. A G-invariant probability measure $\eta \in \mathcal{M}^1_G(M_2)$ does not need to be simulated 'directly' on M_2. Instead, we may simulate any $\nu \in \mathcal{M}^1_G(M_1)$ with $\nu^\pi = \eta$. Applying the equivariant mapping π to the generated pseudorandom elements on M_1 one finally obtains the wanted η-distributed pseudorandom elements. Of particular importance, of course, is a situation as considered in 2.14(v). As $\nu^\pi = \eta$ for all $\nu \in \mathcal{M}^1_G(M_1)$ one may choose any of these distributions without further computations. Example 2.16 will illustrate the central idea.

Definition 2.15. *The terms λ, λ_n and λ_C denote the* Lebesgue measure *on $\mathbb{R}$, the n-dimensional Lebesgue measure on $\mathbb{R}^n$ and the restriction of λ_n to a Borel subset $C \subseteq \mathbb{R}^n$, i.e. $\lambda_C(B) = \lambda_n(B)$ for $B \in \mathcal{B}(C) = \mathcal{B}(\mathbb{R}^n) \cap C$. The term $N(\mu, \sigma^2)$ stands for the* normal *or* Gauss distribution. *The $(n-1)$-sphere S^{n-1} consists of all $\mathbf{x} \in \mathbb{R}^n$ with Euclidean norm 1. The normed geometric surface measure on S^2 is denoted with $\mu_{(\mathrm{S}^2)}$. The* Dirac measure *$\varepsilon_m \in \mathcal{M}^1(M)$ has its total mass concentrated on $m \in M$, i.e. $\varepsilon_m(B) = 1$ iff $m \in B$ and $\varepsilon_m(B) = 0$ else.*

Assume that $(\Omega, \mathcal{A})$ and $(\Omega', \mathcal{A}')$ are measurable spaces and that P is a probability measure on $\mathcal{A}$. The triple $(\Omega, \mathcal{A}, P)$ is called a measure space*, and a measurable mapping $X\colon \Omega \to \Omega'$ is a* random variable*. The image measure X^P is also called the* distribution *of X.*

Example 2.16. In this example we use Lemma 2.14 to derive an algorithm for the generation of equidistributed pseudorandom vectors on the surface of the unit ball in $\mathbb{R}^3$, that is, to simulate the normed geometric surface measure $\mu_{(\mathrm{S}^2)}$. First, the special orthogonal group $\mathrm{SO}(3)$ acts transitively on S^2 via $(\mathbf{T}, \mathbf{x}) \mapsto \mathbf{Tx}$. Further, $\mu_{(\mathrm{S}^2)} \in \mathcal{M}^1_{\mathrm{SO}(3)}(\mathrm{S}^2)$. As S^2 is a Hausdorff space 2.14(v) implies that $\mu_{(\mathrm{S}^2)}$ is the only $\mathrm{SO}(3)$-invariant probability measure on S^2. The $\mathrm{SO}(3)$ also acts on $\mathbb{R}^3 \setminus \{0\}$ by left multiplication, though not transitive. The non-transitivity on M_1 is definitely desirable since one can choose between many G-invariant measures. Clearly, one chooses such a $\nu \in \mathcal{M}^1_{\mathrm{SO}(3)}(\mathbb{R}^3 \setminus \{0\})$ which is suitable for simulation purposes. The restricted normal distribution $\nu := N(0,1) \otimes N(0,1) \otimes N(0,1)_{|\mathbb{R}^3 \setminus \{0\}}$ has the three-dimensional Lebesgue density $f(\mathbf{x}) := \mathrm{e}^{-\mathbf{x}\cdot\mathbf{x}/2}/\sqrt{2\pi}^3$ where '$\cdot$' denotes the scalar product of vectors. The transformation theorem in $\mathbb{R}^3$ yields

$$\frac{1}{\sqrt{2\pi}^3}\int_B \mathrm{e}^{-\mathbf{x}\cdot\mathbf{x}/2}\,\mathrm{d}\mathbf{x} = \frac{1}{\sqrt{2\pi}^3}\int_B |\det \mathbf{T}| \mathrm{e}^{-(\mathbf{Tx})\cdot(\mathbf{Tx})/2}\,\mathrm{d}\mathbf{x} = \frac{1}{\sqrt{2\pi}^3}\int_{\mathbf{T}B} \mathrm{e}^{-\mathbf{y}\cdot\mathbf{y}/2}\,\mathrm{d}\mathbf{y}$$

for all $(\boldsymbol{T}, B) \in \mathrm{SO}(3) \times \mathcal{B}(\mathbb{R}^3 \setminus \{0\})$. This shows that ν is $\mathrm{SO}(3)$-invariant. Since $\pi\colon \mathbb{R}^3 \setminus \{0\} \to \mathrm{S}^2$, $\pi(x) := x/\|x\|$ is $\mathrm{SO}(3)$-equivariant 2.14(v) finally

implies $\nu^\pi = \mu_{(\mathrm{S}^2)}$. This yields the following pseudoalgorithm. The pseudorandom vector $(\widetilde{X}_1, \widetilde{X}_2, \widetilde{X}_3) \in \mathbb{R}^3 \setminus \{0\}$ generated in Step 1 may be viewed as ν-distributed.

Algorithm (Equidistribution on S^2)

1. Generate independent $N(0,1)$-distributed pseudorandom numbers $\widetilde{X}_1$, $\widetilde{X}_2$, $\widetilde{X}_3$ until $\widetilde{N} := \widetilde{X}_1^2 + \widetilde{X}_2^2 + \widetilde{X}_3^2 > 0$.
2. Return $\widetilde{Z} := \frac{(\widetilde{X}_1, \widetilde{X}_2, \widetilde{X}_3)}{\sqrt{\widetilde{N}}}$.
3. END.

Of course, this simulation algorithm is not new at all. However, the calculus of group action, equivariant mappings and invariant measures illustrate its mode of functioning and verifies its correctness without exhaustive computations. A proof which uses the symmetry of the sphere and specific properties of normal distributions is given in [18], p. 230. We point out that due to the outstanding symmetry of the sphere there exist many other efficient algorithms for the simulation of $\mu_{(\mathrm{S}^2)}$ (cf. [18], p. 230 ff., for example).

Section 4.2 considers the equidistribution on the Grassmannian manifold. Again, Lemma 2.14 will be applied to deduce efficient simulation algorithms. These algorithms were used to confirm a conjecture from the field of combinatorial geometry (cf. Section 4.2). As the dimension of the Grassmannian manifold of particular interest is high the straight-forward approach, namely using a chart mapping, would have hardly been practically feasible.

The transformation of G-invariant measures with equivariant mappings constitutes the central idea of our concept. We will face it multiply, though in more sophisticated ways. Lemma 2.14(ii) provides a collection of sufficient conditions that Borel measures are transformed into Borel measures. The following counterexamples show that these conditions may not be dropped, even if $\pi\colon M_1 \to M_2$ is continuous and injective (cf. Counterexample 2.17(ii)).

Counterexample 2.17. (i) Let $G = \mathbb{Z}_2 := \{1, -1\}$, the multiplicative subgroup of $\mathbb{R}^* := \mathbb{R} \setminus \{0\}$ which merely consists of two elements. Let $\mathbb{Z}_2$ act on $M_1 := \mathbb{R}^*$ and $M_2 := \mathbb{Z}_2$ by $\Theta_1(z, r) := zr$, resp. by $\Theta_2(z, z_1) := zz_1$. The mapping $\pi\colon \mathbb{R}^* \to \mathbb{Z}_2$ is given by the signum function, i.e. $\pi(r) := \mathrm{sgn}(r)$, which is equivariant with respect to these $\mathbb{Z}_2$-actions. Clearly, $\mathcal{M}_{Z_2}(\mathbb{R}^*) = \{\nu \in \mathcal{M}(\mathbb{R}^*) \mid \nu(B) = \nu(-B) \text{ for all } B \in \mathcal{B}(\mathbb{R}^*)\}$ and $\mathcal{M}_{Z_2}(\mathbb{Z}_2) = \{\eta \in \mathcal{M}(\mathbb{Z}_2) \mid \eta(\{1\}) = \eta(\{-1\}) \in [0, \infty)\}$. In particular, the Lebesgue measure $\lambda \in \mathcal{M}_{Z_2}(\mathbb{R}^*)$ and $\lambda^\pi(\{1\}) = \lambda((0,\infty)) = \infty = \lambda((-\infty, 0)) = \lambda^\pi(\{-1\})$, i.e. $\lambda^\pi \in \mathcal{M}^+_{Z_2}(\mathbb{Z}_2)$ but $\lambda^\pi \notin \mathcal{M}_{Z_2}(\mathbb{Z}_2)$.
(ii) Let $M_1 = \mathbb{Q}$, $M_2 = [-2, 2] \subseteq \mathbb{R}$ and $\pi(q) := \arctan q$. If we equip $\mathbb{Q}$ with the discrete topology and the interval $[-2, 2]$ with the common (Euclidean) topology, then the spaces M_1 and M_2 are second countable and locally compact, and π is continuous and injective. Suppose that a group G

acts trivially on M_1 and M_2. As a first consequence, the group action need not be considered in the following, and $\pi\colon M_1 \to M_2$ is G-equivariant. Further, $\mathcal{M}_G(\mathbb{Q}) = \mathcal{M}(\mathbb{Q})$ and $\mathcal{M}_G([-2,2]) = \mathcal{M}([-2,2])$. Let $\nu \in \mathcal{M}(\mathbb{Q})$ be defined by $\nu(q) := 1$ for each $q \in \mathbb{Q}$. Then $\nu^\pi([-\varepsilon,\varepsilon]) = \nu(\tan([-\varepsilon,\varepsilon]) \cap \mathbb{Q}) = \infty$ for each $\varepsilon > 0$ as $\mathbb{Q}$ is dense in $\mathbb{R}$. Hence ν^π is not a Borel measure on M_2.

Theorem 2.18. *Suppose that the compact group G acts on a second countable locally compact space M. Then*
(i) *(Uniqueness property) For $\nu_1, \nu_2 \in \mathcal{M}_G(M)$ we have*

$$\left(\nu_{1|\mathcal{B}_G(M)} = \nu_{2|\mathcal{B}_G(M)}\right) \Rightarrow (\nu_1 = \nu_2)\,. \tag{2.1}$$

(ii) *If $\tau \in \mathcal{M}(M)$ then*

$$\tau^*(B) := \int_G \tau(gB)\,\mu_G(\mathrm{d}g) \qquad \text{for all } B \in \mathcal{B}(M) \tag{2.2}$$

defines a G-invariant Borel measure on M with $\tau^*{}_{|\mathcal{B}_G(M)} = \tau_{|\mathcal{B}_G(M)}$. The mapping

$$*\colon \mathcal{M}(M) \to \mathcal{M}_G(M) \subseteq \mathcal{M}(M), \qquad \tau \mapsto \tau^* \tag{2.3}$$

is idempotent with $*_{|\mathcal{M}_G(M)} = \mathrm{id}_{|\mathcal{M}_G(M)}$. These statements are also valid for $\mathcal{M}^1(M)$ and $\mathcal{M}^1_G(M)$ instead of $\mathcal{M}(M)$ and $\mathcal{M}_G(M)$.
(iii) *Let $\nu \in \mathcal{M}_G(M)$ and $\tau \in \mathcal{M}(M)$ with $\tau = f \cdot \nu$. Then*

$$\tau^* = f^* \cdot \nu \qquad \text{with } f^*(m) := \int_G f(gm)\,\mu_G(\mathrm{d}g). \tag{2.4}$$

In particular, $\nu(\{m \in M \mid f^*(m) = \infty\}) = 0$, and f^* is $(\mathcal{B}_G(M), \mathcal{B}(\overline{\mathbb{R}}))$-measurable.
(iv) *Let M be a group and $\Theta_g\colon M \to M, \Theta_g(m) := gm$ a group homomorphism for all $g \in G$. Then Θ_g is an isomorphism, and $(\mathcal{M}^1_G(M), \odot)$ is a sub-semigroup of the convolution semigroup $(\mathcal{M}^1(M), \odot)$.*

Proof. Let $\nu_1, \nu_2 \in \mathcal{M}_G(M)$ be finite measures with $\nu_{1|\mathcal{B}_G(M)} = \nu_{2|\mathcal{B}_G(M)}$. Now assume that $\nu_1 \neq \nu_2$. Without loss of generality we may assume $\nu_1 = f \cdot \nu_2$ since otherwise we could replace ν_2 by $(\nu_1 + \nu_2)/2 \neq \nu_1$. As $\nu_1 \neq \nu_2$ but $\nu_1(M) = \nu_2(M) < \infty$ we conclude $\nu_2(\{m \in M \mid f(m) \neq 1\}) > 0$. In particular, there exists an index $n \in \mathbb{N}$ for which $E_n := \{m \in M \mid f(m) > 1 + 1/n\}$ has positive ν_2-measure. As M is second countable and locally compact the measure ν_2 is regular ([5], p. 213 (Satz 29.12)) and hence there exists a continuous function $\overline{f}\colon M \to [0,\infty)$ with $\int_M |\overline{f}(m) - f(m)|\,\nu_2(\mathrm{d}m) < \varepsilon := \nu_2(E_n)/12n$ ([5], p. 215 (Satz 29.14)). Now let $\overline{\overline{f}}\colon M \to [0,\infty]$ be defined by $\overline{\overline{f}}(m) := \int_G \overline{f}(gm)\,\mu_G(\mathrm{d}g)$. Let $\varepsilon' > 0$. As $\overline{f} \circ \Theta$ is continuous for each $g \in G$ there exist open neighbourhoods $U_g \subseteq G$ of g and $V_g \subseteq M$ of m such that $\left|\overline{f} \circ \Theta(g,m) - \overline{f} \circ \Theta(h,m')\right| < \varepsilon'$ for all $(h,m') \in U_g \times V_g$. As $G \times \{m\}$ is compact there exists a finite subcover $\bigcup_{j \le k} \left(U_{g_j} \times V_{g_j}\right) \supseteq G \times \{m\}$. Clearly,

$\bigcup_{j\le k} U_{g_j} = G$, and for $V := \bigcap_{j\le k} V_{g_j}$ we have $\sup_{m'\in V} \left|\overline{f}(hm) - \overline{f}(hm')\right| < \varepsilon'$ for all $h \in G$. In particular, this implies the continuity of $\overline{\overline{f}}$. We point out that the group action $\Theta\colon G \times M \to M$ is $(\mathcal{B}(G)\otimes\mathcal{B}(M), \mathcal{B}(M))$-measurable by 2.11(v). The Theorem of Fubini and the G-invariance of $\nu_1 = f\cdot\nu_2$ and ν_2 further imply

$$\begin{aligned}\left|\int_B \left(\overline{\overline{f}}(m) - f(m)\right)\,\nu_2(\mathrm{d}m)\right| &= \left|\int_B\int_G \left(\overline{f}(gm) - f(m)\right)\,\mu_G(\mathrm{d}g)\,\nu_2(\mathrm{d}m)\right| \\ &= \left|\int_G \left(\int_B \overline{f}(gm)\,\nu_2(\mathrm{d}m) - \int_B f(m)\,\nu_2(\mathrm{d}m)\right)\,\mu_G(\mathrm{d}m)\right| \\ &\le \int_G\int_{gB} \left|\overline{f}(m) - f(m)\right|\,\nu_2(\mathrm{d}m)\,\mu_G(\mathrm{d}g) < \varepsilon \qquad \text{for each } B\in\mathcal{B}(M).\end{aligned}$$

This implies $\int_C \left|\overline{\overline{f}}(m) - f(m)\right|\,\nu_2(\mathrm{d}m) < 2\varepsilon$ for each measurable $C \subseteq M$. Let $\overline{\overline{E}}_n := \{m\in M \mid \overline{\overline{f}}(m) > 1 + 1/2n\}$. As $f(m) - \overline{\overline{f}}(m) > 1/2n$ on $E_n \setminus \overline{\overline{E}}_n$ we obtain

$$2\varepsilon > \int_{E_n\setminus\overline{\overline{E}}_n} \left|\overline{\overline{f}}(m) - f(m)\right|\,\nu_2(\mathrm{d}m) \ge \left(\nu_2(E_n) - \nu_2(E_n\cap\overline{\overline{E}}_n)\right)/2n.$$

Inserting $\varepsilon = \nu_2\left(E_n\right)/12n$ yields

$$\nu_2\left(\overline{\overline{E}}_n\right) \ge \nu_2\left(E_n\cap\overline{\overline{E}}_n\right) \ge 2\nu_2(E_n)/3 > 0.$$

As $\overline{\overline{f}}$ is continuous, $\overline{\overline{E}}_n$ and hence the union $\overline{\overline{E}}_n^* := \bigcup_{g\in G} g\overline{\overline{E}}_n$ are open. In particular $\overline{\overline{E}}_n^* \in \mathcal{B}_G(M)$. As $\overline{\overline{f}}$ and ν_2 are G-invariant $\nu_3 := \overline{\overline{f}}\cdot\nu_2$ is also a finite G-invariant measure on $\mathcal{B}(M)$. In particular, $\nu_3(S) \ge (1+1/2n)\nu_2(S)$ holds for each measurable subset $S \subseteq \overline{\overline{E}}_n$. By 2.4(iv) there exist countably many $g_1, g_2, \ldots \in G$ with $\overline{\overline{E}}_n^* = \bigcup_{j\in N} g_j\overline{\overline{E}}_n = \sum_{j\in N} F_j$ with $F_1 := g_1\overline{\overline{E}}_n$ and $F_{j+1} = \bigcup_{i\le j+1} g_i\overline{\overline{E}}_n \setminus \sum_{i\le j} F_i$. The G-invariance of ν_3 and ν_2 guarantees $\nu_3(\overline{\overline{E}}_n^*) \ge (1+1/2n)\nu_2(\overline{\overline{E}}_n^*) > 0$. Finally, as $\overline{\overline{E}}_n^* \in \mathcal{B}_G(M)$ we obtain

$$\begin{aligned}0 = \nu_1(\overline{\overline{E}}_n^*) - \nu_2(\overline{\overline{E}}_n^*) &= \int_{\overline{\overline{E}}_n^*} \left(\left(\overline{\overline{f}}(m) - 1\right) + \left(f(m) - \overline{\overline{f}}(m)\right)\right)\,\nu_2(\mathrm{d}m) \\ &\ge \int_{\overline{\overline{E}}_n^*} \left(\overline{\overline{f}}(m) - 1\right)\,\nu_2(\mathrm{d}m) - \left|\int_{\overline{\overline{E}}_n^*} \left(f(m) - \overline{\overline{f}}(m)\right)\,\nu_2(\mathrm{d}m)\right| \\ &> \nu_2(\overline{\overline{E}}_n^*)\left(\frac{1}{2n} - \frac{3}{2}\frac{1}{12n}\right) = \frac{3}{8n}\nu_2(\overline{\overline{E}}_n^*) > 0\end{aligned}$$

which leads to the desired contradiction. That is, assertion (i) is proved for finite measures. Now let ν_1, ν_2 be arbitrary Borel measures. Due to 2.11(iv)

there exists a decomposition $M = \sum_{j=1}^{\infty} A_j$ of M into disjoint relatively compact G-invariant subsets $A_1, A_2, \ldots \in \mathcal{B}(M)$. For $i \in \{1,2\}$ and $j \in \mathbb{N}$ we define the finite measures $\nu_{i;j} \in \mathcal{M}(M)$ by $\nu_{i;j}(B) := \nu_i(B \cap A_j)$ for all $B \in \mathcal{B}(M)$. Clearly, as $A_j \in \mathcal{B}_G(M)$ the finite measures $\nu_{1;j}$ and $\nu_{2;j}$ coincide on $\mathcal{B}_G(M)$ and hence are equal. As $\nu_i = \sum_{j\in N} \nu_{i;j}$ we conclude $\nu_1 = \nu_2$ which finishes the proof of (i).
For the proof of (ii) we assume for the moment that $\tau(M) < \infty$. Recall that the group action $\Theta\colon G \times M \to M$ is $(\mathcal{B}(G) \otimes \mathcal{B}(M), \mathcal{B}(M))$-measurable (2.11(v)). Since $\mathrm{inv}_1\colon G \times M \to G \times M$, $\mathrm{inv}_1(g,m) := (g^{-1}, m)$ is measurable with respect to $(\mathcal{B}(G) \otimes \mathcal{B}(M))$ the composition $1_B \circ \Theta \circ \mathrm{inv}_1$ is $(\mathcal{B}(G) \otimes \mathcal{B}(M), \mathcal{B}(\mathbb{R}))$-measurable for each $B \in \mathcal{B}(M)$ and, as $0 \le 1_B \circ \Theta \circ \mathrm{inv}_1 \le 1$, also $\mu_G \otimes \tau$-integrable. The Theorem of Fubini implies

$$\begin{aligned}\int_{G\times M} 1_B \circ \Theta \circ \mathrm{inv}_1(g,m)\, \mu_G \otimes \tau(\mathrm{d}g, \mathrm{d}m) &= \int_G \int_M 1_B\left(g^{-1}m\right)\, \tau(\mathrm{d}m)\, \mu_G(\mathrm{d}g) \\ = \int_G \int_M 1_{gB}(m)\, \tau(\mathrm{d}m)\, \mu_G(\mathrm{d}g) &= \int_G \tau(gB)\, \mu_G(\mathrm{d}g)\end{aligned}$$

i.e. τ^* is well-defined for each $B \in \mathcal{B}(G)$. Obviously, $\tau^*(\emptyset) = 0$ and $\tau^*(B) \ge 0$ for all $B \in \mathcal{B}(M)$. For disjoint Borel subsets $B_1, B_2, \ldots \in \mathcal{B}(M)$ we have

$$\begin{aligned}\tau^*\left(\sum_{j=1}^{\infty} B_j\right) &= \int_G \tau\left(g \sum_{j=1}^{\infty} B_j\right)\, \mu_G(\mathrm{d}g) = \int_G \sum_{j=1}^{\infty} \tau\left(gB_j\right)\, \mu_G(\mathrm{d}g) \\ &= \sum_{j=1}^{\infty} \int_G \tau\left(gB_j\right)\, \mu_G(\mathrm{d}g) = \sum_{j=1}^{\infty} \tau^*(B_j)\end{aligned}$$

and hence $\tau^* \in \mathcal{M}^+(M)$. (The second equation follows from the fact that Θ_g is a homeomorphism, the third from the Theorem of Fubini.) For fixed $B \in \mathcal{B}(M)$ let temporarily $\varphi\colon G \to \mathbb{R}$ be defined by $\varphi(g) := \tau(gB)$. As μ_G is left- and right-invariant we conclude $\tau^*(hB) = \int_G \varphi(gh)\, \mu_G(\mathrm{d}g) = \int_G \varphi(g)\, \mu_G(\mathrm{d}g) = \tau^*(B)$ for all $h \in G$. As $B \in \mathcal{B}(M)$ was arbitrary and since $\tau(M) < \infty$ this yields $\tau^* \in \mathcal{M}_G(M)$. By definition $\tau(A) = \tau^*(A)$ for all $A \in \mathcal{B}_G(M)$, and $(\tau^*)^*(A) = \int_G \tau^*(gA)\, \mu_G(\mathrm{d}g) = \int_G \tau^*(A)\, \mu_G(\mathrm{d}g) = \tau^*(A)$. From (i) we have $(\tau^*)^* = \tau^*$. This proves (ii) for finite τ. Now let $\tau \in \mathcal{M}(M)$ be an arbitrary Borel measure. Due to 2.11(iv) there exists a decomposition $M = \sum_{j=1}^{\infty} A_j$ of M in disjoint relatively compact G-invariant subsets $A_1, A_2, \ldots \in \mathcal{B}(M)$. As in the proof of (i) for $j \in \mathbb{N}$ we temporarily define the measure $\tau_j \in \mathcal{M}(M)$ by $\tau_j(B) := \tau(B \cap A_j)$ for all $B \in \mathcal{B}(M)$. In particular, all τ_j are finite and $\tau = \sum_{j=1}^{\infty} \tau_j$. As the integrand is non-negative the Theorem of Fubini implies $\tau^*(B) = \int_G \sum_{j=1}^{\infty} \tau_j(gB)\, \mu_G(\mathrm{d}g) = \sum_{j=1}^{\infty} \int_G \tau_j(gB)\, \mu_G(\mathrm{d}g) = \sum_{j=1}^{\infty} \tau_j^*(B)$. In particular, the term $\tau^*(B)$ is well-defined ($\tau^*(B) \in [0,\infty]$). The measure τ is the sum of countably many G-invariant measures τ and hence itself G-invariant. Similarly, one concludes $(\tau^*)^* = \tau^*$. As $\tau^*(K) \le \tau^*(GK) = \tau(GK) < \infty$ for all compact subsets

$K \subseteq M$ we have $\tau^* \in \mathcal{M}_G(M)$. The assertion concerning the probability measures are obvious since $\tau^*(M) = \tau(M) = 1$ in this case.
Let $\tau = f \cdot \nu$, and let the measures τ_j be defined as in the proof of (ii), i.e. $\tau_j(B) = \tau(B \cap A_j)$. For all $B \in \mathcal{B}(M)$ we obtain the inequality

$$\begin{aligned}\infty > \tau_j^*(B) &= \int_G \left(\int_{gB} f(m) 1_{A_j}(m)\, \nu(\mathrm{d}m) \right) \mu_G(\mathrm{d}g) \\ &= \int_G \left(\int_B f(g^{-1}m) 1_{A_j}(g^{-1}m)\, \nu^{\Theta_{g^{-1}}}(\mathrm{d}m) \right) \mu_G(\mathrm{d}g) \\ &= \int_B \left(\int_G f(gm)\, \mu_G(\mathrm{d}g) \right) 1_{A_j}(m)\, \nu(\mathrm{d}m) = \int_B f^*(m) 1_{A_j}(m)\, \nu(\mathrm{d}m)\end{aligned}$$

where the third equation follows from the G-invariance of A_j and ν (in particular, $g^{-1}m \in A_j$ iff $m \in A_j$ and $\nu = \nu^{\Theta_{g^{-1}}}$) and the fact that μ_G invariant under inversion ([63], p. 123 (Theorem 5.14) and [5], pp. 180 (Lemma 26.2) and 193 (Korollar 27.3)). As $B \in \mathcal{B}(M)$ was arbitrary this implies $\tau_j^* = f^* \cdot 1_{A_j} \cdot \nu$ and hence $\tau^* = f^* \cdot \nu$. The integral $\int_G f(gm)\, \mu_G(\mathrm{d}g)$ can only be infinite on a ν-zero set (Fubini). This implies $\nu((f^*)^{-1}(\{\infty\})) = \sum_{j=1}^{\infty} \nu(\{m \in A_j \mid \int_G f(gm)\, \mu_G(\mathrm{d}g) = \infty\}) = 0$. The continuity of the group action guarantees the $\mathcal{B}(G \times M)$-measurability of the function $F(g, m) := f(gm)$. Applying Tonelli's theorem ([5], p. 157 (Satz 23.6)) to the product measure $\mu_G \otimes \nu$ proves the $\mathcal{B}(M)$-measurability of $f^*(\cdot) = \int_G F(g, \cdot)\, \mu_G(\mathrm{d}g)$. As f^* is constant on all G-orbits this in particular implies that f^* is also $\mathcal{B}_G(M)$-measurable which completes the proof of assertion (iii).
The first assertion of (iv) is obvious as $\Theta_g \colon M \to M$, $m \mapsto gm$, is a homeomorphism. The group G acts on the product space $M \times M$ by $(\Theta, \Theta) \colon G \times (M \times M)$, $(\Theta, \Theta)(g, (m_1, m_2)) := (\Theta(g, m_1), \Theta(g, m_2))$. For all $g \in G$ and $B_1, B_2 \in \mathcal{B}(M)$ and $\nu_1, \nu_2 \in \mathcal{M}_G(M)$ we immediately obtain $(\nu_1 \otimes \nu_2)^{(\Theta_g, \Theta_g)}(B_1 \times B_2) = (\nu_1 \otimes \nu_2)(g^{-1}B_1 \times g^{-1}B_2) = (\nu_1 \otimes \nu_2)(B_1 \times B_2)$. As M is second countable ν is σ-finite by 2.8(iii). Using ([5], p. 152 (Satz 22.2)) one concludes that the measures $\nu_1 \otimes \nu_2$ and $(\nu_1 \otimes \nu_2)^{(\Theta_g, \Theta_g)}$ coincide, i.e. that $\nu_1 \otimes \nu_2 \in \mathcal{M}_G(M \times M)$. Since Θ_g is a group homomorphism the mapping $\pi \colon M \times M \to M$, $\pi(m_1, m_2) := m_1 m_2$ is equivariant: $g\pi(m_1, m_2) = g(m_1 m_2) = \Theta_g(m_1 m_2) = \Theta_g(m_1)\Theta_g(m_2) = \pi(gm_1, gm_2)$. Hence the convolution product $\nu_1 \odot \nu_2 \in \mathcal{M}_G^1(M)$ which proves (iv). □

Theorem 2.18 says that on a second countable locally compact space M the G-invariant Borel measures are uniquely determined by their values on the sub-σ-algebra $\mathcal{B}_G(M)$. In fact, this insight will be crucial in the following. Counterexample 2.19 shows that this need not be true if these measures are not Borel measures.

Counterexample 2.19. Let $G = M = ([0, 1), \oplus)$ (cf. Example 2.13(iii)) and $\Theta(g, m) := g \oplus m$. This action is transitive so that $\mathcal{B}_G(M) = \{\emptyset, M\}$. Let further $\nu_1, \nu_2 \in \mathcal{M}^+(M)$ be given by

$$\nu_1(B) := \begin{cases} 0 & \text{if } \lambda(B) = 0 \\ \infty & \text{else} \end{cases} \qquad \text{and}$$
$$\nu_2(B) := \begin{cases} 0 & \text{if } B = \emptyset \\ \infty & \text{else} \end{cases}$$

for $B \in \mathcal{B}([0,1),\oplus)) = \mathcal{B}([0,1))$. The measures ν_1 and ν_2 are G-invariant with $\nu_1([0,1)) = \nu_2([0,1)) = \infty$, i.e. $\nu_{1|B_G(M)} = \nu_{2|B_G(M)}$, but $\nu_1 \neq \nu_2$.

Let $\mathcal{A}_0 \subseteq \mathcal{A} \subseteq A_1$ denote σ-algebras over M and $\eta \in \mathcal{M}^+(M,\mathcal{A})$. Clearly, its restriction $\eta_{|\mathcal{A}_0}$ is a measure on the sub-σ-algebra $\mathcal{A}_0$. It does always exist and is unique. In contrast, an extension of η to $\mathcal{A}_1$ may not exist, and if an extension exists it need not be unique. The concept of measure extensions will be important in the following section. Example 2.20 should help the reader to get familiar with it.

Example 2.20. (i) Let $\mathcal{A}_1 \supseteq \mathcal{A} := \{\emptyset, \Omega\}$ be a σ-algebra over a non-empty set Ω. Then $\eta(\emptyset) := 0$ and $\eta(\Omega) := 1$ defines a probability measure on $\mathcal{A}$, and each $\nu \in \mathcal{M}^1(\Omega,\mathcal{A}_1)$ is an extensions of η.
(ii) The Lebesgue measure on $[0,1]$ cannot be extended from $\mathcal{B}([0,1])$ to the power set $\mathcal{P}([0,1])$ ([44], p. 50).
(iii) Let $\mathcal{A}$ be a σ-algebra over a non-empty set Ω and $\eta \in \mathcal{M}^+(\Omega,\mathcal{A})$. Further, let $\mathcal{N}_\eta := \{N \subseteq \Omega \mid N \subseteq A \text{ for an } A \in \mathcal{A} \text{ with } \eta(A) = 0\}$, the system of the η-zero sets. Then $\mathcal{A}_\eta := \{A + N \mid A \in \mathcal{A},\, N \in \mathcal{N}_\eta\}$ is also a σ-algebra over Ω. There exists a unique extension $\eta_v \in \mathcal{M}^+(\Omega,\mathcal{A}_\eta)$ of η, called the *completion* of η, and this extension is given by $\eta_v(A+N) := \eta(A)$ ([28], pp. 34f.). In the same way one concludes that an extension $\eta_1 \in \mathcal{M}(\Omega,\mathcal{A}_1)$ of η on an intermediate σ-algebra $\mathcal{A}_1$, i.e. $\mathcal{A} \subseteq \mathcal{A}_1 \subseteq \mathcal{A}_\eta$, is also unique. In particular, $\eta_1 = \eta_{v|\mathcal{A}_1}$.

Combining 2.18(ii) and (i) yields an interesting corollary from the field of measure extensions.

Corollary 2.21. *Suppose that a compact group G acts on a second countable locally compact space M, and assume that $\kappa_1 \in \mathcal{M}(M)$ is an extension of $\kappa \in \mathcal{M}^+(M,\mathcal{B}_G(M))$. Then there exists a unique G-invariant extension $\kappa_0 \in \mathcal{M}_G(M)$ of κ.*

In particular, one can simulate G-invariant distributions without determining them explicitly.

Corollary 2.22. *Suppose that a compact group G acts on a second countable locally compact space M. Further, let X and Y denote independent random variables on G and M, resp., which are μ_G- and τ-distributed, resp. As μ_G is invariant under inversion (cf. 2.12) we obtain*

$$\begin{aligned} \text{Prob}(Z := \Theta(X,Y) \in B) &= \int_G \int_M 1_B(gm)\, \mu_G(\mathrm{d}g)\, \tau(\mathrm{d}m) \qquad (2.5) \\ &= \int_G \tau(g^{-1}B)\, \mu_G(\mathrm{d}g) = \int_G \tau(hB)\, \mu_G(\mathrm{d}h) = \tau^*(B) \end{aligned}$$

for all $B \in \mathcal{B}(M)$, i.e. Z is τ^*-distributed.

2.2 Definition of Property (∗) and Its Implications (Main Results)

In this section the main results of this book are derived. Property (∗) formulates the general assumptions which will be considered in the remainder. In particular, (∗) makes the assumptions (V) from the introduction precise.

(∗) The 5-tupel (G, S, T, H, φ) is said to have Property (∗) if the following conditions are fulfilled:
- -a- G, S, T, H are second countable.
- -b- T and H are locally compact, and S is a Hausdorff space.
- -c- G is a compact group which acts on S and H.
- -d- G acts transitively on S.
- -e- $\varphi\colon S \times T \to H$ is a surjective measurable mapping.
- -f- The mapping φ is equivariant with respect to the G-action $g(s,t) := (gs,t)$ on the product space $S \times T$ and the G-action on H, i.e. $g\varphi(s,t) = \varphi(gs,t)$ for all $(g,s,t) \in G \times S \times T$.

At the first sight Property (∗) might seem to be very restrictive. However, the topological conditions -a- and -b- are fulfilled for nearly all spaces of practical relevance (cf. Example 2.5 and Chapter 4). In particular, conditions -a- and -b- are easy to verify. In general, also condition -e- is fulfilled as continuous mappings are measurable. As will become clear in Chapter 4 at hand of several examples also the verification of -c-, -d- and -f- usually needs no more than straight-forward arguments. The transitivity of the G-action (condition -d-) implies the compactness of S. By 2.14(v) there is a unique G-invariant probability measure on S.

Definition 2.23. *A second countable space M is called* Polish *if M is completely metrizable. Let M_1 and M_2 be locally compact. A continuous mapping $\chi\colon M_1 \to M_2$ is* proper *if the pre-image $\chi^{-1}(K) \subseteq M_1$ is compact for each compact $K \subseteq M_2$. The unique G-invariant probability measure on S is denoted with $\mu_{(S)}$. Further, $E_G(\{h\})$ denotes the orbit $Gh \subseteq H$ and, more generally, $E_G(F) := GF$ for $F \subseteq H$.*

Remark 2.24. (i) Suppose that the 5-tuple (G, S, T, H, φ) has Property (∗). By 2.4(iii) and [59], p. 149 (Satz 13.16), the spaces G, S, T and H are Polish.
(ii) On a second countable locally compact space the terms 'Borel measure' and 'Radon measure' coincide (cf. 2.47 and 2.48(i)).

Lemma 2.25. *Assume that the 5-Tupel (G, S, T, H, φ) has Property (∗). Further, let $s_0 \in S$ be arbitrary but fixed.*
(i) The embedding

$$\iota\colon T \to S \times T \qquad\qquad \iota(t) := (s_0, t) \tag{2.6}$$

is continuous.
(ii) $\varphi^{-1}(A) = S \times \iota^{-1}(\varphi^{-1}(A))$ *for all* $A \in \mathcal{B}_G(H)$*. In particular,* $\mathcal{B}_0(T) := \{\iota^{-1}(\varphi^{-1}(A)) \mid A \in \mathcal{B}_G(H)\}$ *is a sub-*σ*-algebra of* $\mathcal{B}(T)$*. The assignment* $A \mapsto \iota^{-1}(\varphi^{-1}(A))$ *defines an isomorphism of* σ*-algebras between* $\mathcal{B}_G(H)$ *and* $\mathcal{B}_0(T)$*. Its inverse is given by* $D \mapsto \varphi(S \times D)$*.*
(iii) $\varphi(S \times \{t\}) = E_G(\{\varphi(s_0, t)\})$*, i.e.* $\varphi(S \times \{t\})$ *is the G-Orbit of* $\varphi(s_0, t)$*.*
(iv) $(t_1 \sim t_2) \overset{\text{def}}{\Longleftrightarrow} (\varphi(S \times \{t_1\}) = \varphi(S \times \{t_2\}))$ *defines an equivalence relation on* T*. For each* $F \subseteq T$ *let* $E_T(F) := \{t \in T \mid t \sim t_0$ *for a* $t_0 \in F\}$*. Then* $\mathcal{B}_E(T) := \{C \in \mathcal{B}(T) \mid C = E_T(C)\}$ *is also a sub-*σ*-algebra of* $\mathcal{B}(T)$*.*
(v) $\mathcal{B}_0(T) \subseteq \mathcal{B}_E(T) \subseteq \mathcal{B}(T)$*.*
(vi) The assignment $\kappa \mapsto \kappa_*$ *given by* $\kappa_*(D) = \kappa(\varphi(S \times D))$ *for all* $D \in \mathcal{B}_0(T)$ *induces bijections between* $\mathcal{M}^1(H, \mathcal{B}_G(H))$ *and* $\mathcal{M}^1(T, \mathcal{B}_0(T))$*, resp. between* $\mathcal{M}^+(H, \mathcal{B}_G(H))$ *and* $\mathcal{M}^+(T, \mathcal{B}_0(T))$*.*

Proof. The set $\{U_1 \times U_2 \mid U_1$ open in S, U_2 open in $T\}$ is a basis for the product topology of $S \times T$. The pre-image $\iota^{-1}(U_1 \times U_2)$ is a open subset of T, namely $= U_2$ if $s_0 \in U_1$ and $= \emptyset$ else. This proves (i). The inclusion $\varphi(s,t) \in A \in B_G(H)$ implies $g\varphi(s,t) = \varphi(gs,t) \in A$ for all $g \in G$. As G acts transitively on S this implies $\varphi^{-1}(A) = S \times F$ for a particular subset $F \subseteq T$. Clearly, $F = \mathrm{pr}_T\left(\varphi^{-1}(A)\right) = \iota^{-1}\left(\varphi^{-1}(A)\right) \in B(T)$ where $\mathrm{pr}_T\colon S \times T \to T$ denotes the projection onto the second component. Trivially, $\emptyset = \iota^{-1}\left(\varphi^{-1}(\emptyset)\right)$ is contained in $\mathcal{B}_0(T)$. Further, $\varphi^{-1}(A^c) = (S \times F)^c = S \times F^c$, i.e. $\iota^{-1}\left(\varphi^{-1}(A^c)\right) = \left(\iota^{-1}\left(\varphi^{-1}(A)\right)\right)^c$. Similarly, one verifies $\bigcup_{j=1}^{\infty} \iota^{-1}\left(\varphi^{-1}(A_j)\right) = \iota^{-1}\left(\varphi^{-1}\left(\bigcup_{j=1}^{\infty} A_j\right)\right)$. Hence $\mathcal{B}_0(T)$ is a σ-algebra. The mapping $\mathcal{B}_G(H) \to \mathcal{B}_0(T)$, $A \mapsto \iota^{-1}\left(\varphi^{-1}(A)\right)$ is injective and, due to the construction of $\mathcal{B}_0(T)$, also bijective. In particular, one immediately concludes that $A \mapsto \iota^{-1}\left(\varphi^{-1}(A)\right)$ induces an isomorphism of σ-algebras. The equality $\varphi(\varphi^{-1}(A)) = A$ verifies the final assertion of (ii). As G operates transitively on S one immediately verifies $E_G(\{\varphi(s_0,t)\}) = G\varphi(s_0,t) = \varphi(Gs_0 \times \{t\}) = \varphi(S \times \{t\})$. Since two orbits are either identical or disjoint the equivalence relation $\sim$ on T is well-defined. With straight-forward arguments one verifies that $\mathcal{B}_E(T)$ is also a σ-algebra. Let $A \in \mathcal{B}_G(H)$ and $t \in E_T\left(\iota^{-1}\left(\varphi^{-1}(A)\right)\right)$. In particular, $\varphi(S \times \{t\}) \subseteq A$, i.e. $t \in \iota^{-1}\left(\varphi^{-1}(A)\right)$ and hence $\iota^{-1}\left(\varphi^{-1}(A)\right) = E_T\left(\iota^{-1}\left(\varphi^{-1}(A)\right)\right) \in \mathcal{B}_E(T)$. Because $\mathcal{B}_G(H)$ and $\mathcal{B}_0(T)$ are isomorphic assertion (vi) follows immediately from (ii). □

Theorem 2.26. *Suppose that the* 5*-Tupel* (G, S, T, H, φ) *has Property (∗). As in Lemma 2.25* $s_0 \in S$ *is arbitrary but fixed. Further,*

$$\Phi\colon \mathcal{M}^\sigma(T) \to \mathcal{M}_G^+(H), \qquad\qquad \Phi(\tau) := \left(\mu_{(\mathrm{S})} \otimes \tau\right)^\varphi. \tag{2.7}$$

Then the following statements are valid:
(i) *Let* $\nu \in \mathcal{M}_G(H)$ *and* $\tau \in \mathcal{M}^\sigma(T)$*. Then* $\nu_* \in \mathcal{M}^\sigma(T, \mathcal{B}_0(T))$*, and* $\tau \in$

$\Phi^{-1}(\nu)$ *iff* $\tau_{|\mathcal{B}_0(T)} = \nu_*$.
(ii) $\tau(T) = (\Phi(\tau))(H)$ *for each* $\tau \in \mathcal{M}^\sigma(T)$.
(iii) *The following properties are equivalent:*

(α) $\Phi(\mathcal{M}^1(T)) = \mathcal{M}^1_G(H)$.
(β) For each $\nu \in \mathcal{M}^1_G(H)$ *the induced measure* $\nu_* \in \mathcal{M}^1(T, \mathcal{B}_0(T))$ *has an extension* $\tau_\nu \in \mathcal{M}^1(T)$.

Any of these conditions implies $\Phi(\mathcal{M}^\sigma(T)) \supseteq \mathcal{M}_G(H)$.
(iv) *If each* $h \in H$ *has a* G*-invariant neighbourhood* U_h *with relative compact pre-image* $(\varphi \circ \iota)^{-1}(U_h) \subseteq T$ *then* $\Phi(\mathcal{M}(T)) \subseteq \mathcal{M}_G(H)$.
(v) *If* $\varphi\colon S \times T \to H$ *is continuous then* $\Phi^{-1}(\mathcal{M}_G(H)) \subseteq \mathcal{M}(T)$.
(vi) *Suppose that*

$$\Phi(\mathcal{M}(T)) \subseteq \mathcal{M}_G(H) \qquad \text{and} \qquad \Phi^{-1}(\mathcal{M}_G(H)) \subseteq \mathcal{M}(T) \tag{2.8}$$

(cf. (iv) and (v)). Then the following properties are equivalent:

(α) $\Phi(\mathcal{M}(T)) = \mathcal{M}_G(H)$.
(β) For each $\nu \in \mathcal{M}^1_G(H)$ *the induced probability measure* $\nu_* \in \mathcal{M}^1(T, \mathcal{B}_0(T))$ *has an extension* $\tau_\nu \in \mathcal{M}^1(T)$.
(γ) For each $\nu \in \mathcal{M}_G(H)$ *the induced measure* $\nu_* \in \mathcal{M}^\sigma(T, \mathcal{B}_0(T))$ *has an extension* $\tau_\nu \in \mathcal{M}^+(T)$.

Any of these conditions implies $\Phi^{-1}(\mathcal{M}_G(H)) = \mathcal{M}(T)$.
(vii) *Let* $\nu \in \mathcal{M}_G(H)$, $f \geq 0$ *a* G*-invariant numerical function and* $\nu = \left(\mu_{(\mathrm{S})} \otimes \tau_\nu\right)^\varphi$. *Then* $\eta := f \cdot \nu \in \mathcal{M}^+_G(H)$. *If* f *is locally* ν*-integrable then* $\eta \in \mathcal{M}_G(H)$ *and*

$$\eta = \left(\mu_{(\mathrm{S})} \otimes f_T \cdot \tau_\nu\right)^\varphi \qquad \text{with} \qquad f_T(t) := f\left(\varphi(s_0, t)\right). \tag{2.9}$$

If $T \subseteq H$ *and* $E_T(\{t\}) = \varphi(S \times \{t\}) \cap T$ *(i.e.* $= E_G(\{\varphi(s_0, t)\}) \cap T$*) then* $f_T(t) = f(t)$, *i.e.* f_T *is given by the restriction* $f_{|T}$ *of* $f\colon G \to [0, \infty]$ *to* T.

Proof. For each $\tau \in \mathcal{M}^\sigma(T)$ the product measure $\left(\mu_{(\mathrm{S})} \otimes \tau\right) \in \mathcal{M}^\sigma(S \times T)$ is invariant under the G-action on $S \times T$. The G-invariance of φ implies $\left(\mu_{(\mathrm{S})} \otimes \tau\right)^\varphi \in \mathcal{M}^+_G(H)$, i.e. Φ is well-defined. By 2.11(iv) there exist disjoint G-invariant relatively compact subsets $A_1, A_2, \ldots \in \mathcal{B}_G(H)$ with $H = \sum_{j \in \mathbb{N}} A_j$. By the definition of Borel measures for any $\nu \in \mathcal{M}_G(H)$ we have $\nu_*(\iota^{-1}(\varphi^{-1}(A_j))) = \nu_j(A_j) < \infty$ for all $j \in \mathbb{N}$ which proves the first assertion of (i). Let $\tau_{|\mathcal{B}_0(T)} = \nu_*$. For all $A \in \mathcal{B}_G(H)$ we have

$$\begin{aligned}\left(\mu_{(\mathrm{S})} \otimes \tau\right)^\varphi(A) &= \left(\mu_{(\mathrm{S})} \otimes \tau\right)\left(S \times \iota^{-1}\left(\varphi^{-1}(A)\right)\right) = \tau\left(\iota^{-1}\left(\varphi^{-1}(A)\right)\right) \\ &= \nu_*\left(\iota^{-1}\left(\varphi^{-1}(A)\right)\right) = \nu(A),\end{aligned}$$

i.e. $\left(\mu_{(\mathrm{S})} \otimes \tau\right)^\varphi{}_{|\mathcal{B}_G(H)} = \nu_{|\mathcal{B}_G(H)}$, and the uniqueness property of G-invariant Borel measures (2.18(i)) implies $\left(\mu_{(\mathrm{S})} \otimes \tau\right)^\varphi = \nu$. On the other hand, if $\nu = \left(\mu_{(\mathrm{S})} \otimes \tau\right)^\varphi$ Lemma 2.25(vi) yields the equation

$$\nu_*\left(\iota^{-1}\left(\varphi^{-1}(A)\right)\right)=\nu(A)=\left(\mu_{(\mathrm{S})}\otimes\tau\right)\left(S\times\iota^{-1}\left(\varphi^{-1}(A)\right)\right)=\tau\left(\iota^{-1}\left(\varphi^{-1}(A)\right)\right)$$

for all $A \in \mathcal{B}_G(H)$, that is $\tau_{|\mathcal{B}_0(T)} = \nu_*$ which completes the proof of (i). From $\left(\mu_{(\mathrm{S})}\otimes\tau\right)^{\varphi}(H) = \mu_{(\mathrm{S})}(S)\cdot\tau(T) = \tau(T)$ we immediately obtain (ii). As H is locally compact each h has a compact neighbourhood $K'_h \subseteq U_h$. As U_h is G-invariant the set $K_h := GK'_h \subseteq U_h$ is a G-invariant neighbourhood of h with relatively compact pre-image $(\varphi\circ\iota)^{-1}(K_h)$. Hence $\varphi^{-1}(K_h) = S\times\iota^{-1}\left(\varphi^{-1}(K_h)\right)$. As $\iota^{-1}\left(\varphi^{-1}(K_h)\right)$ is relatively compact the set $\varphi^{-1}(K_h)$ is relatively compact, too, and (iv) follows from 2.14(ii)(δ). Suppose that $\left(\mu_{(\mathrm{S})}\otimes\tau\right)^{\varphi} = \nu \in \mathcal{M}_G(H)$ and φ is continuous. If $K_0 \subseteq T$ is compact $\tau(K_0) = \left(\mu_{(\mathrm{S})}\otimes\tau\right)(S\times K_0) \le \nu\left(\varphi\left(S\times K_0\right)\right) < \infty$ since $\varphi\left(S\times K_0\right)$ is a continuous image of a compact subset and hence itself compact. This finishes the proof of (v). Now assume that (iii) (α) holds where φ need not be continuous. Then for each $\nu \in \mathcal{M}^1_G(H)$ there exists a probability measure $\tau_\nu \in \mathcal{M}^1(T)$ with $\nu = \left(\mu_{(\mathrm{S})}\otimes\tau_\nu\right)^{\varphi}$. Due to (i) τ_ν must be an extension of ν_*. On the other hand, each extension τ of ν_* fulfills the equation $\nu = \left(\mu_{(\mathrm{S})}\otimes\tau\right)^{\varphi}$ which proves the equivalence of (iii) (α) and (iii) (β). Let the subsets $A_j \in \mathcal{B}_G(H)$ be defined as in the proof of (i), and set $\nu_j(B) := \nu(B\cap A_j)$ for each $B \in \mathcal{B}(H)$. Then ν_j is a finite measure on M, and hence (iii)(α) guarantees the existence of a finite $\tau_j \in \mathcal{M}(T)$ with $\Phi(\tau_j) = \nu_j$. Consequently, $(\mu_{(\mathrm{S})}\otimes\tau)^{\varphi}$ for $\tau := \sum_{j\in\mathbb{N}}\tau_j$. By (i) $\tau_{|\mathcal{B}_0(T)} = \nu_*$, and hence $\tau \in \mathcal{M}^\sigma(T)$ which completes the proof of (iii). Next, we are going to prove (vi). Due to (ii) condition (vi) (α) implies the equality $\Phi(\mathcal{M}^1(T)) = \mathcal{M}^1_G(H)$, and due to (iii) this is equivalent to condition (vi) (β). From (iii) condition (β) in turn implies $\Phi(\mathcal{M}^\sigma(T)) \supseteq \mathcal{M}_G(H)$ and with (i) this implies (γ). Assume that (γ) holds. By (i) this is equivalent to $\Phi(\mathcal{M}^\sigma(T)) \supseteq \mathcal{M}_G(H)$. Together with the assumption (2.8) this yields (iii)(α) which proves the equivalence of (α), (β) and (γ). The final assertion of (vi) follows immediately from (α) and the assumption $\Phi^{-1}(\mathcal{M}_G(H)) \subseteq \mathcal{M}(T)$. Applying 2.14(iv) yields the first and the second assertion of (vii). In particular, $f_T = f\circ\varphi\circ\iota$ is measurable. Recall that G acts transitively on S. As f is G-invariant we conclude $f(\varphi(s,t)) = f(g\varphi(s_0,t)) = f(\varphi(s_0,t)) = f_T(t)$ for all $(s,t) \in S\times T$. This in turn implies

$$\int_B 1\,\eta(\mathrm{d}h) = \int_B f(h)\,\nu(\mathrm{d}h) = \int_{\varphi^{-1}(B)} f\circ\varphi(s,t)\,\left(\mu_{(\mathrm{S})}\otimes\tau_\nu\right)(\mathrm{d}s,dt)$$

$$\int_{\varphi^{-1}(B)} f_T(t)\,\left(\mu_{(\mathrm{S})}\otimes\tau_\nu\right)(\mathrm{d}s,dt)$$

for all $B \in \mathcal{B}(H)$, i.e. $\eta = \left(\mu_{(\mathrm{S})}\otimes f_T\cdot\tau_\nu\right)^{\varphi}$. The final assertion of (vii) is trivial. □

Remark 2.27. (i) For continuous $\varphi\colon S\times T\to H$ the condition that each $h \in H$ has a G-invariant neighbourhood U_h with relatively compact pre-image $(\varphi\circ\iota)^{-1}(U_h)$ (cf. 2.26(iv)) is fulfilled iff φ is proper. Suppose that the left-hand condition holds. Any compact subset $K \subseteq H$ can be covered by finitely

many G-invariant neighbourhoods U_h which implies that φ is proper. The inverse direction follows from the definition of local compactness and the fact that for G-invariant U_h the pre-image $(\varphi\circ\iota)^{-1}(U_h)$ equals $S\times F$ for a suitable $F\subseteq T$.
(ii) Example 2.28 illustrates the meaning of the conditions (2.8). In particular, 2.28(iii) shows that $\Phi(\mathcal{M}(T)) = \mathcal{M}_G(H)$ does not necessarily mean that $\Phi^{-1}(\mathcal{M}_G(H)) = \mathcal{M}(T)$. Note, however, that $\Phi(\mathcal{M}^1(T)) = \mathcal{M}^1_G(H)$ implies $\Phi^{-1}(\mathcal{M}^1_G(H)) = \mathcal{M}^1(T)$ as $\tau(T) = \Phi(\tau)(H)$.

Example 2.28. (i) Let $G = \{e_G\}$, $S = \{s_0\}$, $T = \mathbb{N}$, $H = \{0\}$ and $\varphi(s,t) := 0$ for all $t \in \mathbb{N}$. The group G acts trivially on S and H, and as $\mathbb{N}$ is usually equipped with the discrete topology. The 5-tuple (G,S,T,H,φ) has Property $(*)$. Clearly, $\mathcal{B}(\mathbb{N}) = \mathcal{P}(\mathbb{N})$ and $\Phi^{-1}(\mathcal{M}(H)) \subseteq \mathcal{M}(\mathbb{N}) = \mathcal{M}^\sigma(\mathbb{N})$ but $\Phi(\mathcal{M}(\mathbb{N})) = \mathcal{M}^+(H)$ which is a proper superset of $\mathcal{M}(H)$, i.e. the right-hand condition of (2.8) is not fulfilled. To be precise, the set of all finite measures on $\mathbb{N}$ is mapped surjectively onto $\mathcal{M}(H)$ wheras all infinite Borel measures on $\mathbb{N}$ are mapped onto the single measure $\eta \in \mathcal{M}^+(H) \setminus \mathcal{M}(H)$ which is given by $\eta(0) = \infty$.
(ii) Let $G = \{e_G\}$, $S = \{s_0\}$, $T = \{0\} \cup \{1/n \mid n \in \mathbb{N}\}$, $H = \mathbb{N}_0$ and $\varphi(s_0,t) := 1/t$ for $t \neq 0$ while $\varphi(s_0,0) = 0$. The group G acts trivially on S and H, and as H is usually equipped with the discrete topology. We view T as a subset of $\mathbb{R}$ and equip T with the induced topology. Then $\{1/n\}$ is open in T for each $n \in \mathbb{N}$, and any open neighbourhood of 0 contains a set $\{1/n \mid n \geq m\}$ with $m \in \mathbb{N}$. In particular, T is compact, $\mathcal{B}(T) = \mathcal{P}(T)$, and the mapping $\varphi\colon S \times T \to H$ is measurable. The 5-tuple (G,S,T,H,φ) has Property $(*)$. As φ is bijective the pre-image $\Phi^{-1}(\nu)$ is singleton for each $\nu \in \mathcal{M}_G(\mathbb{N}_0) = \mathcal{M}(\mathbb{N}_0) = \mathcal{M}^\sigma(\mathbb{N}_0)$. To be precise, $\nu = \Phi(\tau_\nu)$ where $\tau_\nu(t) := \nu(\varphi(s_0,t))$ for each $t \in T$. Let $\nu' \in \mathcal{M}(\mathbb{N}_0)$ be given by $\nu'(n) := 1$ for each $n \in \mathbb{N}_0$. For each compact neighbourhood K of 0 we have $\tau_{\nu'}(K) = \infty$ and hence $\tau_{\nu'} \notin \mathcal{M}(T)$. Note that all $\tau \in \mathcal{M}(T)$ are finite since T is compact. In particular, $\Phi(\mathcal{M}(T)) \subseteq \mathcal{M}_G(H)$, i.e. the left-hand condition from (2.8) is fulfilled, but $\Phi^{-1}(\mathcal{M}_G(H)) \not\subseteq \mathcal{M}(T)$. In particular $\Phi(\mathcal{M}^\sigma(T)) = \mathcal{M}^\sigma(\mathbb{N}_0) = \mathcal{M}(\mathbb{N}_0)$ whereas $\Phi(\mathcal{M}(T))$ equals the set of all finite Borel measures on $\mathbb{N}_0$.
(iii) Let G, S and H be defined as in (i) while $T = \{0\} \cup \{1/n \mid n \in \mathbb{N}\} \cup \{2,3,\ldots\}$ and $\varphi(s_0,0) := 0$, $\varphi(s_0,1/n) := n$ and $\varphi(s_0,n) := n$ for $n \in \mathbb{N}$. Equipped with the induced topology (from $\mathbb{R}$) the space T is second countable and locally compact, and the 5-tuple (G,S,T,H,φ) has Property $(*)$. Suppose that $\nu' \in \mathcal{M}(\mathbb{N}_0)$ is defined as in (i). Further, let $\tau_1, \tau_2 \in \mathcal{M}^\sigma(T)$ be given by $\tau_1(0) = 1$, $\tau_1(1/n) = 1$ for $n \in \mathbb{N}$ and $\tau_1(n) = 0$ for $n \geq 2$ while $\tau_2(0) = 1$, $\tau_2(n) = 1$ for $n \in \mathbb{N}$ and $\tau_1(1/n) = 0$ for $n \geq 2$. Then $\Phi(\tau_1) = \Phi(\tau_2) = \nu'$ and $\tau_1 \notin \mathcal{M}(T)$ while $\tau_2 \in \mathcal{M}(T)$. In particular $\Phi(\mathcal{M}(T)) = \mathcal{M}(\mathbb{N}_0)$ (for $\eta \in \mathcal{M}(\mathbb{N}_0)$ set $\tau_\eta(n) := \eta(n)$ for $n \in \mathbb{N}_0$ and $\tau_\eta(1/n) := 0$ for $n \geq 2$) but $\Phi^{-1}(\mathcal{M}_G(H)) = \mathcal{M}^\sigma(T)$ which is a proper superset of $\mathcal{M}(T)$.

In applications the spaces G and H are usually given, and the interest lies in the G-invariant measures on H, often motivated by simulation or integration problems. To exploit the results of this chapter one has to find suitable spaces S and T and a surjective mapping $\varphi\colon S \times T \to H$ such that (G, S, T, H, φ) has Property (∗).

Theorem 2.26 gives necessary and sufficient conditions that there exists a Borel measure $\tau_\nu \in \mathcal{M}(T)$ with $\nu = \left(\mu_{(\mathrm{S})} \otimes \tau_\nu\right)^\varphi$ for each $\nu \in \mathcal{M}_G(H)$, resp. for each $\nu \in \mathcal{M}_G^1(H)$. If $\nu \in \mathcal{M}_G(H)$ and $\tau_\nu \in \Phi^{-1}(\nu)$ then

$$\int_H f(h)\,\nu(\mathrm{d}h) = \int_T \int_S f \circ \varphi(s,t)\,\mu_{(\mathrm{S})}(\mathrm{d}s)\tau_\nu(\mathrm{d}t) \tag{2.10}$$

for all ν-integrable functions $f\colon H \to \mathbb{R}$. If f is G-invariant, too, then the right-hand side simplifies to $\int_T f_T(t)\,\tau_\nu(dt)$.

It was already pointed out in the introduction that in the great majority of the applications integrals over T can be evaluated more easily than integrals over H. Also the benefit of the symmetry concept for stochastic simulations and statistical applications has been sketched. These aspects will be extended in Chapters 3 and 4. Depending on the concrete case it may be rather complicated to determine a pre-image $\tau_\nu \in \Phi^{-1}(\nu)$ explicitly (cf. Chapter 4). For this purpose 2.26(vii) will turn out to be very useful. In fact, if $\eta \in \mathcal{M}_G(H)$ with $\eta \ll \nu$ then $\eta = f \cdot \nu$ with a G-invariant density f (2.18(ii),(iii)). Using 2.26(vii) yields an explicit expression for τ_η if such a term is known for τ_ν. We will make intensive use of Theorem 2.26(vii) in Chapter 4.

The existence problem, namely whether $\Phi^{-1}(\nu) \neq \emptyset$, is equivalent to a measure extension problem on $\mathcal{B}_0(T)$ which hopefully is easier to solve. However, depending on the concrete situation measure extension problems may also be very difficult. For continuous φ, however, the situation is much better. Below we will prove that $\Phi\left(\mathcal{M}^1(T)\right) = \mathcal{M}_G^1(H)$ for continuous φ. If additionally each $h \in H$ has a G-invariant neighbourhood U_h with relatively compact pre-image $(\varphi \circ \iota)^{-1}(U_h)$ then also $\Phi(\mathcal{M}(T)) = \mathcal{M}_G(H)$. Theorem 2.32 uses results from the field of analytical sets which are collected in Lemma 2.31. We mention that Theorem 2.51 gives an alternate proof of the same result. The proof of Theorem 2.51 uses techniques from functional analysis and it is not constructive. It can be found at the end of Section 2.3.

Definition 2.29. *Let $\sim_0$ denote an equivalence relation on the space Ω. A subset $R_0 \subseteq \Omega$ is called a* section *if it contains exactly one element of each equivalence class. A mapping* $\mathrm{sel}\colon \Omega \to \Omega$ *is called a* selection *(with respect to $\sim_0$) if* sel *is constant on the equivalence classes and if* $\mathrm{sel}(\omega)$ *is contained in its own equivalence class for each $\omega \in \Omega$. Further, we introduce the σ-algebra $\mathcal{B}(H)_F := \bigcap_{\nu \in \mathcal{M}_G^1(H)} \mathcal{B}(H)_\nu$ (cf. Example 2.20(iii)).*

Remark 2.30. If $0 \neq \nu \in \mathcal{M}_G(H)$ is finite then obviously $\nu/\nu(H) \in \mathcal{M}_G^1(H)$ and $\mathcal{B}(H)_\nu = \mathcal{B}(H)_{\nu/\nu(H)}$. If $\nu \in \mathcal{M}_G(H)$ is not finite there exist non-zero finite G-invariant Borel measures $\nu_1, \nu_2 \ldots \in \mathcal{M}_G(H)$ with $\nu = \sum_{j=1}^\infty \nu_j$

(cf. the proof of 2.26(iii)). If $E \in \bigcap_{j=1}^{\infty} \mathcal{B}(H)_{\nu_j}$ then for each $j \in \mathbb{N}$ there exist measurable sets $B_j, N_j \in \mathcal{B}(H)$ with $\nu_j(N_j) = 0$ and $E \subseteq B_j + N_j$. Let temporarily $B := \bigcup_{j=1}^{\infty} B_j$ and $N := \bigcap_{j=1}^{\infty} N_j$. Then $B, N \in \mathcal{B}(H)$ and $\nu(N) = 0$. If $h \in E$ then $h \in B$ or $h \in N$, i.e. $E \subseteq B \cup N = B + (B^c \cap N)$, and hence $E \in \mathcal{B}(H)_\nu$. This implies the inclusions $\mathcal{B}(H)_\nu \supseteq \bigcap_{j=1}^{\infty} \mathcal{B}(H)_{\nu_j} \supseteq \mathcal{B}(H)_F$, and hence each G-invariant Borel measure can be uniquely extended to $\mathcal{B}(H)_F$ (cf. Example 2.20(iii)). In particular, $\mathcal{B}(H)_F = \bigcap_{\nu \in \mathcal{M}_G(H)} \mathcal{B}(H)_\nu$. We denote this extension briefly with ν_v instead of $\nu_{v|\mathcal{B}(H)_F}$.

Lemma 2.31. *(i) Suppose that $(\Omega, \mathcal{A})$ is a measurable space and that E is a Polish space. Further, let $\pi_\Omega \colon E \times \Omega \to \Omega$ denote the projection onto the second component, and suppose that the set $N \subseteq E \times \Omega$ fulfils the following conditions:*

(α) For each $\omega \in \Omega$ the set $N(\omega) := \{e \in E \mid (e, \omega) \in N\} \subseteq E$ is closed.
(β) For each open subset $U \subseteq E$ the image $\pi_\Omega(N \cap (U \times \Omega)) \in \mathcal{A}$.

Then there exists a $(\mathcal{A}, \mathcal{B}(E))$-measurable mapping $f \colon \Omega \to E$ with $f(\omega) \in N(\omega)$ for each $\omega \in \Omega$ with $N(\omega) \neq \emptyset$.
(ii) Assume that $\sim_0$ defines an equivalence relation on the Polish space E which fulfils the following conditions:

(α) The equivalence classes (with respect to $\sim_0$) are closed subsets of E.
(β) For each open subset $U \subseteq E$ the set $\{e \in E \mid e \sim_0 u \text{ for a } u \in U\} \in \mathcal{B}(E)$.

Then there exists a measurable selection $\mathrm{sel} \colon E \to E$ *with respect to $\sim_0$.*

Proof. [17], pp. 221 (Théorème 15) and 234 (Théorème 30). □

Theorem 2.32. *Suppose that the 5-tupel (G, S, T, H, φ) has Property (*), and let $s_0 \in S$.*
(i) *Assume that $\psi \colon H \to T$ is a $(\mathcal{B}(H)_F, \mathcal{B}(T))$-measurable mapping with $\varphi(s_0, \psi(h)) \in E_G(\{h\})$ for each $h \in H$. Then for each $\nu \in \mathcal{M}_G^1(H)$ the image measure ${\nu_v}^\psi \in \mathcal{M}(T)$ is an extension of ν_*, or equivalently, $\nu = \left(\mu_{(\mathrm{S})} \otimes {\nu_v}^\psi\right)^\varphi$. In particular, $\Phi\left(\mathcal{M}^1(T)\right) = \mathcal{M}_G^1(H)$. Under the additional assumptions (2.8) also $\Phi(\mathcal{M}(T)) = \mathcal{M}_G(H)$ and $\Phi^{-1}(\mathcal{M}_G(H)) = \mathcal{M}(T)$. If the mapping $\psi \colon H \to T$ is measurable (i.e. $(\mathcal{B}(H), \mathcal{B}(T))$-measurable) then clearly ${\nu_v}^\psi = \nu^\psi$.*
(ii) *If φ is continuous then there exists a measurable mapping $\psi \colon H \to T$ with $\varphi(s_0, \psi(h)) \in E_G(\{h\})$ for each $h \in H$.*
Consequently, $\Phi(\mathcal{M}^1(T)) = \mathcal{M}_G^1(H)$ for continuous φ. If additionally each $h \in H$ has a G-invariant neighbourhood U_h with relatively compact pre-image $(\varphi \circ \iota)^{-1}(U_h) \subseteq T$ then also $\Phi(\mathcal{M}(T)) = \mathcal{M}_G(H)$ and $\Phi^{-1}(\mathcal{M}_G(H)) = \mathcal{M}(T)$.
(iii) *The G-orbits induce an equivalence relation on H. There exists a measurable selection* $\mathrm{sel} \colon H \to H$ *with respect to this equivalence relation.*
(iv) *Suppose that $\psi \colon H \to T$ is a measurable mapping with $\varphi(s_0, \psi(h)) \in E_G(\{h\})$ for each $h \in H$. Then the composition $\psi \circ \mathrm{sel}$ has the same property. Additionally, $\psi \circ \mathrm{sel}$ is constant on each G-orbit.*

Proof. For each $A \in \mathcal{B}_G(H)$ we have

$$\psi^{-1}\left((\varphi\circ\iota)^{-1}(A)\right) = \left\{h \in H \mid \psi(h) \in (\varphi\circ\iota)^{-1}(A)\right\} = A.$$

For $\nu \in \mathcal{M}^1_G(H)$ this yields ${\nu_v}^{\psi}\left((\varphi\circ\iota)^{-1}(A)\right) = \nu(A) = \nu_*\left((\varphi\circ\iota)^{-1}(A)\right)$, i.e. ${\nu_v}^{\psi}$ is an extension of ν_* to $\mathcal{B}(T)$. The second, third and forth assertion of (i) follow from the first and 2.26 (i), (iii) and (vi), resp. The final assertion of (i) is trivial. The locally compact spaces T and H are second countable. By 2.4(iii) they are metrizable and σ-compact. In particular, T and H are Polish spaces ([Qu], 149 (Satz 13.16)). To prove the first assertion of (ii) we apply Lemma 2.31(i) with $(\Omega, \mathcal{A}) = (H, \mathcal{B}(H))$, $E = T$ and $N := \{(t,h) \mid \psi(s_0,t) \in E_G(\{h\})\}$. (As G acts transitively on S and as φ is equivariant the definition of N does not depend on the particular choice of s_0.) Then

$$(t \in N(h)) \Longleftrightarrow (\varphi(S\times\{t\}) = E_G(\{h\})) \Longleftrightarrow \left(t \in (\varphi\circ\iota)^{-1}(E_G(\{h\}))\right).$$

As φ is assumed to be continuous the composition $(\varphi\circ\iota)$ is also continuous and, consequently, for each $h \in H$ the set $N(h) \subseteq T$ is the pre-image of a compact subset and hence closed. In particular, condition (α) of 2.8(i) is fulfilled. As φ is surjective $N(h) \neq \emptyset$. Let U be an open subset of T. As T is locally compact for each $t \in U$ there exist subsets U_t and K_t with $t \in U_t \subseteq K_t \subseteq U$ where U_t is an element of a fixed countable topological base and K_t is a compact neighbourhood of t. As the base is countable there exists a sequence $(t_n)_{n\in N}$ with $t_n \in U$ and $U = \bigcup_{j=1}^{\infty} U_{t_j} \subseteq \bigcup_{j=1}^{\infty} K_{t_j} \subseteq U$. Altogether, one concludes

$$\begin{aligned}\pi_H(N \cap (U\times H)) &= \{h \mid \text{there exists a } t \in U \text{ with } \varphi(S\times\{t\}) = E_G(\{h\})\}\\ &= \varphi(S\times U) = \varphi\left(S \times \bigcup_{j=1}^{\infty} K_{t_j}\right) = \bigcup_{j=1}^{\infty} \varphi(S\times K_{t_j}) \in \mathcal{B}(H)\end{aligned}$$

since each $\varphi(S\times K_{t_j})$ is a continuous image of a compact set and hence itself compact. Lemma 2.31(i) guarantees the existence of a measurable mapping $\psi\colon H \to T$ with $\psi(h) \in N(h)$ for all $h \in H$. The remaining statements of (ii) follow from (i), 2.26(iv), (v) and (vi). For the proof of (iii) we apply 2.31(ii) with $E = H$. The orbit $E_G(\{h\}) = Gh$ is compact and hence closed for each $h \in H$. If $U \subseteq H$ is open $E_G(U) = \bigcup_{g\in G} gU$ is a union of open subsets and hence itself open. In particular, it is measurable. Lemma 2.31(ii) guarantees the existence of a measurable selection $\mathrm{sel}\colon H \to H$ with respect to the G-orbits. Statement (iv) is obvious. □

The following corollary is an immediate consequence of Theorem 2.32.

Corollary 2.33. *Suppose that the* 5*-tupel* (G,S,T,H,φ) *has Property* (∗) *with continuous* φ*. Then each of the following properties implies the others:*

(α) *For* $\kappa \in \mathcal{M}^1(H, B_G(H))$ *there exists an extension* $\kappa' \in \mathcal{M}^1(H)$.
(β) *For* $\kappa \in \mathcal{M}^1(H, B_G(H))$ *there exists a* G*-invariant extension* $\kappa'' \in \mathcal{M}^1_G(H)$.
(γ) *For* $\kappa_* \in \mathcal{M}^1(T, \mathcal{B}_0(T))$ *there exists an extension* ${\kappa_*}' \in \mathcal{M}^1(T)$.

Proof. The implication '(α) $\Rightarrow$ (β)' equals Corollary 2.21. Condition (β) implies (γ) by 2.32(i),(ii) and 2.26(i). Recall that $\kappa \mapsto \kappa_*$ defines a bijection between $\mathcal{M}^1(H, \mathcal{B}_G(H))$ and $\mathcal{M}^1(T, B_0(T))$ (Lemma 2.25(vi)). Theorem 2.26(i) yields $\left(\mu_{(\mathrm{S})} \otimes {\kappa_*}'\right)^{\varphi}{}_{|\mathcal{B}_G(H)} = \kappa$ and hence (γ) implies (α). □

In 2.32(i) sufficient conditions are given that $\Phi\left(\mathcal{M}^1(T)\right) = \mathcal{M}^1_G(H)$ and $\Phi\left(\mathcal{M}(T)\right) = \mathcal{M}_G(H)$, resp. Provided that a mapping $\psi\colon H \to T$ with the demanded properties exists the surjectivity of Φ could be shown constructively. For continuous φ such a mapping does always exist (2.32(ii)).

Unlike Theorem 2.26 and Theorem 2.32 the statements 2.34(ii) and (iii) refer to single Borel measures rather than to $\mathcal{M}^1_G(H)$ and $\mathcal{M}_G(H)$, resp. On the positive side the sufficient conditions from 2.34 are less restrictive than those of 2.32, and usually they are easy to verify. Note that 2.34(iv) extends 2.32(ii) to non-continuous $\varphi\colon S \times T \to H$ (cf. Remark 2.35).

Theorem 2.34. *Suppose that the* 5*-tupel* (G, S, T, H, φ) *has Property* $(*)$. *Then*
(i) $\{(s,t) \in S \times T \mid \varphi$ *is continuous in* $(s,t)\} = S \times F_c$ *for a suitable subset* $F_c \subseteq T$.
(ii) *Let* $\nu \in \mathcal{M}_G(H)$ *be finite and suppose that there exist countably many compact subsets* $K_1, K_2, \ldots \subseteq T$ *which meet the following conditions:*

(α) *The restriction* $\varphi_{|S\times K_j}\colon S \times K_j \to H$ *is continuous for each* $j \in \mathbb{N}$.
(β) $\nu\left(H \setminus \bigcup_{j\in N} \varphi(S \times K_j)\right) = 0$.

Then there exists an measure $\tau_\nu \in \mathcal{M}(T)$ *with* $\left(\mu_{(\mathrm{S})} \otimes \tau_\nu\right)^{\varphi} = \nu$. *If, additionally,* $\Phi\left(\mathcal{M}_G(H)\right)^{-1} \subseteq \mathcal{M}(T)$ *the same is true for non-finite* $\nu \in \mathcal{M}_G(H)$.
(iii) *Let* $\nu \in \mathcal{M}_G(H)$ *be finite and suppose that there exist countably many subsets* $C_1, C_2, \ldots \in \mathcal{B}_0(T)$ *which meet the following conditions:*

(α) $\nu_*\left(T \setminus \bigcup_{j\in N} C_j\right) = 0$.
(β) *The restriction* $\varphi_{|S\times C_j}\colon S \times C_j \to H$ *is continuous for each* $j \in \mathbb{N}$.
(γ) *For each* C_j *there exist countably many compact subsets* $K_{j;1}, K_{j;2}, \ldots \subseteq T$ *with* $C_j = \bigcup_{i\in\mathbb{N}} K_{j;i}$.

Then there exists an measure $\tau_\nu \in \mathcal{M}(T)$ *with* $\left(\mu_{(\mathrm{S})} \otimes \tau_\nu\right)^{\varphi} = \nu$. *If, additionally,* $\Phi\left(\mathcal{M}_G(H)\right)^{-1} \subseteq \mathcal{M}(T)$ *the same is true for non-finite* $\nu \in \mathcal{M}_G(H)$.
(iv) *Suppose that there exist countably many compact subsets* $K_1, K_2, \ldots \subseteq T$ *which meet the following conditions.*

(α) *The restriction* $\varphi_{|S\times K_j}\colon S \times K_j \to H$ *is continuous for each* $j \in \mathbb{N}$.

(β) $T = \bigcup_{j\in N} K_j$.

Then $\Phi(\mathcal{M}^1(T)) = \mathcal{M}^1_G(H)$. *Under the additional assumptions (2.8) also* $\Phi(\mathcal{M}(T)) = \mathcal{M}_G(H)$ *and* $\Phi^{-1}(\mathcal{M}_G(H)) = \mathcal{M}(T)$.

Proof. Suppose that φ is continuous in (s,t). We have to show that for each neighbourhood V' of $\varphi(s',t)$ there exists a neighbourhood U' of (s',t) with $\varphi(U') \subseteq V'$. Let $s' = gs$. As φ is continuous in (s,t) there exists a neighbourhood U of (s,t) with $\varphi(U) \subseteq g^{-1}V'$. Then gU is a neigbourhood of (s',t) and $\varphi(gU) = g\varphi(U) \subseteq gg^{-1}V' = V'$ which proves (i). Let $A_1 := \varphi(S \times K_1)$ and inductively $A_j := \bigcup_{i\le j} \varphi(S \times K_i) \setminus \bigcup_{i<j} \varphi(S \times K_i)$. Condition (ii)$(\alpha)$ implies that the image $\varphi(S \times K_j)$ is compact. By construction, the sets $A_1, A_2 \ldots$ are disjoint and contained in $\mathcal{B}_G(H)$. In particular, $\nu_j(B) := \nu(B \cap A_j)$ for each $B \in \mathcal{B}(H)$ defines a G-invariant measure on H with total mass on $\varphi(S \times K_j)$. We point out that the 5-tuple $(G, S, K_j, \varphi(S \times K_j), \varphi_j := \varphi_{|S\times K_j})$ has Property (∗). Theorem 2.32(ii) hence guarantees the existence of a measure $\tau'_j \in \mathcal{B}(K_j)$ with $\left(\mu_{(\mathrm{S})} \otimes \tau'_j\right)^{\varphi_j} = \nu_{j\,|\mathcal{B}(\varphi(S\times K_j))}$. Clearly, $\tau_j(B) := \tau'_j(B \cap K_j)$ defines a Borel measure on T which coincides with τ'_j on $\mathcal{B}(K_j)$, and $\tau_j(K_j^c) = 0$. Consequently, $\nu_j = (\mu_{(\mathrm{S})} \otimes \tau_j)^{\varphi}$ for each $j \in \mathbb{N}$. Condition (ii)(β) implies

$$\nu(B) = \sum_{j\in\mathbb{N}} \nu(B \cap A_j) + \nu\left(B \cap \left(H \setminus \sum_{j\in\mathbb{N}} A_j\right)\right) = \sum_{j\in\mathbb{N}} \nu_j(B) + 0$$

for each $B \in \mathcal{B}(H)$ and hence $\nu = \left(\mu_{(\mathrm{S})} \otimes \tau_\nu\right)^{\varphi}$ with $\tau_\nu := \sum_{j\in\mathbb{N}} \tau_j$. If $\nu \in \mathcal{M}_G(H)$ is not finite then there exist countably many disjoint subsets $A_1, A_2, \ldots \in \mathcal{B}_G(H)$ with $H = \sum_{j\in\mathbb{N}} A_j$ and $\nu(A_j) < \infty$ for all $j \in \mathbb{N}$. Let $\nu_j \in \mathcal{M}_G(H)$ be defined by $\nu_j(B) := \nu(B \cap A_j)$ for each $B \in \mathcal{B}(H)$. The condition $\nu(H \setminus \bigcup \varphi(S \times K_j)) = 0$ clearly implies $\nu_j(H \setminus \bigcup \varphi(S \times K_j)) = 0$ for each $j \in \mathbb{N}$ and hence the conditions (α) and (β) guarantee the existence of a finite $\tau_j \in \mathcal{M}(T)$ with $\nu_j = \left(\mu_{(\mathrm{S})} \otimes \tau_j\right)^{\varphi}$. Hence $\nu = \left(\mu_{(\mathrm{S})} \otimes \tau\right)^{\varphi}$ for $\tau := \sum_{j\in\mathbb{N}} \tau_j$. The additional condition $\Phi^{-1}(\mathcal{M}_G(H)) \subseteq \mathcal{M}(T)$ implies $\tau \in \mathcal{M}(T)$ which completes the proof of (ii). As $K_{j;i} \subseteq C_j$ the restriction $\varphi_{|K_{j;i}}$ clearly is continuous. As $C_1, C_2, \ldots \in \mathcal{B}_0(T)$ condition (iii)(β) implies

$$0 = \nu_*\left(T \setminus \bigcup_{j\in\mathbb{N}} C_j\right) = \nu\left(H \setminus \bigcup_{j\in\mathbb{N}} \varphi(S \times C_j)\right) = \nu\left(H \setminus \bigcup_{j,i\in\mathbb{N}} \varphi(S \times K_{j;i})\right).$$

Consequently, the compact subsets $K_{j;i}$ meet the the conditions (ii) (α) and (β) which completes the proof of (iii). Clearly, (iv)(β) implies (ii)(β) for each $\nu \in \mathcal{M}^1_G(H)$, and hence $\Phi(\mathcal{M}(T)) = \mathcal{M}^1_G(H)$. The second assertion of (iv) is an immediate consequence from Theorem 2.26(iii) and (vi). □

Remark 2.35. (i) Theorem 2.34(iv) generalizes 2.32(ii) as there the continuity of φ implies $\Phi^{-1}(\mathcal{M}_G(H)) \subseteq \mathcal{M}(T)$ (cf. 2.26(v)). In fact, we may set $K_1 := T$

for continuous φ.
(ii) The sets $K_j \subseteq T$ and $C_j \in \mathcal{B}_0(T)$ from 2.34 need not be contained in the set F_c from 2.34(i) (cf. Example 2.38(ii)).
(iii) The conditions 2.34(iii)(α) to (γ) imply 2.34(ii)(α) and (β). Example 2.38(iv) shows that they are strictly weaker. On the positive side 2.34(iii)(α) to (γ) do not explicitly refer to the mapping φ and the measure ν. In specific situations they may be easier to verify than 2.34(ii)(α) and (β).

For statistical applications, for instance, one should be able to determine the image measures ${\nu_v}^\psi$ or ν^ψ (cf. Theorem 2.32), resp., for all admissible hypotheses, but at least for the null hypothesis.

If a mapping $\psi' \colon H \to T$ with the properties demanded in 2.32(i) exists the composition $\psi := \psi' \circ \mathrm{sel}$ has the same properties. Additionally, ψ is constant on each G-orbit (2.32(iii)). If ψ is constant on each G-orbit then the image $\psi(H) \subseteq T$ is a section with respect to the equivalence relation $\sim$ (cf. 2.25(iv)). However, this section may not be measurable which would complicate concrete computations. To determine the image measures ${\nu_v}^\psi$ or ν^ψ explicitly the mapping ψ, resp. the image $\psi(H)$, should be chosen in a suitable way. For concrete computations it hence may be favourable to determine a section $R_G \subseteq T$ first which in turn induces a mapping $\psi \colon H \to R_G \subseteq T$. Unfortunately, the induced mapping ψ may not be measurable. The following theorem illuminates the situation. It may be viewed as the counterpart of 2.32(i).

Theorem 2.36. *Suppose that the 5-tupel (G, S, T, H, φ) has Property $(*)$ and that $R_G \subseteq T$ is a measurable section with respect to the equivalence relation $\sim$ (cf. 2.25(iv)). To this section R_G there corresponds a unique mapping*

$$\psi \colon H \to R_G \subseteq T, \qquad \psi(h) := t_h \tag{2.11}$$

where $t_h \in R_G$ denotes the unique element with $\varphi \circ \iota(t_h) \in E_G(\{h\})$. Finally, let $\nu \in \mathcal{M}_G(H)$.
(i) *$\psi^{-1}(C) = \varphi\,(S \times (C \cap R_G))$ for each $C \in \mathcal{B}(T)$.*
(ii) *Suppose that $\varphi\,(S \times (C \cap R_G)) \in \mathcal{B}(H)_F$ for all $C \in \mathcal{B}(T)$. Then*

$$\tau_\nu(C) := \nu_v\,(\varphi\,(S \times (C \cap R_G))) \qquad \text{for each } C \in \mathcal{B}(T) \tag{2.12}$$

defines a σ-finite extension of ν_* to $\mathcal{B}(T)$ with $\tau_\nu\,(T \setminus R_G) = 0$. (In particular, $\nu = \left(\mu_{(S)} \otimes \tau_\nu\right)^\varphi$.) If $\tau_1 \in \mathcal{B}(T)$ is a further extension of ν_* with $\tau_1\,(T \setminus R_G) = 0$ then $\tau_1 = \tau_\nu$.
(iii) *The following properties are equivalent:*

(α) $\varphi\,(S \times (C \cap R_G)) \in \mathcal{B}(H)_F$ for all $C \in \mathcal{B}(T)$.
(β) The mapping $\psi \colon H \to T$ is $(\mathcal{B}(H)_F, \mathcal{B}(T))$-measurable.

Each of these properties implies ${\nu_v}^\psi = \tau_\nu$.
(iv) *Let φ be continuous and R_G be locally compact in the induced topology. Then $\varphi\,(S \times (C \cap R_G)) \in \mathcal{B}(H)$ for each $C \in \mathcal{B}(T)$, i.e. the mapping $\psi \colon H \to T$ is $(\mathcal{B}(H), \mathcal{B}(T))$-measurable.*

Proof. By assumption, R_G is a section with respect to '$\sim$' and φ is surjective. For each $h \in H$ there hence exists a unique element $t_h \in R_G$ with the demanded properties. In particular, the mapping ψ is well-defined. From the definition of ψ one immediately obtains

$$\psi^{-1}(\{t\}) = \{h \in H \mid h \in E_G(\{\varphi \circ \iota(t)\})\} = E_G(\{\varphi \circ \iota(t)\}) = \varphi(S \times \{t\})$$

for each $t \in R_G$. Hence $\psi^{-1}(C) = \psi^{-1}(C \cap R_G) = \varphi(S \times (C \cap R_G))$ for all $C \in \mathcal{B}(T)$ which proves (i). If $C_1, C_2, \ldots \in \mathcal{B}(T)$ are disjoint the images $\varphi(S \times (C_1 \cap R_G)), \varphi(S \times (C_2 \cap R_G)), \ldots$ are disjoint, too, since R_G is a section. Since ν_v is a measure τ_ν is σ-additive and in particular $\tau_\nu \in \mathcal{M}^+(T)$. Further, $\varphi(S \times (D \cap R_G)) = \varphi(S \times D) \in \mathcal{B}_G(H)$ for all $D \in \mathcal{B}_0(T) \subseteq \mathcal{B}_E(T)$, i.e. $\tau_\nu(D) = \nu(\varphi(S \times D)) = \nu_*(D)$ and hence $\tau_{\nu|\mathcal{B}_0(T)} = \nu_*$. As $\nu_* \in \mathcal{M}^\sigma(T, \mathcal{B}_0(T))$ (Theorem 2.26(i)) and $\tau_\nu(T \setminus R_G) = 0$ all statements of (ii) besides the uniqueness property have been shown. Clearly, any other extension τ_1 of ν_* must also be σ-finite. If R_G is equipped with the induced topology the restriction $\iota_r \colon R_G \to S \times R_G$ is continuous and $\varphi_r := \varphi_{|S \times R_G}$ is surjective. As $\mathcal{B}(T)$ is countably generated $\varphi_r^{-1}(B) = \varphi^{-1}(B) \cap (S \times R_G) \in \mathcal{B}(S \times T) \cap (S \times R_G) \in (\mathcal{B}(S) \otimes \mathcal{B}(T)) \cap (S \times R_G) = \mathcal{B}(S) \otimes \mathcal{B}(R_G)$ for all $B \in \mathcal{B}(H)$ which proves the measurability of φ_r. In particular, $\nu = \left(\mu_{(S)} \otimes \tau_{\nu|\mathcal{B}(R_G)}\right)^{\varphi_r} = \left(\mu_{(S)} \otimes \tau_{1|\mathcal{B}(R_G)}\right)^{\varphi_r}$. For the moment let $\mathcal{A}_1$ denote that σ-algebra over H which is generated by $\mathcal{B}(H)$ and $\{\varphi(S \times F) \mid F \in \mathcal{B}(R_G)\}$. As $\varphi_r^{-1}(\mathcal{B}(H)) \in \mathcal{B}(S \times R_G)$ and $\iota_r^{-1} \circ \varphi_r^{-1}(\varphi(S \times F)) = F \in \mathcal{B}(R_G)$ for all $F \in \mathcal{B}(R_G)$ the mapping $\varphi_r \colon S \times R_G \to H$ is $(\mathcal{B}(S \times R_G), \mathcal{A}_1)$-measurable. By assumption, $\mathcal{B}(H) \subseteq \mathcal{A}_1 \subseteq \mathcal{B}(H)_F \subseteq \mathcal{B}(H)_\nu$ which implies $\left(\mu_{(S)} \otimes \tau_\nu\right)^{\varphi_r}{}_v = \left(\mu_{(S)} \otimes \tau_1\right)^{\varphi_r}{}_v =: \nu_1 \in \mathcal{M}^+(H, \mathcal{A}_1)$ (cf. Example 2.20(iii)). From the definition of the image measure and since R_G is a section $\left(\mu_{(S)} \otimes \tau_\nu\right)(S \times F) = \nu_1(\varphi_r(S \times F)) = \left(\mu_{(S)} \otimes \tau_1\right)(S \times F)$ for all $F \in \mathcal{B}(R_G)$, i.e. $\tau_\nu = \tau_1$, which completes the proof of (ii). The equivalence of (iii)(α) and (iii)(β) is an immediate consequence of (i). Since

$$\begin{aligned} \nu_v{}^\psi(D) &= \nu_v\left(\psi^{-1}(D)\right) = \nu_v(\varphi(S \times (D \cap R_G))) \\ &= \nu_v(\varphi(S \times D)) = \nu(\varphi(S \times D)) = \nu_*(D) \end{aligned}$$

for all $D \in \mathcal{B}_0(T)$, the second assertion of (iii) follows from 2.26(i). If R_G is locally compact its compact subsets generate the σ-algebra $\mathcal{B}(R_G)$. (Recall that R_G is second countable and hence σ-compact.) For continuous φ the restriction φ_r is also continuous. For each compact $K_0 \subseteq R_G$ the pre-image $\psi^{-1}(K_0) = \varphi_r(S \times K_0)$ is compact, too, and hence contained in $\mathcal{B}(H)$. □

Remark 2.37. Suppose that $R_G \subseteq T$ is a measurable section and let $\psi \colon H \to T$ be defined as in Theorem 2.36. Assume further that ψ is $(\mathcal{B}(H)_F, \mathcal{B}(T))$-measurable. For each $\tau \in \mathcal{M}^1(T)$ with $\tau(R_G{}^c) = 0$ the uniqueness property from 2.36(ii) implies the identity $\tau = \left((\mu_{(S)} \otimes \tau)^\varphi\right)_v{}^\psi$. On the other hand, for $\nu \in \mathcal{M}^1_G(H)$ we have $\nu = \left(\mu_{(S)} \otimes \nu_v{}^\psi\right)^\varphi$. As $R_G \in \mathcal{B}(T)$ there is an $1-1$-correspondance between the probability measures on R_G and the probability

measures on T whose total mass is concentrated on R_G. In particular, $\nu \mapsto (\nu_v)^{\psi}{}_{|\mathcal{B}(R_G)}$ defines a bijection between $\mathcal{M}^1_G(H)$ and $\mathcal{M}^1(R_G)$.

The following example demonstrates the use and the usefulness of the theorems derived in this section. Example 2.38(i) treats the introductory example.

Example 2.38. (i) (Polar coordinates) Assume that S equals the unit circle in $\mathbb{R}^2$, $T = [0, \infty)$ and $H = \mathbb{R}^2$. Suppose further that the group $\mathrm{SO}(2)$ acts on S and H by left multiplication while

$$\varphi\colon S \times [0,\infty) \to \mathbb{R}^2, \qquad \varphi\left((\cos\alpha, \sin\alpha)^t, r\right) := (r\cos\alpha, r\sin\alpha)^t. \quad (2.13)$$

Clearly, φ is continuous and surjective, and

$$\mathbf{T}\varphi\left((\cos\alpha, \sin\alpha)^t, r\right) = \mathbf{T}(r\cos\alpha, r\sin\alpha)^t = \varphi\left(\mathbf{T}(\cos\alpha, \sin\alpha)^t, r\right) \quad (2.14)$$

for all $\mathbf{T} \in \mathrm{SO}(2)$, i.e. φ is $\mathrm{SO}(2)$-equivariant. As the remaining conditions are also fulfilled the 5-tuple $(\mathrm{SO}(2), \mathrm{S}^1, [0,\infty), \mathbb{R}^2, \varphi)$ has Property $(*)$. We choose $s_0 = (1,0) \in S$. The pre-image $(\varphi \circ \iota)^{-1}\left(\left\{\mathbf{x} \in \mathbb{R}^2 \mid \|\mathbf{x}\| \le 2\|\mathbf{y}\|\right\}\right) = [0, 2\|\mathbf{y}\|]$ is compact for each $\mathbf{y} \in \mathbb{R}^2$. In particular, as φ is continuous the assumptions (2.8) are also fulfilled. We point out that all equivalence classes on T with respect to $\sim$ are singleton. Hence $[0,\infty)$ itself is a measurable section. In particular, 2.32(ii) implies $\Phi\left(\mathcal{M}([0,\infty))\right) = \mathcal{M}_{\mathrm{SO}(2)}\left(\mathbb{R}^2\right)$, and 2.36(ii) ensures that each $\nu \in \mathcal{M}_{\mathrm{SO}(2)}(\mathbb{R}^2)$ has exactly one pre-image $\tau_\nu \in \mathcal{M}([0,\infty))$. Applying 2.36(ii) yields the pre-image τ_{λ_2}. In particular, $\tau_{\lambda_2}(C) := \lambda_2\left(\varphi(S \times C)\right)$ for all $C \in \mathcal{B}([0,\infty))$. For $0 < a < b$ this leads to

$$\tau_{\lambda_2}\left((a,b)\right) = \lambda_2\left(\left\{\mathbf{x} \in \mathbb{R}^2 \mid a < \|\mathbf{x}\| < b\right\}\right) = \pi(b^2 - a^2) = \int_a^b 2\pi t \,\mathrm{d}t, \quad (2.15)$$

i.e. τ_{λ_2} has the Lebesgue density $g_{\lambda_2}(t) := 2\pi t$. (Of course, the same result can be derived by applying the well-known transformation theorem.) Let $f(\mathbf{x}) := \mathrm{e}^{-\mathbf{x}\cdot\mathbf{x}/2}/2\pi$. Then $\eta = f \cdot \lambda_2 \in \mathcal{M}_{\mathrm{SO}(2)}(\mathbb{R}^2)$ defines a two-dimensional normal distribution with pre-image $\tau_\eta = f_T \cdot \tau_{\lambda_2} = (f_T \cdot g_{\lambda_2}) \cdot \lambda$ with $f_T(t) = (e^{-t^2/2})/2\pi$ (Theorem 2.26(vii)). This yields $f_T \cdot g_{\lambda_2}(t) = re^{-t^2/2}$. Finally, we point out that the mapping $\psi\colon \mathbb{R}^2 \to [0,\infty)$ from Theorem 2.36 is given by $\psi(\mathbf{x}) := \|\mathbf{x}\|$.
(ii) Let $G = \{e_G\}$, $S = \{s_0\}$, $T = \mathbb{R}$, $H = [0,1)$ and $\varphi(s_0, t) := t - \lfloor t \rfloor$ where $\lfloor t \rfloor$ denotes the largest integer $\le t$. As G is singleton it acts trivially on S and H, and the 5-tuple (G, S, T, H, φ) has Property $(*)$. Further, $\mathcal{B}_G(H) = \mathcal{B}(H)$, and $\mathcal{B}_0(T) = \{A + \mathbb{Z} \mid A \in \mathcal{B}_G(H)\}$, and $\mathcal{M}^1_G(H) = \mathcal{M}^1(H)$. Note that the set $F_c \subseteq T$ from 2.34(i) equals $F_c = \mathbb{R} \setminus \mathbb{Z}$. The subsets $C_1 := \mathbb{Z} \in \mathcal{B}_0(T)$ and $C_2 := \mathbb{R} \setminus \mathbb{Z} \in \mathcal{B}_0(T)$ meet the conditions 2.34(iii)(α) to (γ) with $K_{1;1} := \{-1, 0, 1\}$, $K_{1;j} := \{-j, j\}$ for $j \ge 2$ and $K_{2;j} := [-j, j] \setminus \bigcup_{-j \le i \le j}(i - 1/j, i + 1/j)$. Theorem 2.34(iii) implies $\Phi^{-1}(\nu) \neq \emptyset$ for each

finite $\nu \in \mathcal{M}_G([0,1))$. Note that we could have also applied 2.34(ii) or (iv) with the compact subsets $K_{i;j}$ from above. Clearly, in this example a preimage $\tau_\nu \in \Phi^{-1}(\nu)$ can be determined immediately. As $[0,1) \subseteq \mathbb{R}$ is a section $\tau_{\nu|[0,1)} = \nu$ and $\tau_\nu(\mathbb{R} \setminus [0,1)) = 0$. In particular, $\Phi^{-1}(\mathcal{M}_G(H)) \subseteq \mathcal{M}(T)$. Note, however, that $\Phi(\lambda) \in \mathcal{M}^+(H)$ is not a Borel measure. In particular, $\Phi(\mathcal{M}(T))$ is not contained in $\mathcal{M}_G(H)$.
(iii) Suppose that the 5-tupel (G,S,T,H,φ) has Property (∗). Assume further that G acts transitively on H. Clearly, $\mathcal{B}_G(H) = \{\emptyset, H\}$ and hence $\mathcal{B}_0(T) = \{\emptyset, T\}$. Further, 2.14(v) implies $\mathcal{M}^1(H) = \{\mu\}$ where μ denotes the unique G-invariant probability measure on H. For each $t \in T$ the set $\{t\} \subseteq T$ is a measurable section. In particular, $\Phi\left(\mathcal{M}^1(T)\right) = \mathcal{M}^1_G(H) = \{\mu\}$, i.e. each probability measure $\tau \in \mathcal{M}^1(T)$ is mapped onto μ.
(iv) Let $G = \mathbb{Z}_2 := \{-1,1\}$ (cf. 2.17(i)), $S = \mathbb{Z}_2$, $T = [0,1)$, $H = \mathbb{Z}_2$. Further, $\varphi(s,t) := 1$ iff ($s = 1$ and $t \in [0,0.5)$) or ($s = 0$ and $t \in [0.5,1)$) while $\varphi(s,t) := -1$ else. The group Z_2 acts on S and H by multiplication. It is easy to verify that the 5-tupel $(\mathbb{Z}_2, \mathbb{Z}_2, [0,1), \mathbb{Z}_2, \varphi)$ has Property (∗). As G acts transitively on H we conclude that $\mathcal{B}_0(T) = \{0,T\}$ and $\mathcal{M}^1_G(H) = \{\mu\}$ with $\mu(1) = \mu(-1) = 0.5$. The mapping φ is not continuous in $\mathbb{Z}_2 \times \{0.5\}$ and hence 2.34(iii) cannot be applied. However, the compact sets $K_0 := \{0.5\}$ and $K_j := [0, 0.5 - 1/j] \cup [0.5 + 1/j, 1 - 1/j]$ meet 2.34(iv)(α) and (β). In particular, $\Phi(\tau) = \mu$ for each $\tau \in \mathcal{M}^1(T)$.

Remark 2.39. Usually, the mapping $\psi\colon H \to T$ is not only $(\mathcal{B}(H)_F, \mathcal{B}(T))$-measurable but $(\mathcal{B}(H), \mathcal{B}(T))$-measurable. Hence the weaker measurability assumptions in 2.32 and 2.36 may look artificially at the first sight. In Section 4.3 for a class of examples a measurable section R_G is constructed for which the induced mapping $\psi\colon H \to R_G \subseteq T$ is measurable. For an arbitrary measurable section R'_G, however, the induced mapping $\psi'\colon H \to R'_G \subseteq T$ can only be proved to be $(\mathcal{B}(H)_F, \mathcal{B}(T))$-measurable. Anyway, weakening the measurability assumption does not cause additional difficulties as the extension ν_v is unique.

Statistics represents an important branch in applied mathematics. Roughly speaking, the statistician observes a sample $z_1, \ldots, z_m \in H$ which he interprets as realizations of random variables $Z_1, \ldots, Z_m$ with unknown distribution. His goal is to estimate this unknown distribution or to decide whether it belongs to the null hypothesis. If H is a high-dimensional manifold usually costly computations are necessary to determine and apply powerful statistical tests. In many cases it is hardly practically feasible to compute the distribution of the test variable $\Psi\colon H \times \cdots \times H \to [0,1]$ for the admissible hypotheses or at least for the null hypothesis, especially if the null and the alternative hypothesis are composite. In some exceptional cases the distribution of the test variable is known but it is not possible to determine the test value $\Psi(z_1, z_2, \ldots, z_m)$ explicitly (cf. [29], for example). For these reasons the well-known χ^2-test (or more precisely: the χ^2-goodness-of-fit test) is often applied. The χ^2-test merely considers the probability that the sample

assume values in particular subsets of H. However, it does not exploit any additional structure of the test problem. Consequently, this simple straightforward approach usually is not very efficient.

In the sequel we assume that the random variables $Z_1, \dots, Z_m$ are independent and that their (unknown) distributions $\nu_1, \dots, \nu_m$ are G-invariant, i.e. $\nu_1, \dots, \nu_m \in \mathcal{M}^1_G(H)$. As it does not cause any additional difficulties the probability measures $\nu_1, \dots, \nu_m$ are not assumed to be identical. The cases of most practical relevance are $\nu_1 = \cdots = \nu_m$ (one-sample problem) and $\nu_1 = \cdots = \nu_k$, $\nu_{k+1} = \cdots = \nu_m$ (two-sample problem). In the sequel we will give a formal proof that under weak additional assumptions symmetries of the observed experiment (expressed by the fact that $\nu_1, \dots, \nu_m \in \mathcal{M}^1_G(H)$) can be exploited to transform the test problem on $H \times \cdots \times H$ into a test problem on $T \times \cdots \times T$ without loss of any information. This will confirm our intuitive argumentation from the introduction. In Section 4.5 the usefulness of this transformation will be illustrated at hand of two examples where the necessary computations become considerably easier. At first, we introduce some definitions.

Definition 2.40. *Let $(V, \mathcal{V})$ be a measurable space, and let 1_{V_0} denote the indicator function of $V_0 \subseteq V$. A* test *on the measurable space $(V, \mathcal{V})$ is a measurable mapping $\Psi\colon V \to [0,1]$. The terms Γ, $\Gamma_0 \subseteq \Gamma$ and $\Gamma \setminus \Gamma_0$ denote non-empty parameter sets. In the sequel we identify Γ_0 and $\Gamma \setminus \Gamma_0$ with disjoint sets of probability measures $\mathrm{P}_{\Gamma_0}, \mathrm{P}_{\Gamma\setminus\Gamma_0} \subseteq \mathcal{M}^1(V, \mathcal{V})$, i.e. $\gamma \mapsto p_\gamma$ defines a bijection $\Gamma \to \mathrm{P}_\Gamma := \mathrm{P}_{\Gamma_0} \cup \mathrm{P}_{\Gamma\setminus\Gamma_0}$. A* test problem *on $(V, \mathcal{V})$ is given by a tupel $\big(\mathrm{P}_{\Gamma_0}, \mathrm{P}_{\Gamma\setminus\Gamma_0}\big)$. We call P_Γ the* admissible hypotheses, *P_{Γ_0} the* null hypothesis *and $\mathrm{P}_{\Gamma\setminus\Gamma_0}$ the* alternative hypothesis. *The mapping $\mathrm{pow}_\Psi\colon \mathrm{P}_\Gamma \to [0,1]$, $\mathrm{pow}_\Psi(p_\gamma) := \int_V \Psi \,\mathrm{d}p_\gamma$ is called the* power function *of Ψ. Let $(W, \mathcal{W})$ denote a measurable space. A $(\mathcal{V}, \mathcal{W})$-measurable mapping $\chi\colon V \to W$ is called a* sufficient statistic *if for each $B \in \mathcal{V}$ exists a $(\mathcal{W}, \mathcal{B}(\mathbb{R}))$-measurable mapping $h_B\colon W \to \mathbb{R}$ with $\int_V 1_B \, dp_\gamma = \int_W h_B \,\mathrm{d}p_\gamma{}^\chi$ for all $\gamma \in \Gamma$.*

A test problem $\mathrm{P}_{\Gamma_0} \cup \mathrm{P}_{\Gamma\setminus\Gamma_0}$ with sample space V is called invariant *(or, more precisely: G-*invariant*) if G acts on V and ${\mathrm{P}_{\Gamma_0}}^{\Theta_g} = \mathrm{P}_{\Gamma_0}$ and ${\mathrm{P}_{\Gamma\setminus\Gamma_0}}^{\Theta_g} = \mathrm{P}_{\Gamma\setminus\Gamma_0}$ for each $g \in G$. A mapping $\varrho\colon V \to V'$ is a* maximal invariant *if $\varrho(v_1) = \varrho(v_2)$ iff v_1 and v_2 lie in the same G-orbit.*

If the null hypothesis, resp. the alternative hypothesis, is true then the observed sample v can be interpreted as a realization of a V-valued random variable whose (unknown) distribution p_γ is contained in P_{Γ_0}, resp. in $\mathrm{P}_{\Gamma\setminus\Gamma_0}$. If $\Psi(V) \in \{0,1\}$ the test Ψ is called *deterministic*. The following example illustrates the use of the introduced definitions. Readers who are completely unfamiliar with statistics are referred to introductory works ([51, 82] etc.)

Example 2.41. (χ^2 test) Assume that (with regard to the random experiment) it is reasonable to interpret the observed sample $x_1, x_2, \dots x_m \in \mathbb{R}$ as a realization of iid random variables $X_1, \dots, X_m$ with unknown distribution η.

Further, let $\mathrm{P}_{\Gamma_0} = \{\eta_0 \otimes \cdots \otimes \eta_0\}$ and $\mathrm{P}_{\Gamma\setminus\Gamma_0} \subseteq \{\tau \otimes \cdots \otimes \tau \mid \tau \in \mathcal{M}^1(\mathbb{R})\}$. To apply a χ^2 test the real line $\mathbb{R}$ is partitioned into s disjoint intervals $I_1, \ldots, I_s$ (or more generally, into s disjoint Borel subsets). The statistician computes the test value $\mathrm{Chi} := \sum_{i=1}^{s} (\mathrm{Hfg}[i] - m \cdot \eta_0(I_i))^2 / (m \cdot \eta_0(I_i))$ where $\mathrm{Hfg}[i] := |\{j \leq m \mid x_j \in I_i\}|$. If the null hypothesis is true then Chi can be viewed as a realization of a random variable which is approximately χ^2-distributed with $s-1$ degrees of freedom. In our notation this reads as follows: $\Psi(x_1, \ldots, x_m) := 1$ if the test value Chi is larger than a particular real number $\chi_{\alpha;\, s-1}$ which depends on the significance level α and the number of intervals s. Otherwise, $\Psi(x_1, \ldots, x_m) := 0$. We reject the null hypothesis iff $\Psi(\cdot) = 1$. In particular, $\mathrm{pow}_\Psi(\eta_0 \otimes \cdots \otimes \eta_0) = \alpha$, and Ψ defines a deterministic test.

If $\chi\colon V \to W$ is a sufficient statistic then by the definition of sufficiency for each $B \in \mathcal{V}$ there exists a common factorisation $h_B\colon W \to \mathbb{R}$ for all $\gamma \in \Gamma$. Obviously, the same is true for each step function $f := \sum_{j=1}^{n} c_j 1_{B_j}$ and, consequently, for each test $\Psi\colon V \to [0,1]$ there exists a mapping $h_\Psi\colon W \to \mathbb{R}$ with $\int_V \Psi \,\mathrm{d}p_\gamma = \int_W h_\Psi \,\mathrm{d}p_\gamma{}^\chi$ for all $\gamma \in \Gamma$ as Ψ can be represented as a limit of step functions. Suppose that $\left(\mathrm{P}_{\Gamma_0}, \mathrm{P}_{\Gamma\setminus\Gamma_0}\right)$ defines a test problem on V (e.g. $V = H^m$ and $\mathcal{V} = \mathcal{B}(H^m)$) with sample $v_0 \in V$. If $\chi\colon V \to W$ is a sufficient statistic (for this test problem) then one instead may consider the transformed test problem $\left(\overline{\mathrm{P}}_{\Gamma_0}, \overline{\mathrm{P}}_{\Gamma\setminus\Gamma_0}\right)$ on W with sample $w_0 := \chi(v_0)$ and $\bar{p}_\gamma := p_\gamma{}^\chi$ without loss of information. In particular, the determination of a powerful test and the computation of its distribution under the admissible hypotheses can be transferred to the second test problem. In many cases this turns out to be much easier (cf. Chapter 4). Several examples of sufficient statistics can be found in [51] and [82], for instance. Although we will not reinforce this aspect we point out that maximal invariants are of particular relevance in statistics (for basics, see e.g. [51], pp. 285 ff.). For the moment we just mention that $\mathrm{P}_{\Gamma_0}{}^{\Theta_g} = \mathrm{P}_{\Gamma_0}$ does not necessarily mean $p_\gamma{}^{\Theta_g} = p_\gamma$ for each $p_\gamma \in \mathrm{P}_{\Gamma_0}$, i.e. p_γ need not be G-invariant. For the sake of readability we introduce some abbreviations.

Definition 2.42. *The m-fold cartesian product $D \times \cdots \times D$ of a set D will be abbreviated with D^m. Similarly, $\mathcal{V}^m := \mathcal{V} \otimes \cdots \otimes \mathcal{V}$ if $\mathcal{V}$ is a σ-algebra and $\mathcal{M}^1(V,\mathcal{V})^m := \{\tau_1 \otimes \cdots \otimes \tau_m \mid \tau_j \in \mathcal{M}^1(V,\mathcal{V})\}$. If $\tau_j = \tau$ for all $j \leq m$ then τ^m stands for $\tau_1 \otimes \cdots \otimes \tau_m$. The m-fold product $\chi \times \cdots \times \chi\colon V_1^m \to V_2^m$ of a mapping $\chi\colon V_1 \to V_2$ is denoted with χ^m.*

Lemma 2.43. *(i) If the mapping $\psi\colon H \to T$ is $(\mathcal{B}(H)_F, \mathcal{B}(T))$-measurable then $\psi^m\colon H^m \to T^m$ is $(\mathcal{C}_F := \bigcap_{\eta \in \mathcal{M}_G^1(H)^m} \mathcal{B}(H^m)_\eta, \mathcal{B}(T^m))$-measurable.*
(ii) Suppose that the assumptions from Theorem 2.36 are fulfilled and let $\nu_1, \ldots, \nu_m \in \mathcal{M}_G^1(H)$. If $\psi\colon H \to T$ is $(\mathcal{B}(H)_F, \mathcal{B}(T))$-measurable then

$$(\nu_1 \otimes \cdots \otimes \nu_m)_v{}^{\psi^m} = (\nu_1 \otimes \cdots \otimes \nu_m)_{v\,|\,\mathcal{C}_F}{}^{\psi^m} = \tau_1 \otimes \cdots \otimes \tau_m \qquad (2.16)$$

where $\tau_j \in \Phi^{-1}(\nu_j)$ denotes the unique pre-image of ν_j with $\tau_j(R_G) = 1$.

Proof. Let $C_1, \dots, C_m \in \mathcal{B}(T)$ and $\eta := \nu_1 \otimes \dots \otimes \nu_m \in \mathcal{M}_G^1(H)^m$. From the definition of ψ^m we obtain $(\psi^m)^{-1}(T \times \dots \times T \times C_k \times T \times \dots \times T) = H \times \dots \times H \times \psi^{-1}(C_k) \times H \times \dots \times H$ with $\psi^{-1}(C_k) \in \mathcal{B}(H)_F$. Hence there exist disjoint subsets $A_k, B_k \in \mathcal{B}(H)$ with $\nu_k(B_k) = 0$ and $\psi^{-1}(C_k) \subseteq A_k + B_k$. In particular, $H \times \dots \times \psi^{-1}(C_k) \times \dots \times H \subseteq H \times \dots \times A_k \times \dots \times H + H \times \dots \times B_k \times \dots \times H$ and $\eta(H \times \dots \times B_k \times \dots \times H) = \nu_1(H) \cdots \nu_k(B_k) \cdots \nu_m(H) = \nu_k(B_k) = 0$, i.e. $(\psi^m)^{-1}(T \times \dots \times C_k \times \dots \times T) \in \mathcal{B}(H^m)_\eta$. As η was arbitrary $(\psi^m)^{-1}(T \times \dots \times C_k \times \dots \times T) \in \mathcal{C}_F$. Since $\mathcal{B}(T) \times T \times \dots \times T \cup T \times \mathcal{B}(T) \times T \times \dots \times T \cup \dots \cup T \times \dots \times T \times \mathcal{B}(T)$ generates the product-σ-algebra $\mathcal{B}(T^m)$ this verifies (i). Similarly, one concludes $(\psi^m)^{-1}(C_1 \times \dots \times C_k) = \psi^{-1}(C_1) \times \dots \times \psi^{-1}(C_m) \subseteq (A_1 + B_1) \times \dots \times (A_m + B_m) = (A_1 \times \dots \times A_m) +$ $(2^m - 1)$ cartesian products of the type $D_1 \times \dots \times D_m$ with $D_j \in \{A_j, B_j\}$ where $D_{j_0} = B_{j_0}$ for at least one index $j_0 \le m$. In particular, each of these terms is a $(\nu_1 \otimes \dots \otimes \nu_m)$-zero set. As $\eta_{v|\mathcal{B}(H^m)} = \nu_1 \otimes \dots \otimes \nu_m$ Theorem 2.36(ii) implies

$$\begin{aligned}
&\eta_v\left((\psi^m)^{-1}(C_1 \times \dots \times C_k)\right) = (\nu_1 \otimes \dots \otimes \nu_m)(A_1 \times \dots \times A_m) \\
&= \nu_1(A_1) \cdots \nu_m(A_m) = (\nu_1)_v\left(\psi^{-1}(C_1)\right) \cdots (\nu_m)_v\left(\psi^{-1}(C_m)\right) \\
&= \tau_1(C_1) \cdots \tau_m(C_m).
\end{aligned}$$

Since $\mathcal{B}(T) \times \dots \times B(T)$ generates the product σ-algebra $\mathcal{B}(T)^m = \mathcal{B}(T^m)$ and since it is stable under finite intersection this implies ${\eta_{v|\mathcal{C}_F}}^{\psi^m} = \tau_1 \otimes \dots \otimes \tau_m$. □

Theorem 2.44. *Suppose that $\psi\colon H \to T$ is a $(B(H)_F, \mathcal{B}(T))$-measurable mapping with $\varphi \circ \iota \circ \psi(h) \in E_G(\{h\})$ for all $h \in H$. Then*

$$\psi^m\colon H^m \to T^m, \quad \psi^m(h_1, \dots, h_m) := (\psi(h_1), \dots, \psi(h_m)) \tag{2.17}$$

is a sufficient statistic for all test problems $\left(\mathrm{P}_{\Gamma_0}, \mathrm{P}_{\Gamma \setminus \Gamma_0}\right)$ on H^m with $\mathrm{P}_\Gamma \subseteq \mathcal{M}_G^1(H)^m$.
More precisely: For each $(\mathcal{B}(H^m), \mathcal{B}(\mathbb{R}))$-measurable function $\Psi\colon H^m \to [0,1]$ we have

$$\int_{H^m} \Psi(\mathbf{h})\, p_\gamma(\mathrm{d}\mathbf{h}) = \int_{T^m} \Psi^* \circ (\varphi \circ \iota)^m(\mathbf{t})\, p_{\gamma_v}{}^{\psi^m}(\mathrm{d}\mathbf{t}) \quad \text{for all } \gamma \in \Gamma \tag{2.18}$$

$$\text{with } \Psi^*(h_1, \dots, h_m) := \int_{G^m} \Psi(g_1 h_1, \dots, g_m h_m)\, \mathrm{d}(\mu_G)^m(g_1, \dots, g_m).$$

Proof. The product group $G^m := G \times \dots \times G$ acts componentwise on H^m via $(g_1, \dots, g_m) \,.\, (h_1, \dots, h_m) := (g_1 h_1, \dots, g_m h_m)$. Clearly, $\mathrm{P}_\Gamma \subseteq \mathcal{M}_G^1(H)^m \subseteq \mathcal{M}_{G^m}^1(H^m)$. Let $\Psi\colon H^m \to [0,1]$ be any $(\mathcal{B}(H^m), \mathcal{B}(\mathbb{R}))$-measurable mapping and let $\eta \in \mathrm{P}_\Gamma$. In particular, $\Psi \cdot \eta \in \mathcal{M}(H^m)$ is a finite measure and Theorem 2.18(iii) implies $\Psi^* \cdot \eta \in \mathcal{M}_{G^m}(H^m)$ with $\Psi^*(h_1, \dots, h_m) :=$

$\int_{G^m} \Psi(g_1 h_1, \ldots, g_m h_m)\, \mathrm{d}\mu_G^m(g_1, \ldots, g_m)$. In particular, $\int_{H^m} \Psi(\mathbf{h})\, \eta(\mathrm{d}\mathbf{h}) = (\Psi \cdot \eta)(H^m) = (\Psi^* \cdot \eta)(H^m) = \int_{H^m} \Psi^*(\mathbf{h})\, \eta(\mathrm{d}\mathbf{h})$, i.e. $\mathrm{pow}_\Psi(p_\gamma) = \mathrm{pow}_{\Psi^*}(p_\gamma)$ for all $p_\gamma \in \mathrm{P}_\Gamma$. Let temporarily $\Psi^{**}\colon T^m \to [0,1]$ be defined by $\Psi^{**} := \Psi^* \circ (\varphi \circ \iota)^m$. Since the composition $(\varphi \circ \iota)$ is $(\mathcal{B}(T), \mathcal{B}(H))$-measurable $(\varphi \circ \iota)^m$ is $(\mathcal{B}(T^m), \mathcal{B}(H^m))$-measurable and hence Ψ^{**} is $(\mathcal{B}(T^m), \mathcal{B}(\mathbb{R}))$-measurable. By assumption, $\varphi \circ \iota \circ \psi(h_j) \in E_G(\{h_j\})$ for all $h_j \in H$, and hence $(\varphi \circ \iota)^m \circ \psi^m(h_1, \ldots, h_m) \in E_G(\{h_1\}) \times \cdots \times E_G(\{h_m\})$ for all $h_1, \ldots, h_m \in H$. As Ψ^* is constant on each G-orbit $E_G(\{h_1\}) \times \cdots \times E_G(\{h_m\})$ we conclude $\Psi^* = \Psi^{**} \circ \psi^m$. The transformation formula for integrals finally yields

$$\begin{aligned}\int_{T^m} \Psi^{**}(\mathbf{t})\, (\eta_{v|\mathcal{C}_F})^{\psi^m}(\mathrm{d}\mathbf{t}) &= \int_{H^m} \Psi^*(\mathbf{h})\, \eta_{v|\mathcal{C}_F}(\mathrm{d}\mathbf{h}) = \int_{H^m} \Psi^*(\mathbf{h})\, \eta(\mathrm{d}\mathbf{h}) \\ &= \int_{H^m} \Psi(\mathbf{h})\, \eta(\mathrm{d}\mathbf{h})\end{aligned}$$

which completes the proof of Theorem 2.44. □

Remark 2.45. (i) The G^m-orbit of $(h_1, \ldots, h_m) \in H^m$ is given by the product set $E_G(\{h_1\}) \times \cdots \times E_G(\{h_m\})$.
(ii) The mapping $\psi\colon H \to T$ from Theorem 2.44 need not be constant on the G-orbits of H. If $\mathrm{sel}\colon H \to H$ denotes a measurable selection (cf. 2.32(iii)) then $\psi' := \psi \circ \mathrm{sel}$ induces a sufficient statistic which additionally is constant on each G^m-orbit. More precisely, $\psi'^m(h_1, \ldots, h_m) = \psi'^m(h'_1, \ldots, h'_m)$ iff $(h_1, \ldots, h_m)$ and $(h'_1, \ldots, h'_m)$ lie in the same G^m-orbit. In particular, ψ'^m is a *maximal invariant*.
(iii) The mapping ψ'^m transfers the test problem $\left(\mathrm{P}_{\Gamma_0}, \mathrm{P}_{\Gamma \setminus \Gamma_0}\right)$ on H^m to a test problem on $R_G^m \subseteq T^m$. In many applications the uniqueness property from Theorem 2.36(ii) facilitates the computation of ν^ψ or ${\nu_v}^\psi$, resp.
(iv) If φ is continuous a $(\mathcal{B}(H)^m, \mathcal{B}(T)^m)$-measurable sufficient statistic $\psi^m\colon H^m \to T^m$ does always exist.
(v) By Definition 2.40 the sufficient statistic $\psi^m\colon H^m \to T^m$ should be $(\mathcal{B}(H^m), \mathcal{B}(T^m))$-measurable. As each probability measure $\eta \in \mathcal{M}_G^1(H)^m$ can be uniquely extended to $\mathcal{C}_F$ this weakening of the measurability assumptions does not cause any ambiguities.
(vi) The condition $\mathrm{P}_\Gamma \subseteq \mathcal{M}_G^1(H)^m$ is more restrictive than the definition of an invariant test problem (cf. Definition 2.40).

2.3 Supplementary Expositions and an Alternate Existence Proof

First, we illuminate the impact of the various assumptions of Property (∗). It follows an excursion through the mathematical literature. We address various research subfields and sketch results on invariant measures which are related to those treated in this book. Theorem 2.51 provides an alternate proof of

2.32(ii). The proof of Theorem 2.51 does not use results from analytical sets but uses techniques from functional analysis.

After the definition of Property (∗) we have already pointed out that the topological conditions are not very restrictive and met by nearly all examples of practical relevance. In general Property (∗) is easy to verify (cf. Sections 4.3 to 4.11). Next, we briefly explain why the particular assumptions are necessary. The respective assumption is given in brackets.

As G, S, T, and H are second countable (-a-) we have $\mathcal{B}(M_1 \times M_2) = \mathcal{B}(M_1) \otimes \mathcal{B}(M_2)$ for all $M_1, M_2 \in \{G, S, T, H\}$. The spaces G, S, T and H are compact or locally compact, resp. (-b-,-c-,-d-; combined with 2.14(v)). We devote our main attention to Borel measures as they have various pleasant properties. In particular, the G-invariant Borel measures on H are uniquely determined by their values on the sub-σ-algebra $\mathcal{B}_G(H)$ (cf. Counterexample 2.19). This is essential for the characterization of $\tau_\nu \in \Phi^{-1}(\nu)$ to be an arbitrary extension of ν_* (cf. 2.26(i)). The proof of Theorem 2.51 makes intensive use of the fact that H is second countable. Moreover, Counterexample 2.52 shows that this condition is indeed indispensable. In combination with the compactness of G (-c-) this in particular ensures the existence of a disjoint decomposition of H into countably many relatively compact G-invariant Borel sets $A_1, A_2, \ldots \in \mathcal{B}(H)$ which has repeatedly been used to extend results on finite measures to arbitrary Borel measures. Due to (-d-) and (-f-) the pre-image $\varphi^{-1}(A)$ is a cartesian product $S \times \iota^{-1}(\varphi^{-1}(A)) \subseteq S \times T$ for each $A \in \mathcal{B}_G(H)$. This property is 'responsible' that $\nu \in \mathcal{M}_G(H)$ admits a representation $\nu = \left(\mu_{(S)} \otimes \tau_\nu\right)^{\dot{\varphi}}$ where the left factor equals the unique G-invariant probability measure on S. If φ was not surjective (-e-) then the complement of $\varphi(S \times T) = \varphi(gS \times T) = g\varphi(S \times T)$, i.e. $H \setminus \varphi(S \times T)$, was a non-empty G-invariant subset. Consequently, any $\nu \in \mathcal{M}_G(H)$ with positive mass on this complement could not be contained in $\Phi(\mathcal{M}(T))$. The equivariance of φ (-f-) is necessary to transform G-invariant measures into G-invariant measures.

Haar measures on locally compact groups and invariant measures on homogeneous spaces are unique up to a scalar multiple. This is a consequence of the fact that the respective group actions are transitive. In our context transitive group actions on H are of little interest as each singleton $\{t\} \subseteq T$ is a section and hence even $(G, S, \{t\}, H, \varphi_{|S\times\{t\}})$ has Property (∗). In general, group actions are not transitive. Loosely speaking, the number of G-invariant measures increases with the number of G-orbits. At the same time pleasant properties get lost. Even if G and H are differentiable manifolds and the G-action is differentiable G-invariant measures can in general not be expressed as differential forms. This limits the tools and techniques which are at disposal for the proofs, at least if the respective statements concern the entity of all G-invariant measures.

In the past a lot of the research work on invariant measures has been devoted to Haar measures and to invariant measures on homogeneous spaces. We recommend [54] as a comprehensive introductory work. Many researchers

studied symmetries in connection with harmonic analysis ([19, 35] etc.). Of particular interest for applications are classes of measures on $\mathbb{R}^n$ or on $\mathrm{Mat}(n,m)$, the vector space of real $(n \times m)$-matrices, with specific symmetry properties ([25, 26, 31, 32]). Further research work concerns the existence of measures which are invariant under certain transformations, on particular properties thereof or possible extensions ([16, 42, 55, 58, 62] etc.). A large number of current research work on invariant measures is motivated by statistical applications. Of particular significance are maximal invariants for invariant test problems and their distribution under the admissible hypotheses ([20, 21, 36, 80]). Of particular relevance for the theoretical foundations is a monograph of Wijsman ([81]) which is referenced by nearly all current publications dealing with maximal invariants or (cross) sections. In the following extensive remark its main theorems are briefly sketched and compared with the results from Section 2.2.

Remark 2.46. In [81] Wijsman first introduces the reader into the basics of differential geometry, Lie groups, group actions, Haar measures and invariant measures. In Chapters 8–13 he treats problems which are related with those considered in the present work. Its reference list contains a number of relevant papers on this topic. In the sequel we briefly sketch Wijsman's main results and compare them with our main theorems from Section 2.2. Within this remark we use the notation from [81].

In [81] a locally compact group G acts properly on a locally compact space $\mathcal{X}$. Wijsman is interested in the G-invariant and, more generally, in the relatively G-invariant measures on $\mathcal{X}$ (cf. [81], Sect. 7.3; note that the terms 'relatively invariant' and 'invariant' coincide for compact G). Loosely speaking, Wijsman is interested in spaces $\mathcal{Y}$ and $\mathcal{T}$ for which a 1–1-mapping $\varrho\colon \mathcal{X} \to \mathcal{Y} \times \mathcal{T}$ exists. More precisely, it is assumed that the homogeneous space $\mathcal{Y}$ is a copy of each orbit $Gx \subseteq \mathcal{X}$ and that there exists a bijection between $\mathcal{T}$ and the G-orbits on $\mathcal{X}$, or equivalently, a bijection between $\mathcal{T}$ and a section $\mathcal{Z} \subseteq \mathcal{X}$ with respect to the G-orbits. (In [81] $\mathcal{Z}$ is called a *(global) cross section*. The terms 'section' and 'cross section' are both common in literature.) If such spaces $\mathcal{Y}$ and $\mathcal{T}$ really exist G acts on $\mathcal{Y}$ by left multiplication as the homogeneous space $\mathcal{Y}$ is isomorphic to the coset G/G_0 for a particular subgroup G_0 of G, and $\varrho^{-1}(gy,t) = g\varrho^{-1}(y,t)$ for all $(g,y,t) \in G \times \mathcal{Y} \times \mathcal{T}$, i.e. ϱ^{-1} is G-equivariant. Under suitable assumptions ϱ maps (relatively) invariant Borel measures on $\mathcal{X}$ to product measures on $\mathcal{Y} \times \mathcal{T}$ where the left factor is a (relatively) invariant measure on $\mathcal{Y}$. In principal, it is sufficient that ϱ is bi-measurable. Relative to [2, 10], for example, Wijsman demands stronger conditions which in turn yield more concrete results. Under very special assumptions a section $\mathcal{Z}$ can be determined explicitly.

Theorem 8.6 is the first main result in [81]. It is assumed that the group G, the spaces $\mathcal{Y}$, $\mathcal{T}$, $\mathcal{X}$ and the cross section $\mathcal{Z}$ are differentiable manifolds. Theorem 8.6 provides properties of the images of relatively G-invariant measures on $\mathcal{X}$ under the mapping ϱ. (These image measures are particular product

measures.) Essential for the existence of a mapping ϱ^{-1} (denoted with φ in [81]) with the desired properties is that all $x \in \mathcal{Z}$ have the same isotropy group and that the Jacobian of ϱ^{-1} is positive on a particular subset of $\mathcal{Y} \times \mathcal{T}$ (Assumption 8.2). Generically, the content of Theorem 8.6 is to be expected as ϱ is a diffeomorphism. In the second half of Chapter 8 Wijsman considers an even more specific situation. He additionally assumes that there exists a Lie group $K = GH$ which acts transitively on $\mathcal{X}$ and that G and H are closed subgroups of K. Under additional assumptions Theorem 8.12 gives elegant explicit expressions for $\mathcal{Y}$, $\mathcal{Z}$, and $\mathcal{T}$. In particular, the cross section $\mathcal{Z}$ can be chosen as a coset of the subgroup H. Unfortunately, the assumptions of Theorems 8.6 and 8.12 are very restrictive, not so much because the spaces are assumed to be differentiable manifolds but since all $x \in \mathcal{Z}$ are assumed to have the same isotropy group. (For Theorem 8.6 this is demanded explicitly while for Theorem 8.12 this is an immediate consequence from 8.11 (i) and (ii): In fact, for $g_0 \in G_{x_0}$, the isotropy group of x_0, we conclude $g_0(hx_0) = h(h^{-1}g_0h)(x_0) = hx_0$.) In particular, this implies that the isotropy group of all $x \in \mathcal{X}$ are conjugate. This is definitely not the case for the examples discussed in Sections 4.3, 4.4, 4.5, 4.9, 4.10 and Example 4.108 in Section 4.11 below. Apart from Section 4.10 the orbits of the identity element or the identity matrix, resp., are singleton, and its isotropy group hence equals G. In fact, in these examples (apart from Section 4.3) the structure of the isotropy groups depend on the multiplicity of the eigenvalues, resp. the singular values, of the respective matrices. We point out that also the theorems from [10] (which resign on differentiability assumptions), for instance, cannot be applied. More precisely, bi-measurable mappings $\varrho^{-1}\colon \mathcal{Y} \times \mathcal{T} \to \mathcal{X}$ with the desired properties do not exist for these examples.

We remark that Theorem 8.6 of [81] can be applied to the examples presented in Sections 4.6 (Iwasawa decomposition), 4.7 (QR-decomposition) and 4.8 (polar decomposition). In Sections 4.7 and 4.8 $G = \mathrm{O}(n)$ acts on $\mathcal{X} = \mathrm{GL}(n)$ by left multiplication. For each $\boldsymbol{M} \in \mathrm{GL}(n)$ the isotropy group equals the identity matrix. Cross sections are given by $\mathcal{Z} = V$, $\mathcal{Z} = \mathrm{R}(n)$ and $\mathcal{Z} = \mathrm{Pos}(n)$, resp., where V is a submanifold of H while $\mathrm{R}(n)$ and $\mathrm{Pos}(n)$ denote the group of all upper triangular matrices with positive diagonal elements or the set of all symmetric positive definite real matrices, resp. Note that $\mathrm{R}(n)$ is closed in the induced topology of $\mathrm{GL}(n)$. We point out that Theorem 8.12 of [81] can be applied to the QR-decomposition with $G = \mathrm{O}(n)$, $H = \mathrm{R}(n)$ and $x_0 = 1_n$.

However, it should be stressed that both Theorem 8.6 and Theorem 8.12 may be very useful for concrete statistical applications. Of particular interest are distributions of invariant statistics and, more generally, distributions of random variables which are derived therefrom. In many applications, the space $\mathcal{X}$ equals the $\mathbb{R}^n$ or $\mathrm{Mat}(n, m)$ or subsets thereof whereas the measures on $\mathcal{X}$ are absolutely continuous with respect to the Lebesgue measure. Then Lebesgue zero-subsets of $\mathcal{X}$ can be neglected. That is, it suffices to find a

mapping $\varrho' : \mathcal{X} \setminus \mathcal{N} \to \mathcal{Y} \times \mathcal{T}$ with the desired properties where $\mathcal{N} \subseteq \mathcal{X}$ denotes a suitable G-invariant Lebesgue zero-set. In his Examples 8.1 and 8.7 Wijsman considers $\mathrm{O}(n)$-invariant measures on $\mathrm{Pos}(n)$. Unlike in Section 4.9 below he excludes the matrices with multiple eigenvalues, and consequently, Theorem 8.6 from [81] can be applied then. However, it does not cover G-invariant measures on $\mathrm{Pos}(n)$ with positive mass on the subset of all matrices with multiple eigenvalues.

In Chapter 12 Wijsman compares his approach from Chapter 8 (denoted as 'W'-method) with the so-called 'ABJ'-method from [3]. Again, it is assumed that a locally compact group G acts properly on a locally compact space $\mathcal{X}$. The 'ABJ' -method assumes that there exists a homogeneous space $\mathcal{Y}$ and a continuous G-equivariant mapping $u : \mathcal{X} \to \mathcal{Y}$. The mapping u is assumed to be a statistic of interest. If η denotes a relatively G-invariant Borel measure on $\mathcal{X}$ and $\pi : \mathcal{X} \to \mathcal{X}/G$ the projection onto the orbit space then the image measure $\eta^{u \times \pi}$ is a product measure on $\mathcal{Y} \times \mathcal{X}/G$. (Loosely speaking, we obtain the more information on η the 'larger' $\mathcal{Y}$ is, i.e. the 'closer' $u \times \pi$ is at a 1-1-mapping.) As already pointed out in the Sections 4.3, 4.4, 4.5, 4.9, and in Example 4.108 below the orbits of the identity element, resp., the identity matrix are singleton. Consequently, the homogeneous space $\mathcal{Y}$ must be singleton in these cases since u was assumed to be equivariant and hence surjective. Wijsman explains ([81], p. 196) that each method can be derived from the other.

The space T and the equivariant mapping $\varphi : S \times T \to H$ in (G, S, T, H, φ) remind on the space $\mathcal{T}$ and the mapping ϱ^{-1} from [81]. However, unlike for $\mathcal{T}$ there usually does not exist a bijection between T and the G-orbits on H; and even then φ need not be 1–1.

The property of a G-invariant measure ν to be an image of a product measure under a surjective mapping $\varphi : S \times T \to H$ is obviously weaker than that in [81] where the corresponding mapping $\varrho^{-1} : \mathcal{Y} \times \mathcal{T} \to \mathcal{X}$ is a diffeomorphism. As a first consequence, unlike its pendant in [81] the measure τ_ν usually is not unique. Note that the mapping ϱ^{-1} is also G-equivariant. If G is compact and if the spaces $G, \mathcal{Y}, \mathcal{T}$ and $\mathcal{X}$ are second countable the 5-tuple $(G, \mathcal{Y}, \mathcal{T}, \mathcal{X}, \varrho^{-1})$ has Property $(*)$. It represents a special case as ϱ^{-1} is bijective. In particular, assertion (8.10) in Theorem 8.6 ([81]) follows immediately from 2.32(ii). As ϱ^{-1} is invertible 2.18(iii) and 2.26(vii) imply statement (8.12) in [81]. That is, Theorem 8.6 does not give additional information for compact G. However, there is no pendant to Theorem 8.12 in this book.

Clearly, the less restrictive assumptions on T and $\varphi : S \times T \to H$ make the respective proofs more complicated but at the same time they open up new applications. In particular, the space T can be chosen independently of the orbit structure on H. Clearly, subsets of $\mathbb{R}^n$ or $\mathrm{Mat}(n, m)$ which can simply be characterized are particularly suitable for concrete computations. For integration and simulation problems it may sometimes be more advisable to chose T than a section $R_G \subseteq T$.

In Chapters 9–11 and 13 Wijsman uses his theorems to derive explicit expressions for a number of transformed measures which are relevant in applied statistics. In Chapter 11 the assumptions of Theorem 8.12 are somewhat relaxed which in turn yields weaker results. However, also the weakened assumptions do not seem to open up further applications. In Chapter 13 the assumptions on the section $\mathcal{Z}$ are weakened, namely a global condition is replaced by a local one. For particular applications (e.g. to determine density ratios) this is sufficient. However, this does not open up further applications in Chapter 4 below.

After this excursion we come to the alternate existence proof. For continuous φ Theorem 2.32(ii) guarantees $\Phi(\mathcal{M}^1(T)) = \mathcal{M}^1_G(H)$, and if each $h \in H$ has a G-invariant neighbourhood U_h with relatively compact preimage $(\varphi \circ \iota)^{-1}(U_h)$ then also $\Phi(\mathcal{M}(T)) = \mathcal{M}_G(H)$. Whereas the theorems from Section 2.2 use techniques and results from set-based measure theory and analytical sets the proof of Theorem 2.51 uses techniques from functional analysis. At first, we introduce some definitions.

Definition 2.47. *Let M denote a locally compact space. The* support *of a continuous function $f\colon M \to \mathbb{R}$, denoted by* $\operatorname{supp} f$*, is the closure of the subset $\{m \in M \mid f(m) \neq 0\} \subseteq M$. Further, $C_c(M) := \{f\colon M \to \mathbb{R} \mid f$ is a continuous function with compact support$\}$ and, analogously, $C_{c,G}(M) := \{f \in C_c(M) \mid f$ is G-invariant$\}$.*
A measure $\eta \in \mathcal{M}^+(M)$ is called inner regular *if $\eta(B) = \sup\{\eta(K) \mid K \subseteq B$, K is compact$\}$ for all $B \in \mathcal{B}(M)$. It is called* outer regular *if $\eta(B) = \inf\{\eta(U) \mid U \supseteq B$, U is open$\}$ for all $B \in \mathcal{B}(M)$. A measure $\eta \in \mathcal{M}^+(M)$ is* regular *if it is both, inner and outer regular. The measure η is a* Radon measure *if it is inner regular and if each $m \in M$ has a neighbourhood U_m with $\eta(U_m) < \infty$. The* vague topology *is the coarsest (smallest) topology on the set of all Radon measures on M for which the mapping $\eta \mapsto \int f(m)\,\eta(\mathrm{d}m)$ is continuous for all $f \in C_c(M)$. The support of Radon measure η on M, denoted with* $\operatorname{supp}\eta$*, is given by $M \setminus \bigcup_{U \text{ open},\, \eta(U)=0} U$.*

Lemma 2.48. *Suppose that M is a second countable locally compact space. Then*
(i) Each Borel measure $\eta \in \mathcal{M}(M)$ is regular. In particular, $\eta \in \mathcal{M}(M)$ iff η is a Radon measure.
(ii) The vague topology on $\eta \in \mathcal{M}(M)$ is generated by the sets

$$V_{f_1,\ldots,f_n;\varepsilon}(\eta) := \tag{2.19}$$
$$\left\{\eta' \in \mathcal{M}(M)\colon \left|\int_M f_j(m)\,\eta(\mathrm{d}m) - \int_M f_j(m)\,\eta'(\mathrm{d}m)\right| < \varepsilon \text{ for all } j \le n\right\},$$
$$\eta \in \mathcal{M}(M),\ n \in \mathbb{N},\ f_1,\ldots,f_n \in C_c(M),\ \varepsilon > 0.$$

Proof. [5], pp. 213 (Satz 29.12) and 221. □

Remark 2.49. Assume that a compact group G acts on a second countable locally compact space M.
(i) For any $\eta \in \mathcal{M}(M)$ the union U_0 of all open η-zero sets is itself an η-zero set. More precisely, U_0 is the maximal open subset of M with $\eta(U_0) = 0$, and $\operatorname{supp}\eta = M \setminus U_0$.
(ii) The group G acts transitively on each orbit $Gm \subseteq M$. By 2.14(v) the orbit Gm is a homogeneous space, and there exists a unique G-invariant probability measure $\kappa' \in \mathcal{M}^1_G(Gm)$. Note that $\kappa'(U) > 0$ for each non-empty open subset $U \subseteq Gm$ (in the induced topology) since the compactness of Gm and the invariance of κ' would imply $\kappa'(Gm) = 0$ else. In particular, $\operatorname{supp}\kappa' = Gm$. There exists a unique probability measure $\kappa_{(Gm)}$ on M with $\kappa_{(Gm)}(A) = \kappa'(A)$ for $A \in \mathcal{B}(Gm)$ and $\kappa_{(Gm)}(M \setminus Gm) = 0$. In particular, $\kappa_{(Gm)} \in \mathcal{M}^1_G(M)$ and $\operatorname{supp}\kappa_{(Gm)} = Gm$ by the definition of the induced topology.

Lemma 2.50. *Suppose that M is a second countable locally compact space. Then*
(i) Let $f \in C_c(M)$ and let f^* be defined as in 2.18(iii), i.e. $f^*(m) := \int_G f(gm)\,\mu_G(\mathrm{d}g)$. Then $f^* \in C_c(M)$, and $\operatorname{supp} f^*$ is G-invariant.
(ii) For $\eta \in \mathcal{M}(M)$ the following statements are equivalent:

(α) $\eta \in \mathcal{M}_G(M)$.
(β) $\int_M f(m)\,\eta(\mathrm{d}m) = \int_M f(gm)\,\eta(\mathrm{d}m)$ for all $(g,f) \in G \times C_c(M)$.

In particular, $\operatorname{supp}\nu$ is G-invariant for each $\nu \in \mathcal{M}_G(M)$, and

$$\int_M f(m)\,\nu(\mathrm{d}m) = \int_M f^*(m)\,\nu(\mathrm{d}m) \qquad \text{for all } f \in C_c(M). \quad (2.20)$$

(iii) Equipped with the vague topology $\mathcal{M}(M)$ is a second countable Hausdorff space. The set $\mathcal{M}_G(M)$ is a closed in $\mathcal{M}(M)$. The subset $\mathcal{M}^{\le 1}(M) := \{\eta \in \mathcal{M}(M) \mid \eta(M) \le 1\}$ is vague compact, i.e. compact in the vague topology.
(iv) The induced topology on $\mathcal{M}^{\le 1}_G(M) := \mathcal{M}_G(M) \cap \mathcal{M}^{\le 1}(M)$ is generated by the sets

$$V^{G;\le 1}_{f_1,\dots,f_n;\varepsilon}(\nu) := \quad (2.21)$$
$$\left\{\nu' \in \mathcal{M}^{\le 1}_G(M)\colon \left|\int_M f_j(m)\,\nu(\mathrm{d}m) - \int_M f_j(m)\,\nu'(\mathrm{d}m)\right| < \varepsilon \text{ for all } j \le n\right\},$$
$$\nu \in \mathcal{M}^{\le 1}(M),\ n \in \mathbb{N},\ f_1,\dots,f_n \in C_{c,G}(M),\ \varepsilon > 0.$$

(v) The set

$$\mathcal{R} := \left\{\eta \mid \eta := \sum_{j=1}^{k} a_j \cdot \kappa_{(Gm_j)};\ a_j \ge 0;\ a_1 + \cdots + a_k \le 1;\ k \in \mathbb{N}\right\} \quad (2.22)$$

is dense in $\mathcal{M}^{\le 1}_G(M)$.

Proof. As M is second countable M is metrizable (2.4(iii)). Assume $f \in C_c(M)$ and define $F\colon M\times G \to \mathbb{R}$, $F(m,g) := f\circ\Theta(g,m) = f(gm)$ for the moment. Clearly, $F = f\circ\Theta$ is continuous and bounded since $f \in C_c(M)$. (Note that $\operatorname{supp} F \subseteq \operatorname{supp} f \times G$.) In particular, the mapping $G \to G$, $g \mapsto F(m,g)$ is μ_G-integrable for each $m \in M$ and the mapping $M \to M$, $F(\cdot,g)$ is continuous for all $g \in G$. The function $r\colon G \to \mathbb{R}$, $r \equiv \sup_{(m,g)\in M\times G} |F(m,g)|$ is bounded and hence μ_G-integrable. Applying the continuity lemma 16.1 in [5], p. 101, with $E = M$, $(\Omega,\mathcal{A},\mu) = (G,\mathcal{B}(G),\mu_G)$, $f = F$ and $h = r$ proves the continuity of $f^* = \int_G F\,d\mu_G$. The union $G\operatorname{supp} f = \Theta(G,\operatorname{supp} f)$ is a G-invariant compact set. Hence its complement $M \setminus G\operatorname{supp} f$ is open and also G-invariant. In particular, $f(gm) = 0$ for all $(g,m) \in G\times(M\setminus G\operatorname{supp} f)$ and hence $f^* \equiv 0$ on this complement. This yields $\operatorname{supp} f^* \subseteq G\operatorname{supp} f$ which proves $f^* \in C_c(M)$. Let $m \in \operatorname{supp} f^*$ and let U' be a neighbourhood of gm. By definition, there exists a $m'' \in g^{-1}U'$ with $f^*(m'') \neq 0$. In particular, $gm'' \in U'$ and the G-invariance of f^* implies $f^*(gm'') \neq 0$. Consequently, $(m \in \operatorname{supp} f^*)$ implies $(gm \in \operatorname{supp} f^*)$. Changing the roles of g and gm proves the inverse direction, and this completes the proof of (i).
Let $\eta \in \mathcal{M}_G(M)$. By definition we obtain $\int_M f(gm)\,\eta(\mathrm{d}m) = \int_M f(m)\,\eta(\mathrm{d}m)$ for all indicator functions $f\colon M \to \mathbb{R}$ and hence for all linear combinations of indicator functions (= step functions). Since each measurable non-negative numerical function can be represented as a limit of monotonously non-decreasing step functions this equation is also valid for all measurable non-negative numerical functions and hence for all $f \in C_c(M)$ as f can be split into its positive part $f^+ := \max\{f,0\}$ and its negative part $f^- := \max\{-f,0\}$, i.e. $f = f^+ - f^-$. This proves the implication $(\alpha) \Longrightarrow (\beta)$.
Now assume that $\eta \in \mathcal{M}(M)$ fulfils (β). Further, let $K \subseteq M$ denote a compact subset and $g \in G$. Since η is regular (2.48(i)) we have $\eta(K) = \inf\{\eta(U) \mid U$ is an open neighbourhood of $K\}$. Assume that $\eta(K) < \eta(gK)$. Then there exists an open neighbourhood U_0 of K with $\eta(U_0) < \eta(gK)$. As M is locally compact there exists a relatively compact open neighbourhood U_1 of K. Set $U_2 := U_0 \cap U_1$. Due to Urysohn's Lemma ([5], p. 193 (Korollar 27.3)) there exists a continuous function $f'\colon M \to \mathbb{R}$ with $0 \le f' \le 1$, $f'_{|K} \equiv 1$ and $f_{|U_2^c} \equiv 0$. In particular, $f' \in C_c(M)$. This leads to a contradiction as $\eta(gK) \le \int f'(g^{-1}m)\,\eta(\mathrm{d}m) = \int f'(m)\,\eta(\mathrm{d}m) \le \eta(U_2) < \eta(gK)$. Similarly, also the assumption $\eta(K) > \eta(gK)$ leads to a contradiction. That is, $\eta(gK) = \eta(K)$ for each compact K. Or equivalently: η und η^{Θ_g} coincide on the compact subsets of M where $\Theta_g\colon M \to M$, $\Theta_g(m) := gm$. As M is σ-compact the compact subsets are a generating system of $\mathcal{B}(M)$ which is stable under intersection, and M can be expressed as the countable union of non-decreasing compact subsets. Satz 1.4.10 from [28], p. 28, implies $\eta = \eta^{\Theta_g}$. As g was arbitrary this implies $\eta \in \mathcal{M}_G(M)$. Let $m \in \operatorname{supp} f^*$ and U' be a neighbourhood of gm. By definition $\nu\left(g^{-1}U'\right) > 0$, and the G-invariance of ν implies $\nu(U') > 0$. As U' was arbitrary $gm \in \operatorname{supp} f^*$. Changing the roles of m and gm we obtain the equivalence $(m \in \operatorname{supp}\nu) \Longleftrightarrow (gm \in \operatorname{supp}\nu)$

which implies the G-equivariance of $\operatorname{supp}\nu$. The final assertion of (ii) is an immediate consequence of the Theorem of Fubini

$$\begin{aligned}&\textstyle\int_M f^*(m)\,\nu(\mathrm{d}m) = \int_M \left(\int_G f(gm)\,\mu_G(\mathrm{d}g)\right)\,\nu(\mathrm{d}m)\\ &\textstyle= \int_G \left(\int_M f(gm)\,\nu(\mathrm{d}m)\right)\,\mu_G(\mathrm{d}g) = \int_M f(m)\,\nu(\mathrm{d}m).\end{aligned}$$

The first and the third assertion of (iii), i.e. the assertions on $\mathcal{M}(M)$ und $\mathcal{M}^{\leq 1}(M)$, are proved in [5], pp. 240 (Satz 31.5) and 237 (Korollar 31.3). (Note that in [5] the term $\mathcal{M}_+(M)$ denotes the Radon measures on M. As M is second countable a measure η on M is a Radon measure iff it is a Borel measure (cf. 2.48(i)).) Let $\eta \in \mathcal{M}(M)\setminus\mathcal{M}_G(M)$. Due to (ii) there exists a pair $(g,f) \in G \times C_c(M)$ with $d := \left|\int_M f(gm)\,\eta(\mathrm{d}m) - \int_M f(m)\,\eta(\mathrm{d}m)\right| > 0$. The triangle inequality yields $V_{f,f\circ\Theta_g;(d/10)}(\eta) \cap \mathcal{M}_G(M) = \emptyset$. As $\eta \in \mathcal{M}(M)\setminus\mathcal{M}_G(M)$ was arbitrary $M(M)\setminus\mathcal{M}_G(M)$ is open which finishes the proof of (iii).
Let $V_{f_1,\ldots,f_n;\varepsilon}(\eta) \subseteq \mathcal{M}(M)$ and $\nu_0 \in \mathcal{M}_G^{\leq 1}(M) \cap V_{f_1,\ldots,f_n;\varepsilon}(\eta)$. There exists an $\varepsilon_0 < \varepsilon$ such that $\left|\int_M f_j\,d\eta - \int_M f_j\,d\nu_0\right| < \varepsilon_0$ for all $j \leq n$, and from the triangle inequality we obtain $V_{f_1,\ldots,f_n;(\varepsilon-\varepsilon_0)}(\nu_0) \subseteq V_{f_1,\ldots,f_n;\varepsilon}(\eta)$. Hence the induced topology on $\mathcal{M}_G^{\leq 1}(M)$ is generated by the sets $V_{f_1,\ldots,f_n;\varepsilon}(\nu)\cap\mathcal{M}_G^{\leq 1}(M)$ with $\nu \in \mathcal{M}_G^{\leq 1}(M)$, $n \in \mathbb{N}$, $f_j \in C_c(M)$ and $\varepsilon > 0$. Due to (ii) we have $\int_M f_j\,\mathrm{d}\nu' = \int_M f_j^*\,\mathrm{d}\nu'$ for all $\nu' \in \mathcal{M}_G(M)$ which proves (iv).
To prove (v) we modify the proof of Satz 30.4 in [5], pp. 223f. Assume $V^{G;\leq 1}_{f_1,\ldots,f_n;\varepsilon}(\nu) \subseteq \mathcal{M}_G^{\leq 1}(M)$. We have to verify that $V^{G;\leq 1}_{f_1,\ldots,f_n;\varepsilon}(\nu) \cap \mathcal{R} \neq \emptyset$. From (i) we conclude that the union $A := \bigcup_{j=1}^n \operatorname{supp} f_j$ is a compact G-invariant subset of M. Hence there exist finitely many open subsets $U_1,\ldots,U_k$ with $A \subseteq \bigcup_{i=1}^k U_i$ and $|f_j(m) - f_j(m')| < \varepsilon$ for all $m,m' \in U_i$ and all $j \leq n$. The set $W_i := GU_i = \bigcup_{g\in G} gU_i$ is a union of open subsets and hence itself open. In particular, for $m_1,m_2 \in W_i$ there exist $g_1,g_2 \in G$ with $g_1m_1, g_2m_2 \in U_i$, and the G-invariance of the functions f_j implies $|f_j(m_1) - f_j(m_2)| = |f_j(g_1m_1) - f_j(g_2m_2)| < \varepsilon$. Next, we define the sets $A_1 := W_1 \cap A$ and $A_s := (W_s \cap A)\setminus\bigcup_{i=1}^{s-1} A_i$ for $2 \leq s \leq k$. In particular, $A_i \in \mathcal{B}_G(M)$, and $A = \sum_{i=1}^k A_i$ is a disjoint decomposition of A into finitely many relative compact subsets. Without loss of generality we may assume that the A_j are non-empty. (Otherwise we cancel the empty sets and relabel the remaining non-empty sets.) For each $i \leq k$ we choose any $m_i \in A_i$, and set $\kappa := \sum_{i=1}^k \nu(A_i)\cdot\kappa_{(Gm_i)}$. Clearly, $\kappa(M) = \sum_{i=1}^k \nu(A_i) = \nu(M) \leq 1$, i.e. $\kappa \in \mathcal{R}$. Since the functions f_j and the sets A_i are G-invariant for each $j \leq n$ we obtain the inequation

$$\begin{aligned}&\left|\int_M f_j(m)\,\nu(\mathrm{d}m) - \int_M f_j(m)\,\kappa(\mathrm{d}m)\right|\\ &= \left|\sum_{i=1}^k \left(\int_{A_i} f_j(m)\,\nu(\mathrm{d}m) - \underbrace{\int_{A_i} f_j(m)\,\kappa(\mathrm{d}m)}_{\nu(A_i)f_j(m_i)}\right)\right|\end{aligned}$$

$$\leq \sum_{i=1}^{k} \int_{A_i} |f_j(m) - f_j(m_i)| \, \nu(\mathrm{d}m) < \sum_{i=1}^{k} \nu(A_j)\varepsilon = \nu(M)\varepsilon \leq \varepsilon.$$

This completes the proof of this Lemma. □

Theorem 2.51. *Suppose that the 5-tuple (G, S, T, H, φ) has Property (∗) and that φ is continuous. As in Section 2.2 the mapping $\Phi\colon \mathcal{M}^\sigma(T) \to \mathcal{M}_G^+(H)$ is defined by $\Phi(\tau) := (\mu_{(\mathrm{S})} \otimes \tau)^\varphi$. Then*
(i) *$\Phi(\mathcal{M}^1(T)) = \mathcal{M}_G^1(H)$.*
(ii) *If each $h \in H$ has a G-invariant neighbourhood U_h with relatively compact pre-image $(\varphi \circ \iota)^{-1}(U_h) \subseteq T$ then also $\Phi(\mathcal{M}(T)) = \mathcal{M}_G(H)$ and $\Phi^{-1}(\mathcal{M}_G(H)) = \mathcal{M}(T)$.*

Proof. We begin with the proof of assertion (ii). For each compact $K \subseteq H$ there exist $h_1, \ldots, h_k \in K$ such that $K \subseteq \bigcup_{i=1}^k U_{h_i}$, and the inclusion $(\varphi \circ \iota)^{-1}(K) \subseteq \bigcup_{i=1}^k (\varphi \circ \iota)^{-1}(U_{h_i})$ proves the relative compactness of $(\varphi \circ \iota)^{-1}(K)$. As $\varphi \circ \iota$ is continuous the pre-image $(\varphi \circ \iota)^{-1}(K)$ is closed and hence compact. In particular, the pre-image $(\varphi \circ \iota)^{-1}(\operatorname{supp} f)$ is compact for each $f \in C_c(H)$. Clearly, $(f \circ \varphi \circ \iota)$ is continuous, and for $t \in T \setminus (\varphi \circ \iota)^{-1}(\operatorname{supp} f)$ we have $f \circ \varphi \circ \iota(t) = 0$. As $T \setminus (\varphi \circ \iota)^{-1}(\operatorname{supp} f)$ is open $\operatorname{supp}(f \circ \varphi \circ \iota) \subseteq (\varphi \circ \iota)^{-1}(\operatorname{supp} f)$ and hence $f \circ \varphi \circ \iota \in C_c(T)$. From 2.26(iv) we conclude $\Phi(\mathcal{M}(T)) \subseteq \mathcal{M}_G(H)$ and hence the restriction $\Phi_{|\mathcal{M}(T)}\colon \mathcal{M}(T) \to \mathcal{M}_G(H)$, $\Phi(\tau) := \big(\mu_{(\mathrm{S})} \otimes \tau\big)^\varphi$ is well-defined. (We consider the restriction $\Phi_{|\mathcal{M}(T)}$ in the following that we can apply the preceding lemma.) First, we prove that $\Phi_{|\mathcal{M}(T)}$ is continuous. As $\mathcal{M}(T)$ is second countable (2.50(iii)) the continuity of any mapping $\chi\colon \mathcal{M}(T) \to M'$ can be characterized using convergent sequences ([59], p. 54). If $\lim_{n\to\infty} \tau_n = \tau \in \mathcal{M}(T)$ then

$$\begin{aligned}
&\int_H f(h)\, \Phi(\tau_n)(\mathrm{d}h) = \int_H f^*(h)\, \Phi(\tau_n)(\mathrm{d}h) \\
&= \int_{S\times T} f^* \circ \varphi(s,t)\, \big(\mu_{(\mathrm{S})} \otimes \tau_n\big)(\mathrm{d}s, \mathrm{d}t) = \int_T \left(\int_S f^* \circ \varphi \circ \iota(t)\, \mu_{(\mathrm{S})}(\mathrm{d}s) \right) \tau_n(\mathrm{d}t) \\
&= \int_T f^* \circ \varphi \circ \iota(t)\, \tau_n(\mathrm{d}t) \longrightarrow_{n\to\infty} \int_T f^* \circ \varphi \circ \iota(t)\, \tau(\mathrm{d}t) \\
&= \int_H f^*(h)\, \Phi(\tau)(\mathrm{d}h) = \int_H f(h)\, \Phi(\tau)(\mathrm{d}h) \qquad \text{for each } f \in C_c(H).
\end{aligned}$$

As $\Phi(\tau_n)$ and $\Phi(\tau)$ are G-invariant the first and the last equation follow from (2.20). The second equation follows from the definition of the image measure, and the G-invariance of f^* implies the third and the fifth. By assumption, the sequence $(\tau_n)_{n\in\mathbb{N}}$ converges vague to $\tau \in \mathcal{M}(T)$. As $f \circ \varphi \circ \iota \in C_c(T)$ this implies the convergence '$\longrightarrow_{n\to\infty}$'. Altogether, we have proved $\Phi(\tau_n) \longrightarrow_{n\to\infty} \Phi(\tau)$. Since τ and the sequence $(\tau_n)_{n\in\mathbb{N}}$ were arbitrary

the restriction $\Phi_{|\mathcal{M}(T)}\colon \mathcal{M}(T) \to \mathcal{M}_G(H)$ has been proved to be continuous ([59], p. 54).

Let $\mathcal{R}$ be defined as in 2.50(v) with $M = H$. For $\eta := \sum_{j=1}^{k} a_j \kappa_{(Gh_j)} \in \mathcal{R}$ there exist $t_1, \ldots, t_k \in T$ with $t_j \in (\varphi \circ \iota)^{-1}(Gh_j)$. From 2.26(i) we obtain $\eta = \Phi\left(\sum_{j=1}^{k} a_j \varepsilon_{t_j}\right)$. As $\Phi_{|\mathcal{M}(T)}$ is continuous and $\mathcal{M}^{\leq 1}(T)$ is compact the image $\Phi(\mathcal{M}^{\leq 1}(T))$ is compact, too. Hence $\Phi(\mathcal{M}^{\leq 1}(T)) \supseteq \mathcal{R}$ implies $\Phi(\mathcal{M}^{\leq 1}(T)) \supseteq \overline{\mathcal{R}} = \mathcal{M}_G^{\leq 1}(H)$. Applying 2.26(ii) yields $\Phi(\mathcal{M}^{\leq 1}(T)) \subseteq \mathcal{M}_G^{\leq 1}(H)$, which proves $\Phi(\mathcal{M}^{\leq 1}(T)) = \mathcal{M}_G^{\leq 1}(H)$. Applying 2.26(ii) again yields $\Phi(\mathcal{M}^1(T)) = \mathcal{M}_G^1(H)$ which proves (i) provided that each $h \in H$ has a G-invariant neighbourhood U_h with relatively compact pre-image $(\varphi \circ \iota)^{-1}(U_h)$. In particular, condition (iii)(β) (= condition (vi)(β)) from 2.26 is fulfilled. Finally, 2.26(iv), (v) and (vi) imply $\Phi(\mathcal{M}(T)) = \mathcal{M}_G(H)$. Further, $\Phi^{-1}(\mathcal{M}_G(H)) = \mathcal{M}(T)$ for $\Phi\colon \mathcal{M}^{\sigma}(T) \to \mathcal{M}_G(H)$ which completes the proof of (ii).

It remains to prove (i) without the additional condition of assertion (ii). If T is compact this condition is automatically fulfilled and nothing has to be proved. Now assume that T is not compact. By Lemma 2.4(iii) T is σ-compact and hence there exists an increasing sequence $T_1, T_2 \cdots$ of compact subsets of T with $\bigcup_{j\in\mathbb{N}} T_j = T$. As φ is continuous $(\varphi(S \times T_j))_{j\in\mathbb{N}}$ is a monotonously non-decreasing sequence of compact G-invariant subsets of H with $\bigcup_{j=1}^{\infty} \varphi(S \times T_j) = H$. We define $A_1 := \varphi(S \times T_1)$ and inductively $A_{j+1} := \varphi(S \times T_{j+1}) \setminus \varphi(S \times T_j)$. The A_j are G-invariant Borel sets which are locally compact in the induced topologies (2.4(ii)), and $H = \sum_{j=1}^{\infty} A_j$. For $\nu \in \mathcal{M}_G^1(H)$ we define measures $\nu_j \in \mathcal{M}(H)$ by $\nu_j(B) := \nu(B \cap A_j)$. In particular, $\nu_j((A_j)^c) := 0$ and $\nu = \sum_{j=1}^{\infty} \nu_j$. As A_j is G-invariant we further conclude $\nu_j \in \mathcal{M}_G(H)$. Obviously, all 5-tupels $\left(G, S, T_j, \varphi(S \times T_j), \varphi_{|S\times T_j}\right)$ have Property $(*)$. As T_j is compact the additional condition of (ii) is trivially fulfilled. With ν_j also $\nu_{j\,|\mathcal{B}(\varphi(S\times T_j))}$ is G-invariant. Assertion (ii) implies that for each $j \in \mathbb{N}$ there exists a $\tau_j' \in \mathcal{M}(T_j)$ with $\nu_{j\,|\mathcal{B}(\varphi(S\times T_j))} = \left(\mu_{(S)} \otimes \tau_j'\right)^{\varphi_{|S\times T_j}}$. On the other hand, $\tau_j(B) := \tau_j'(B \cap T_j)$ defines a probability measure on T with $\tau_j((T_j)^c) = 0$. Since $\tau_{j\,|\mathcal{B}(T_j)} = \tau_j'$ and $\nu_j((\varphi(S \times T_j))^c) = 0$. we conclude $\nu_j = \left(\mu_{(S)} \otimes \tau_j\right)^{\varphi}$. Now let $\tau := \sum_{j=1}^{\infty} \tau_j$. Obviously, $\tau \in \mathcal{M}^{\sigma}(T)$. Theorem 2.26(ii) implies $\tau(T) = \sum_{j=1}^{\infty} \tau_j(T) = \sum_{j=1}^{\infty} \nu_j(H) = \nu(H) = 1$, and hence $\tau \in \mathcal{M}^1(T)$. From the construction of the measures ν_j and τ we obtain

$$(\mu \otimes \tau)^{\varphi}(A) = \tau\left((\varphi \circ \iota)^{-1}(A)\right) = \sum_{j=1}^{\infty} \tau_j\left((\varphi \circ \iota)^{-1}(A)\right) = \sum_{j=1}^{\infty} \nu_j(A) = \nu(A)$$

for all $A \in \mathcal{B}_G(H)$. Theorem 2.18(i) implies $\left(\mu_{(S)} \otimes \tau\right)^{\varphi} = \nu$ which completes the proof of (i). □

The proof of Theorem 2.51 makes extensive use of the fact that T and H are second countable (condition (-a-) of Property $(*)$). The following coun-

terexample shows that not only the proof of Theorem 2.51 breaks down if we drop this assumption. In fact, Theorem 2.51 becomes wrong.

Counterexample 2.52. Let $T = H = [0,1]$, $S = \{s_0\}$ and $\varphi(s_0, x) := x$ while the group $G = \{g_0\}$ acts trivially on S and T. We equip T with the discrete topology and H with the Euclidean topology. In particular, the product space $S \times T$ is discrete and φ is continuous. Further, $\mathcal{B}_0(T) = \mathcal{B}(H)$ equals the 'common' Borel-σ-algebra on $[0,1]$ whereas $\mathcal{B}(T) = \mathcal{P}(T)$. Obviously, $\tau \in \Phi^{-1}(\lambda)$ iff τ is an extension of the Lebesgue measure λ from $\mathcal{B}([0,1])$ to the power set $\mathcal{P}([0,1])$. Hence $\Phi^{-1}(\lambda) = \emptyset$ ([44], p. 50; cf. also Example 2.20(ii)).

3 Significance, Applicability and Advantages

This book investigates invariant probability measures and, more generally, invariant Borel measures on second countable locally compact spaces. The radially symmetric measures in $\mathbb{R}^2$ surely represent the most familiar non-trivial example. Various aspects have intensively been discussed in the introduction.

Apart from the preparatory Section 2.1 throughout this book we assume that a second countable compact group G acts on a second countable locally compact space H. Our interest lies in $\mathcal{M}_G(H)$ and $\mathcal{M}_G^1(H)$ which denote the set of G-invariant Borel measures or the set of G-invariant probability measures on H, resp. In Chapter 4 a number of examples underline that for many applications there exist spaces S and T and a surjective measurable mapping $\varphi\colon S \times T \to H$ where G acts transitively on S. In particular, there exists a unique probability measure $\mu_{(S)}$ on S. Clearly, the G-action on S can be extended to an action on the product space $S \times T$ via $g(s,t) := (gs,t)$. If the mapping $\varphi\colon S \times T \to H$ is G-equivariant, i.e. if $\varphi(gs,t) = g\varphi(s,t)$ for all $(s,t,g) \in S \times T \times G$, then the image measure $\big(\mu_{(S)} \otimes \tau\big)^{\varphi}$ is a G-invariant measure on H. The much more difficult problem is to verify whether

$$\mathcal{M}_G^1(H) = \big(\mu_{(S)} \otimes \mathcal{M}^1(T)\big)^{\varphi} \quad \text{and} \quad \mathcal{M}_G(H) = \big(\mu_{(S)} \otimes \mathcal{M}(T)\big)^{\varphi}, \tag{3.1}$$

i.e. whether each G-invariant Borel measure ν on H can be represented as an image of a particular product measure $\mu_{(S)} \otimes \tau_\nu$ with $\tau_\nu \in \mathcal{M}(T)$. Note that the mapping φ and the left-hand probability measure $\mu_{(S)}$ are the same for all G-invariant measures on H whereas the 'individuality' of ν is encoded in the right-hand measure τ_ν.

If the 5-tupel (G,S,T,H,φ) has Property $(*)$ Theorem 2.26 provides an equivalent characterization of the 'surjectivity problems' $\mathcal{M}_G^1(H) \stackrel{?}{=} \big(\mu_{(S)} \otimes \mathcal{M}^1(T)\big)^{\varphi}$ and $\mathcal{M}_G(H) \stackrel{?}{=} \big(\mu_{(S)} \otimes \mathcal{M}(T)\big)^{\varphi}$, namely as measure extension problems (Theorem 2.26). The first and (under additional assumptions) also the second equation are valid for continuous φ (2.32(ii) and 2.51, resp.). Theorem 2.34 extends this result, while 2.32(i) and 2.36 provide sufficient conditions that (3.1) holds if φ is not assumed to be continuous. In particular, Theorems 2.32 and 2.36 give concrete candidates $\tau_\nu \in \Phi^{-1}(\nu)$. The representation of the G-invariant measures on H as images of particular product measures is the mathematical core of this work. As a by-product we

W. Schindler: LNM 1808, pp. 55–62, 2003.

additionally obtained some results from the field of measure extensions (cf. 2.21, 2.33).

The results derived in Chapter 2 are important of their own, and surely they are aesthetically pleasing. Moreover, these results can fruitfully be applied to simplify and to speed up concrete computations. We have already discussed various aspects in the introduction, and in Chapter 4 we will study a number of examples. Before we come to these examples we briefly recall the main aspects discussed in the introduction. We broaden our reflections on stochastic simulations by considerung new aspects.

If $\nu = \left(\mu_{(\mathrm{S})} \otimes \tau_\nu\right)^\varphi \in \mathcal{M}_G(H)$ the transformation formula for integrals and Fubini's Theorem imply

$$\begin{aligned}\int_H f(h)\,\nu(\mathrm{d}h) &= \int_{S\times T} f\left(\varphi(s,t)\right)\,\left(\mu_{(\mathrm{S})}\otimes\tau_\nu\right)(\mathrm{d}s,\mathrm{d}t) \\ &= \int_T \left(\int_S f\left(\varphi(s,t)\right)\,\mu_{(\mathrm{S})}(\mathrm{d}s)\right)\,\tau_\nu(\mathrm{d}t).\end{aligned} \tag{3.2}$$

for all ν-integrable $f\colon H \to \mathbb{R}$. The displayed formula reminds on polar or, more generally, on spherical coordinates where the variables s and t correspond with the angles and the radius, resp. Indeed, polar and spherical coordinates represent special cases of the more general situation covered by Property $(*)$. Property $(*)$ does not demand that the mapping $\varphi\colon S\times T \to H$ is bijective and hence, unlike for spherical coordinates, the measure τ_ν usually is not unique. Moreover, the spaces S, T and H need not be subsets of the $\mathbb{R}^n$. As the radially symmetric functions for spherical coordinates in the general case the G-invariant functions are of particular interest since the double-integral in (3.2) simplifies to

$$\int_H f(h)\,\nu(\mathrm{d}h) = \int_T f\left(\varphi(s_0,t)\right)\,\tau_\nu(\mathrm{d}t) \qquad \text{for any } s_0 \in S \tag{3.3}$$

if $f\colon H \to \mathbb{R}$ is a G-invariant ν-integrable function. Note that the due to the equivariance of φ the integrand $f\left(\varphi(s,t)\right)$ does not depend on s. The element $s_0 \in S$ can hence be chosen arbitrarily. For each G-invariant integrand f the integral over H can be reduced into an integral over T. For many applications this simplifies concrete computations enormously since the space T normally is a manifold or a subset of $\mathbb{R}^n$, and its dimension is considerably smaller than that of H.

In Section 4.4 we treat the probability measures on the special orthogonal group $\mathrm{SO}(n)$ which are invariant under conjugation ('conjugation-invariant' probability measures). For $n \in \{2k, 2k+1\}$ the space T equals a subgroup of $\mathrm{SO}(n)$ which is isomorphic to k-dimensional cube $[0, 2\pi)^k$ (equipped with the componentwise addition modulo 2π). On the other hand $H = G = \mathrm{SO}(n)$ is an $n(n-1)/2$-dimensional manifold which can be embedded in $\mathbb{R}^{n^2}$. Integrals over $\mathrm{SO}(3)$ with respect to conjugation-invariant Borel measures and $\mathrm{SO}(3)$-invariant integrands, for instance, simplify to integrals over $\mathbb{R}$. The

advantages for analytical and numerical computations are obvious. The general linear group $H = \mathrm{GL}(n)$ appears in many applications. Depending on the degree of symmetry of the measure ν and the integrand, integrals over $\mathrm{GL}(n)$ can be reduced to integrals over the group of upper triangular $(n \times n)$-matrices with positive diagonal elements or over the set of symmetric positive definite $(n \times n)$-matrices, resp., or even to integrals over the diagonal matrices (cf. the Sections 4.7, 4.8 and 4.10). In the last case the dimension of the domain of integration shrinks from n^2 to n and, additionally, in many cases the integrands simplify notably. On $\mathrm{GL}(n)$ the determinant function and the spectral norm, for example, can be represented as complicated and unhandy terms in the matrix components. For diagonal matrices both terms are extremely simple. This opens up a wide range of interesting applications (cf. Section 4.10).

Statistical problems represent a further field of application for the symmetry calculus. Under suitable conditions the determination of a powerful test and of its distribution under the admissible hypotheses can be transferred from $H \times \cdots \times H$ to $T \times \cdots \times T$ without loss of information. The advantages are similar to those for integration problems. Instead of the observed sample $z_1, \ldots, z_m$ itself it suffices to store and to consider their images $\psi(z_1), \ldots, \psi(z_m)$ under a suitable mapping $\psi\colon H \to T$. For test problems on $\mathrm{SO}(3)$ with conjugation-invariant admissible hypotheses, for example, we may restrict our attention to the traces of the observed matrices rather than on the matrices themselves.

In the past three decades stochastic simulations have become an important tool in applied mathematics, information science and various applied sciences (see, e.g., [1]). Exemplary for countless application we point out Monte-Carlo-methods for the evaluation of high-dimensional integrals, the simulation of particular physics experiments, simulations of queing models to estimate the service time in computer networks, and simulations used in the development of new car models where random environmental influences are considered.

Roughly speaking, in a stochastic simulation one generates algorithmically values $\widetilde{V}_1, \ldots, \widetilde{V}_N$ which should have similar statistical properties as 'true' realizations of (i.e., values assumed by) random variables $V_1, \ldots, V_N$. These random variables usually (but not always) are independent and identically distributed (*iid*) with unknown distribution η_0. Due to the Law of large numbers and the convergence theorem of Glivenko and Cantelli for empirical distributions (cf. [28], p. 145) it is reasonable to expect that one can derive 'sound' estimators for η_0 or at least for particular values $\eta_0(A_1), \ldots, \eta_0(A_k)$. In many applications the random variables V_j can be represented as functions of further random variables $R_{1;j}, \ldots, R_{s;j}$ with known distribution, i.e. $V_j := \Psi(R_{1;j}, \ldots, R_{s;j})$ (cf. [18]). In order to unify the giantic number of different simulation problems one usually generates so-called *standard random numbers* $\widetilde{U}_1, \widetilde{U}_2, \ldots \in [0, 1)$ first which should have similar statistical prop-

erties as realization sequences of iid random variables $U_1, U_2, \ldots$ which are equidistributed on the interval $[0,1)$. In the next step one determines transformation(s) which map the sequence $U_1, U_2, \ldots$ onto the random variables $R_{1;1}, \ldots, R_{s;1}, \ldots, R_{1;j}, \ldots, R_{s;N}$. If the $R_{i;j}$ are iid exponentially distributed then $\widetilde{R}_{i;j} := -\ln(\widetilde{U}_{s(j-1)+i})$, for instance. Generally speaking, one derives *pseudorandom elements* $\widetilde{R}_{i;j}$ from the standard random numbers $\widetilde{U}_1, \widetilde{U}_2, \ldots$. Using the $\widetilde{R}_{i;j}$ one computes the desired pseudorandom elements $\widetilde{V}_1, \ldots, \widetilde{V}_N$, and therefrom the latter one draws conclusions on the unknown distribution η_0.

Remark 3.1. (i) Although pseudorandom elements are generated deterministically and hence do not have a distribution in a probabilistic sense the common literature suggestively speaks of η_0-distributed pseudorandom elements $\widetilde{V}_j$. To avoid clumsy formulations and since it should not cause misunderstandings we adopt this convention. If the random variables V_j assume real numbers or values in a real vector space, resp., the pseudorandom elements $\widetilde{V}_j$ are also denoted as *pseudorandom numbers* or *pseudorandom vectors*, resp.
(ii) In very rare cases the standard random numbers are not generated with a pseudorandom number generator but with a physical random number generator. Usually, a physical noise source (typically realized by an electronic switching) generates an analog signal which is digitalized after uniform time intervals, e.g. by a comparator, yielding a random bit. Physical random number generators are much slower than pseudorandom number generators. Moreover, pseudorandom number generators cost no more than some additional lines of code. On the positive side random numbers generated by a physical noise source are 'truly' random.
(iii) Physical random number generators are widely used for cryptographic applications, e.g. for the generation of session keys or signature parameters. Clearly, the suitability of the general design of the noise source has to be verified. Moreover, further aspects have to be considered when the random number generator is in operation: The noise source could break down instantaneously or the quality of the random bits might become worse in the course of the time (see e.g. [72, 73]).

If $\left(\mu_{(\mathrm{S})} \otimes \tau_\nu\right)^\varphi = \nu \in \mathcal{M}_G^1(H)$ the simulation of iid ν-distributed random variables $Z_1, \ldots, Z_N$ can be decomposed into two independent simulation problems on S and T. More precisely, one generates $\mu_{(\mathrm{S})}$-distributed pseudorandom elements $\widetilde{X}_1, \ldots, \widetilde{X}_N \in S$ and τ_ν-distributed pseudorandom elements $\widetilde{Y}_1, \ldots, \widetilde{Y}_N \in T$. Finally, $\widetilde{Z}_1 := \varphi(\widetilde{X}_1, \widetilde{Y}_1), \ldots, \widetilde{Z}_N := \varphi(\widetilde{X}_N, \widetilde{Y}_N)$ yields the desired ν-distributed pseudorandom elements. This approach has various advantages: First of all, due to the outstanding symmetry of $\mu_{(\mathrm{S})}$ generally there exist efficient algorithms for the generation of pseudorandom elements $\widetilde{X}_j$. These algorithms are often based on geometrical considerations (cf. Chapter 4). In many applications the space T is a manifold or a subset of $\mathbb{R}^n$ and its dimension is considerably smaller than that of H. In the rule

simulation problems on S and T are much easier to handle than a 'direct' generation of ν-distributed pseudorandom elements on H and, moreover, in many applications the domain T is more suitable for concrete computations. In this connection we point out that for integration and simulation problems it may not always be advisable to choose a section $R_G \subseteq T$ instead of T itself. In nearly all examples presented in Chapter 4 the space T is a cone in a vector space or a cube in $\mathbb{R}^n$ while H is a high-dimensional manifold. As the first step, i.e. the generation of pseudorandom elements on S, is the same for all G-invariant probability measures on H this approach leads to a partial unification of the simulation algorithms. The decomposition into two independent simulation problems may reduce the needed computation time drastically (cf. Section 4.5). Moreover, normally this reduces the average number of standard random numbers needed per pseudorandom element $\widetilde{Z}_j \in H$. At first sight the last argument may appear artificial and of pure academic nature. However, elementary considerations underline that this is not the case.

The usefulness of a transformation χ for simulation purposes depends essentially on two criteria. First of all, χ should be a simple term which enables a fast generation of pseudorandom elements. A second criterion is the average number of standard random numbers required per pseudorandom element. As already pointed out in general such algorithms are preferable for which this average number is small. To see this one should be aware that the reliability of a stochastic simulation depends essentially on the properties of the used standard random numbers. The standard random numbers are not truly random but algorithmically generated. Occasionally, this may cause unpleasant effects. An important criterion for the quality of a standard random number generator are the approximation properties of the *standard random vectors* $(\widetilde{U}_1, \ldots, \widetilde{U}_n), (\widetilde{U}_2, \ldots, \widetilde{U}_{n+1}), \ldots \in [0,1)^n$ with respect to the n-dimensional Lebesgue measure for $n = 2, 3, \ldots$ ([22], [46], pp. 75–113). As the common pseudorandom number generators are periodic the standard random vectors thin out rapidly when the dimension n increases. Consequently, their approximation properties must become worse. (We mention that this is not the case if a physical random number generator is used.) The smaller the average number of standard random numbers needed per pseudorandom element is the minor should be the influence of the approximation properties of the standard random vectors in high dimensions on the quality, that is, on the reliability of the stochastic simulation.

The following example underlines that this requirement is not absolute. In fact, it sometimes counteracts a third criterion which usually is completely disregarded: namely, whether χ transforms the finite structure of the used standard random vectors in a way that fits to the geometry of the space and to the symmetry of the simulated distribution.

Example 3.2. (Equidistribution on the unit sphere) As in Example 2.16 the terms S^2 and $\mu_{(\mathrm{S}^2)}$ denote the surface of the unit ball in $\mathbb{R}^3$ (unit sphere)

and the equidistribution on S^2, resp. We define the mapping

$$\chi\colon [0,1)^2 \to S^2$$
$$\chi(u,v) := \left(\sqrt{1-(2v-1)^2}\cos(2\pi u)\,,\sqrt{1-(2v-1)^2}\sin(2\pi u)\,,2v-1\right).$$

The restriction $\chi_{|(0,1)^2} \to S^2$ is a diffeomorphism on its image $\chi((0,1))$, and an elementary computation verifies $((D\chi)^t(D\chi))\ (u,v) \equiv 16\pi^2$ where $D\chi$ denotes the Jacobi matrix of χ. The mapping χ transforms $\lambda_{[0,1)^2}$, the Lebesgue measure on $[0,1)^2$, into $\mu_{(S^2)}$, the equidistribution on the unit sphere. Compared with other algorithms for the generation of equidistributed pseudorandom vectors on the unit sphere the mapping χ has two undeniable advantages: The generation of each pseudorandom element on S^2 does only need two standard random numbers, and insights into the structure of two-dimensional standard random vectors can easily be transferred to the pseudorandom vectors on S^2. Nevertheless, the use of the transformation χ also has two grave disadvantages. First, for each pseudorandom vector on S^2 a square root and two trigonometric functions have to be evaluated. Consquently, the transformation χ is of little practical significance since the outstanding geometry of a ball admits a large number of algorithms which are much faster ([18], pp. 230f.). From a theoretical point of view, however, there is another serious problem. If the standard random numbers are generated with a *linear congruential generator*, for instance, the pairs $(\widetilde{U}_1,\widetilde{U}_2),(\widetilde{U}_2,\widetilde{U}_3),\ldots$ form a shifted lattice. The mapping χ transforms the edges of this lattice into spiral lines. Near the poles the distance between neighboured points lying on the same spiral line, that is, the distance between neighboured possible values a pseudorandom vector $\widetilde{R}_j := \chi(\widetilde{U}_{2j-1},\widetilde{U}_{2j})$ can assume, is rather small. Relative to these distances the stripes between neighbouring spirals are very broad. Clearly, these stripes cannot be hit by any pseudorandom vector. This effect counteracts the homogeneity of the sphere and the symmetry of $\mu_{(S^2)}$. This is a consequence of the fact that the absolute values of the partial derivatives $\partial\chi_1(u,v)/\partial v$ and $\partial\chi_2(u,v)/\partial v$ tend to infinity as the vertical components of (u,v) tend to 0 or 1, respectively. At the same time the corresponding partial derivatives in u-direction tend to 0 (for details see [64], pp. 30 ff.).

It is obvious that such unpleasant effects may also occur for other generator types. Of course, if the period of the used standard number generator is sufficiently large the non-reachable regions are so small that this irregularity may be neglected. However, Example 3.2 calls our attention to a more general problem.

For many familiar distributions in $\mathbb{R}$ or $\mathbb{R}^n$, resp., a lot of sophisticated transformations have been published which are usually verified with concrete computations or elementary geometric arguments ([18]). Unless there are no outstanding and obvious symmetries (as for the equidistribution on the sphere, for example) the simulation of a distribution η_0 on an n-dimensional manifold M is usually traced back to a simulation problem on

$\mathbb{R}^n$. Therefore, one chooses a possibly small number t of disjoint Borel subsets $B_1, B_2, \ldots, B_t \subseteq M$ with $\eta_0(B_j) > 0$ which cover M (possibly apart from a set with η_0-measure zero). Each B_j is contained in an open set $U_j \subseteq M$ where $\chi_j : U_j \to V_j \subseteq \mathbb{R}^n$ denotes the corresponding chart mapping (cf. Remark 4.16). The generation of pseudorandom elements on M falls into three steps. At first, one generates a pseudorandom index $j \in \{1, \ldots, t\}$ with respect to the probability vector $(\nu(B_1), \ldots, \nu(B_t))$. In the second step one generates a pseudorandom vector $\widetilde{X}$ on $\chi_j(B_j)$ with respect to the distribution $\nu^{\chi_j}_{|B_j}/\nu(B_j)$. Finally, $\widetilde{Z} := \chi_j^{-1}(\widetilde{X})$ yields the desired pseudorandom element on M. Clearly, for the second step standard techniques can be applied. Unfortunately, the image measure $\nu_{j\,|B_j}{}^{\chi_j}/\nu_j(B_j)$ depends essentially on the choice of the chart mapping χ_j. Normally, these image measures neither have any symmetries nor they are contained in familiar distribution classes for which efficient simulation algorithms are known. To improve the practical feasibility one clearly tries to keep the number t as small as possible. Especially, if the dimension of M is large it may turn out to be problematic to choose large subsets of maximal charts. Near the boundary of $\chi_j(U_j)$ the Jacobian matrices $D\chi_j^{-1}(x)$ and $D\chi_j^{-1}(y)$ may differ extremely even if the arguments x and y lie closely together. Example 3.2 (with $t = 1$ and $B_1 = U_1 = \chi((0,1)^2)$) shows that even if the pseudorandom vectors on $\chi_j(B_j)$ yield an acceptable simulation of $\nu^{\chi_j}_{|B_j}/\nu(B_j)$ this need not be the case for the random elements on M. Clearly, the quantitative meaning of 'close' and 'near' depends on the distances between neighboured pseudorandom random elements on $\chi_j(B_j)$. As the standard random vectors thin out when the dimension n increases this may also affect commonly used generators with large period lengths (e.g. linear congruential generators with period $\geq 2^{64}$) and cause defects similar to those discussed above.

Clearly, under such circumstances one should not trust the results obtained by a stochastic simulation. Dividing M into many small subsets should at least reduce those defects. However, the number of computations and hence the time required to determine handy expressions for the images $\psi_1(B_1), \ldots, \psi_t(B_t)$ and to compute explicit formulas for the distributions $\nu^{\psi_1}/\psi_1(B_1), \ldots, \nu^{\psi_t}/\psi_t(B_t)$ increases linearly in t.

In many cases there are no alternatives. For G-invariant distributions on H, however, the situation is much better. As already set out the mapping $\varphi : S \times T \to H$ can be used to decompose a G-invariant simulation problem on H into two independent simulation problems on S and T which usually are noticeably easier to handle. This decomposition abstains completely from chart mappings. (Note that Property $(*)$ does not demand that H has to be a manifold.) Moreover, φ usually is a 'natural' mapping, e.g. a matrix multiplication, and the G-invariant probability measure $\mu_{(\mathrm{S})}$ on the homogeneous space S has maximal symmetry. In the rule there exist efficient algorithms for a simulation of $\mu_{(\mathrm{S})}$. There are no chart boundaries on H, and at least within the particular orbits the generated pseudorandom elements should not

distinguish any region. It thus remains to avoid unpleasant effects in the simulation of $\tau_\nu \in \mathcal{M}(T)$. Normally, this should be much easier to achieve than in a simulation on H which uses charts. In Section 4.5 the efficiency of this approach will be demonstrated at hand of conjugation-invariant distributions on SO(3). Moreover, applying 2.18(iv) efficient algorithms for a simulation of the composition of random rotations and for the computation of its distribution will be derived. In Section 4.11 the usefulness of the symmetry concept for simulations on large finite sets is explained. In Section 4.2 we will discuss an example from the field of combinatorial geometry. Although it is not of the type (G, S, T, H, φ) it uses group actions and equivariant mappings.

4 Applications

In Chapter 2 the symmetry concept was introduced and we proved a number of theorems. Their importance, their applicability and their benefit for applications have already been worked out in the introduction and the preceding chapter. In the present chapter these results serve as useful tools for a large number of various applications. The main emphasis lies on integration problems, stochastic simulations and statistical problems. We recommend to study Section 4.1 first while the remaining sections may be read in arbitrary order. Therefore, some redundancies in the representation of this chapter had to be accepted.

4.1 Central Definitions, Theorems and Facts

In this section we summarize the main theorems from Chapter 2 (2.18, 2.26, 2.32, 2.34, 2.36, 2.44, 2.51) as far as these results are relevant for applications. Though interesting for themselves we leave out some results from the field of measure extensions and results on the existence of pre-images under the mapping $\Phi\colon \mathcal{M}^\sigma(T) \to \mathcal{M}_G^+(H)$ (cf. Theorem 4.8 or Theorem 2.26, resp.) if the results do not refer to $\mathcal{M}_G^1(H)$ or $\mathcal{M}_G(H)$ but to individual measures $\nu \in \mathcal{M}_G(H)$. However, we encourage the interested reader to study Chapter 2 or at least its main theorems and corollaries (2.21, 2.22, 2.33) which provide more information than restated below. First, we repeat central definitions from Chapter 2 which will be needed in the present chapter.

Definition 4.1. *The restriction of a mapping $\chi\colon M_1 \to M_2$ to $E_1 \subseteq M_1$ is denoted with $\chi_{|E_1}$, the pre-image of $F \subseteq M_2$, i.e. the set $\{m \in M_1 \mid \chi(m) \in F\}$, with $\chi^{-1}(F)$. If χ is invertible χ^{-1} also stands for the inverse of χ. The union of* disjoint *subsets A_j is denoted with $A_1 + \cdots + A_k$ or $\sum_j A_j$, resp.*

Definition 4.2. *Let M be a topological space. A collection of open subsets of M is called a* topological base *if each open subset $O \subseteq M$ can be represented as a union of elements of this collection. The space M is said to be* second countable *if there exists a countable base. A subset $U \subseteq M$ is called a* neighbourhood *of $m \in M$ if there is an open subset $U' \subseteq M$ with $\{m\} \subseteq U' \subseteq U$. A topological space M is called* Hausdorff space *if for each $m \neq m' \in M$*

W. Schindler: LNM 1808, pp. 63–153, 2003.

there exist disjoint open subsets U and U' with $m \in U$ and $m' \in U'$. We call a subset F of a Hausdorff space M compact *if each open cover $\bigcup_{\iota \in I} U_\iota \supseteq F$ has a finite subcover. If M itself has this property then M is called a* compact space. *A subset $F \subseteq M$ is said to be* relatively compact *if its closure $\overline{F}$ (i.e. the smallest closed superset of F) is compact. A Hausdorff space M is called* locally compact *if each $m \in M$ has a compact neighbourhood. A topological space M is called* σ-compact *if it can be represented as a countable union of compact subsets. For $F \subseteq M$ the open sets in the* induced topology *(or synonymously:* relative topology*) on F are given by $\{O \cap F \mid O$ is open in $M\}$. In the* discrete topology *each subset $F \subseteq M$ is an open subset of M.*

Let M_1 and M_2 be topological spaces. The product topology *on $M_1 \times M_2$ is generated by $\{O_1 \times O_2 \mid O_i$ is open in $M_i\}$. A group G is said to be a* topological group *if G is a topological space and if the group operation $(g_1 g_2) \mapsto g_1 g_2$ and the inversion $g \mapsto g^{-1}$ are continuous where $G \times G$ is equipped with the product topology. We call the group G* compact *or* locally compact*, resp., if the underlying space G is compact or locally compact, resp.*

Illustrating examples are given in 2.5. Readers who are completely unfamiliar with the foundations of topology are referred to introductory books (e.g. to [59, 23, 43]). We point out that some topology works (e.g. [43]) do not demand the Hausdorff property in the definition of compactness. For this, one should be careful when combining lemmata and theorems with compactness assumptions from different topology articles or books. Note that each locally compact space is compact. Unless otherwise stated within this chapter all countable sets are equipped with the discrete topology while the $\mathbb{R}^n$ and matrix groups are equipped with the usual Euclidean topology. Next, we recall definitions from measure theory.

Definition 4.3. *The* power set *of Ω is denoted with $\mathcal{P}(\Omega)$. A* σ-algebra *(or synonymously: a* σ-field*) $\mathcal{A}$ on Ω is a subset of $\mathcal{P}(\Omega)$ with the following properties: $\emptyset \in \mathcal{A}$, $A \in \mathcal{A}$ implies $A^c \in \mathcal{A}$, and $A_1, A_2, \ldots \in \mathcal{A}$ implies $\bigcup_{j \in \mathbb{N}} A_j \in \mathcal{A}$. The pair $(\Omega, \mathcal{A})$ is a* measurable space. *A subset $F \subseteq \Omega$ is* measurable *if $F \in \mathcal{A}$. Let $(\Omega_1, \mathcal{A}_1)$ and $(\Omega_2, \mathcal{A}_2)$ be measurable spaces. A mapping $\varphi \colon \Omega_1 \to \Omega_2$ is called* $(\mathcal{A}_1, \mathcal{A}_2)$-measurable *if $\varphi^{-1}(A_2) \in \mathcal{A}_1$ for each $A_2 \in \mathcal{A}_2$. We briefly call φ measurable if there is no ambiguity about the σ-algebras $\mathcal{A}_1$ and $\mathcal{A}_2$. The* product σ-algebra *of $\mathcal{A}_1$ and $\mathcal{A}_2$ is denoted with $\mathcal{A}_1 \otimes \mathcal{A}_2$. A function $\Omega \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$ is called* numerical.

A measure ν on $\mathcal{A}$ is non-negative *if $\nu(F) \in [0, \infty]$ for all $F \in \mathcal{A}$. If $\nu(\Omega) = 1$ then ν is a* probability measure. *The set of all probability measures on $\mathcal{A}$, resp. the set of all non-negative measures on $\mathcal{A}$, are denoted with $\mathcal{M}^1(\Omega, \mathcal{A})$, resp. with $\mathcal{M}^+(\Omega, \mathcal{A})$. Let $\tau_1, \tau_2 \in \mathcal{M}^+(\Omega, \mathcal{A})$. The measure τ_2 is said to be* absolutely continuous *with respect to τ_1, abbreviated with $\tau_2 \ll \tau_1$, if $\tau_1(N) = 0$ implies $\tau_2(N) = 0$. If τ_2 has the τ_1-density f we write $\tau_2 = f \cdot \tau_1$. Let $\varphi \colon \Omega_1 \to \Omega_2$ be a measurable mapping between two measurable spaces Ω_1 and Ω_2, and let $\eta \in \mathcal{M}^+(\Omega_1, \mathcal{A}_1)$. The term η^φ stands for the* image measure

of η under φ, i.e. $\eta^{\varphi}(A_2) := \eta(\varphi^{-1}(A_2))$ for each $A_2 \in \mathcal{A}_2$. A measure $\tau_1 \in \mathcal{M}^+(\Omega, \mathcal{A})$ is called σ-finite *if Ω can be represented as the limit of a sequence of non-decreasing measurable subsets $(A_n)_{n\in\mathbb{N}}$ with $\tau_1(A_n) < \infty$ for all $n \in \mathbb{N}$. The set of all σ-finite measures on $\mathcal{A}$ is denoted with $\mathcal{M}^{\sigma}(\Omega, \mathcal{A})$. The* product measure *of $\eta_1 \in \mathcal{M}^{\sigma}(\Omega_1, \mathcal{A}_1)$ and $\eta_2 \in \mathcal{M}^{\sigma}(\Omega_2, \mathcal{A}_2)$ is denoted with $\eta_1 \otimes \eta_2$.*

Let $\mathcal{A}_0 \subseteq \mathcal{A} \subseteq \mathcal{A}_1$ denote σ-algebras over Ω. The restriction *of $\eta \in \mathcal{M}^+(\Omega, \mathcal{A})$ to the sub-σ-algebra $\mathcal{A}_0$ is denoted by $\eta_{|\mathcal{A}_0}$. A measure $\eta_1 \in \mathcal{A}_1$ is called an* extension *of η if its restriction to $\mathcal{A}$ coincides with η, i.e. if $\eta_{1|\mathcal{A}} = \eta$.*

The Borel σ-algebra *$\mathcal{B}(M)$ over a topological space M is generated by the open subsets of M. If it is unambiguous we briefly write $\mathcal{M}^1(M)$, $\mathcal{M}^{\sigma}(M)$ or $\mathcal{M}^+(M)$, resp., instead of $\mathcal{M}^1(M, \mathcal{B}(M))$, $\mathcal{M}^{\sigma}(M, \mathcal{B}(M))$ and $\mathcal{M}^+(M, \mathcal{B}(M))$, resp., and we call $\nu \in \mathcal{M}^+(M)$ a measure on M. If M is locally compact then $\nu \in \mathcal{M}^+(M)$ is said to be a* Borel measure *if $\nu(K) < \infty$ for each compact subset $K \subseteq M$. The set of all Borel measures is denoted with $\mathcal{M}(M)$.*

For fundamentals of measure theory we refer the interested reader e.g. to [5, 6, 28, 33]. In our applications all spaces will be locally compact and our interest lies in probability measures and Borel measures. However, occasionally it is advisable to consider $\mathcal{M}^+(H)$ instead of $\mathcal{M}(H)$ as this saves clumsy formulations and preliminary distinction of cases that certain mappings are well-defined. The most familiar Borel measure surely is the Lebesgue measure on $\mathbb{R}$. The following lemma will be useful in the subsequent sections.

Lemma 4.4. *(i) Let $\tau_1 \in \mathcal{M}^{\sigma}(M, \mathcal{A})$. Then the following conditions are equivalent:*

(α) $\tau_2 \ll \tau_1$.
(β) τ_2 has a τ_1-density f, i.e. $\tau_2 = f \cdot \tau_1$ for a measurable $f\colon M \to [0, \infty]$.

(ii) Let M be a topological space and $A \in \mathcal{B}(M)$. If A is equipped with the induced topology then $\mathcal{B}(A) = \mathcal{B}(M) \cap A$.
(iii) Let M_1 and M_2 be topological spaces with countable topological bases. Then $\mathcal{B}(M_1 \times M_2) = \mathcal{B}(M_1) \otimes \mathcal{B}(M_2)$.

Proof. cf. Lemma 2.8 □

Definition 4.5. *Let G denote a compact group, e_G its identity element and M a topological space. A continuous mapping $\Theta\colon G \times M \to M$ is said to be a* group action*, or more precisely, a* G-action *if $\Theta(e_G, m) = m$ and $\Theta(g_2, \Theta(g_1, m)) = \Theta(g_2 g_1, m)$ for all $(g_1, g_2, m) \in G \times G \times M$. We say that G* acts *(or synonymously:* operates*) on M and call M a* G-space*. If it is unambiguous we briefly write gm instead of $\Theta(g, m)$. The G-action is said to be* trivial *if $gm = m$ for all $(g, m) \in G \times M$. For $m \in M$ we call $G_m := \{g \in G \mid gm = m\}$ the* isotropy group*. The union $Gm := \bigcup_{g\in G}\{gm\}$*

is called the orbit *of* m. *The* G*-action is said to be* transitive *if for any* $m_1, m_2 \in M$ *there exists a* $g \in G$ *with* $gm_1 = m_2$. *If* G *acts transitively on* M *and if the mapping* $g \mapsto gm$ *is open for all* $m \in M$ *then* M *is a* homogeneous space. *A mapping* $\pi\colon M_1 \to M_2$ *between two* G*-spaces is said to be* G-equivariant *(or short:* equivariant*) if it commutes with the* G*-actions on* M_1 *and* M_2, *i.e. if* $\pi(gm) = g\pi(m)$ *for all* $(g, m) \in G \times M_1$. *A non-negative measure* ν *on* M *is called* G-invariant *if* $\nu(gB) := \nu(\{gx \mid x \in B\}) = \nu(B)$ *for all* $(g, B) \in G \times \mathcal{B}(M)$. *The set of all* G*-invariant probability measures (resp.* G*-invariant Borel measures, resp.* G*-invariant* σ*-finite measures, resp.* G*-invariant non-negative measures) on* M *are denoted with* $\mathcal{M}_G^1(M)$ *(resp.* $\mathcal{M}_G(M)$, *resp.* $\mathcal{M}_G^\sigma(M)$, *resp.* $\mathcal{M}_G^+(M)$*).*

Examples of σ-algebras, invariant measures and group actions are given in 2.7, 2.10 and 2.13. In Chapter 2 Lemma 2.14 was applied within the proofs of some main theorems. We will use parts of 2.14 directly in the Sections 4.2 and 4.11. The relevant statements are summarized in 4.6. We further refer the interested reader to 2.16 and 2.17 which illustrate specific aspects of Lemma 2.14. We point out that the transport of invariant measures is the central idea behind all results of this book.

Theorem 4.6. *Let* G *be a compact group which acts on the topological spaces* M, M_1 *and* M_2. *Assume further that the mapping* $\pi\colon M_1 \to M_2$ *is* G*-equivariant and measurable. Then*
(i) For each $\nu \in \mathcal{M}_G^+(M_1)$ *we have* $\nu^\pi \in \mathcal{M}_G^+(M_2)$. *If* $\nu \in \mathcal{M}_G^1(M_1)$ *then* $\nu^\pi \in \mathcal{M}_G^1(M_2)$.
(ii) $\mathcal{M}_G^1(M) \neq \emptyset$. *If* M *is finite then every* G*-invariant measure* ν *on* M *is equidistributed on each orbit, i.e.* ν *has equal mass on any two points lying to the same orbit.*
(iii) If G *acts transitively on a Hausdorff space* M *then* $|\mathcal{M}_G^1(M)| = 1$. *In particular,* M *is a compact homogeneous* G*-space.*

Proof. Lemma 2.14(i), (iii) and (v). □

Theorem 4.7(i) is crucial for the following. It says that G-invariant Borel measures on a second countable locally compact space M are uniquely determined by their values on the sub-σ-algebra

$$\mathcal{B}_G(M) := \{B \in \mathcal{B}(M) \mid gB = B \text{ for all } g \in G\}. \tag{4.1}$$

Suppose that X and Y are independent random variables on a locally compact group G'. If X and Y are η_1- and η_2-distributed, resp., then the distribution of $Z := XY$ is given by $\eta_1 \odot \eta_2$, the *convolution product* of η_1 and η_2. Equipped with the operation '$\odot$' the set $\mathcal{M}^1(G')$ becomes a semigroup $(\mathcal{M}^1(G'), \odot)$ with identity element $\varepsilon_{e_{G'}}$, the Dirac measure with total mass on the identity element $e_{G'}$. In Section 4.5 we exploit 4.7(iv) to derive efficient algorithms to simulate the composition of random rotations and to compute its distribution. Counterexample 2.19 shows that the uniqueness property 4.7(i) need not be valid if not $\nu_1, \nu_2 \in \mathcal{M}(M)$.

Theorem 4.7. *Suppose that the compact group G acts on a second countable locally compact space M. Then*
(i) *(Uniqueness property) For $\nu_1, \nu_2 \in \mathcal{M}_G(M)$ we have*

$$\left(\nu_{1|\mathcal{B}_G(M)} = \nu_{2|\mathcal{B}_G(M)}\right) \Rightarrow (\nu_1 = \nu_2)\,. \tag{4.2}$$

(ii) *If $\tau \in \mathcal{M}(M)$ then*

$$\tau^*(B) := \int_G \tau(gB)\,\mu_G(\mathrm{d}g) \qquad \text{for all } B \in \mathcal{B}(M) \tag{4.3}$$

defines a G-invariant Borel measure on M with $\tau^*{}_{|\mathcal{B}_G(M)} = \tau_{|\mathcal{B}_G(M)}$. If $\tau \in \mathcal{M}_G(M)$ then $\tau^* = \tau$.
(iii) *Let $\tau \in \mathcal{M}(M), \nu \in \mathcal{M}_G(M)$ and $\tau = f \cdot \nu$. Then*

$$\tau^* = f^* \cdot \nu \qquad \text{with } f^*(m) := \int_G f(gm)\,\mu_G(\mathrm{d}g). \tag{4.4}$$

In particular, $\nu(\{m \in M \mid f^*(m) = \infty\}) = 0$, and f^* is $(\mathcal{B}_G(M), \mathcal{B}(\overline{\mathbb{R}}))$-measurable.
(iv) *Let M be a group and $\Theta_g \colon M \to M, \Theta_g(m) := gm$ a group homomorphism for all $g \in G$. In particular, Θ_g is an isomorphism, and $(\mathcal{M}^1_G(M), \odot)$ is a sub-semigroup of the convolution semigroup $(\mathcal{M}^1(M), \odot)$.*

Proof. Theorem 2.18 □

Property $(*)$ formulates general assumptions which will be the basis of the following theorems.

$(*)$ The 5-tupel (G, S, T, H, φ) is said to have Property $(*)$ if the following conditions are fulfilled:
-a- G, S, T, H are second countable.
-b- T and H are locally compact, and S is a Hausdorff space.
-c- G is a compact group which acts on S and H.
-d- G acts transitively on S.
-e- $\varphi \colon S \times T \to H$ is a surjective measurable mapping.
-f- The mapping φ is equivariant with respect to the G-action $g(s,t) := (gs, t)$ on the product space $S \times T$ and the G-action on H, i.e. $g\varphi(s,t) = \varphi(gs,t)$ for all $(g,s,t) \in G \times S \times T$.

We mention that by 4.6(iii) there exists a unique G-invariant probability measure $\mu_{(S)}$ on G. In applications the group G and the space H are usually given, and the interest lies in the G-invariant probability measures, or more generally, in the G-invariant Borel measures on H, usually motivated by integration, simulation or statistical problems. To apply the theorems formulated below one has to find suitable spaces S and T and a measurable (preferably: continuous) mapping $\varphi \colon S \times T \to H$ such that the 5-tuple (G, S, T, H, φ) has Property $(*)$.

We have already pointed out that Property $(*)$ is not very restrictive. In general, the verification of its particular conditions does not cause serious problems. Moreover, one can hardly resign on them (cf. Section 2.3). In the following sections we will investigate several applications which have Property $(*)$.

Roughly speaking, we are faced with two fundamental problems: The first question is whether each $\nu \in \mathcal{M}_G(H)$ can be represented in the form $(\mu_{(\mathrm{S})} \otimes \tau_\nu)^\varphi$ for a suitable $\tau_\nu \in \mathcal{M}(T)$. Moreover, to a given $\nu \in \mathcal{M}_G(H)$ we have to find an explicit expression for such a $\tau_\nu \in \mathcal{M}(T)$ (provided that a measure with this property really exists). Theorem 4.8 below characterizes the possible candidates τ_ν. However, some explanatory remarks and definitions are necessary.

The inclusion map

$$\iota\colon T \to S \times T, \qquad \iota(t) := (s_0, t) \tag{4.5}$$

is continuous and hence in particular measurable. The element $s_0 \in S$ is arbitrary but fixed in the following. The sub-σ-algebra

$$\mathcal{B}_0(T) := \{\iota^{-1}(\varphi^{-1}(A)) \mid A \in \mathcal{B}_G(H)\} \subseteq \mathcal{B}(T) \tag{4.6}$$

is isomorphic to $\mathcal{B}_G(H) \subseteq \mathcal{B}(H)$. To be precise, $(\varphi \circ \iota)^{-1}$ induces a bijection between $\mathcal{B}_G(H)$ and $\mathcal{B}_0(T)$ which preserves unions and complements (cf. 2.25). In particular, each $\nu \in \mathcal{M}_G(H)$ corresponds with a unique $\nu_* \in \mathcal{M}^+(T, \mathcal{B}_0(T))$ which is given by

$$\nu_*(C) := \nu(\varphi(S \times C)) \qquad \text{for } C \in \mathcal{B}_0(T) \tag{4.7}$$

(cf. 2.25(vi)). The G-orbits on H are denoted with $E_G(\cdot)$. We define an equivalence relation $\sim$ on T by $t_1 \sim t_2$ iff $\varphi(S \times \{t_1\}) = \varphi(S \times \{t_2\})$, i.e. if $E_G(\{\varphi(s_0, t_1)\}) = E_G(\{\varphi(s_0, t_2)\})$. The equivalance classes on T (induced by $\sim$) are denoted with $E_T(\cdot)$. We point out that the equivalence relation $\sim$ does not depend on the specific choice of $s_0 \in S$: As G acts transitively on S to each $s_0' \in S$ there exists a $g_0 \in G$ with $s_0' = g_0 s_0$. This implies $\varphi(s_0', t) = \varphi(g_0 s_0, t) = g_0 \varphi(s_0, t) \in E_G(\{\varphi(s_0, t)\})$.

Let $\mathcal{A}_0 \subseteq \mathcal{A} \subseteq A_1$ denote σ-algebras over M and $\eta \in \mathcal{M}^+(M, \mathcal{A})$. Clearly, its restriction $\eta_{|\mathcal{A}_0}$ is a measure on the sub-σ-algebra $\mathcal{A}_0$. It does always exist and is unique. In contrast, an extension of η to $\mathcal{A}_1$ may not exist, and if an extension exists it need not be unique. The concept of measure extensions is crucial for the understanding of Theorem 4.8. The interested reader is referred to 2.20.

Theorem 4.8. *Suppose that the* 5*-Tupel* (G, S, T, H, φ) *has Property* $(*)$*, and let* $s_0 \in S$ *be arbitrary but fixed. Further,*

$$\Phi\colon \mathcal{M}^\sigma(T) \to \mathcal{M}_G^+(H) \qquad \Phi(\tau) := \left(\mu_{(\mathrm{S})} \otimes \tau\right)^\varphi. \tag{4.8}$$

Then the following statements are valid:
(i) *Let* $\nu \in \mathcal{M}_G(H)$ *and* $\tau \in \mathcal{M}^\sigma(T)$. *Then* $\nu_* \in \mathcal{M}^\sigma(T, \mathcal{B}_0(T))$, *and* $\tau \in \Phi^{-1}(\nu)$ *iff* $\tau_{|\mathcal{B}_0(T)} = \nu_*$.
(ii) $\tau(T) = (\Phi(\tau))(H)$ *for each* $\tau \in \mathcal{M}^\sigma(T)$.
(iii) *The following properties are equivalent:*

(α) $\Phi(\mathcal{M}^1(T)) = \mathcal{M}^1_G(H)$.
(β) For each $\nu \in \mathcal{M}^1_G(H)$ *the induced measure* $\nu_* \in \mathcal{M}^1(T, \mathcal{B}_0(T))$ *has an extension* $\tau_\nu \in \mathcal{M}^1(T)$.

Any of these conditions implies $\Phi(\mathcal{M}^\sigma(T)) \supseteq \mathcal{M}_G(H)$.
(iv) *If each* $h \in H$ *has a* G*-invariant neighbourhood* U_h *with relatively compact pre-image* $(\varphi \circ \iota)^{-1}(U_h) \subseteq T$ *then* $\Phi(\mathcal{M}(T)) \subseteq \mathcal{M}_G(H)$.
(v) *If* $\varphi\colon S \times T \to H$ *is continuous then* $\Phi^{-1}(\mathcal{M}_G(H)) \subseteq \mathcal{M}(T)$.
(vi) *Suppose that*

$$\Phi(\mathcal{M}(T)) \subseteq \mathcal{M}_G(H) \qquad \text{and} \qquad \Phi^{-1}(\mathcal{M}_G(H)) \subseteq \mathcal{M}(T) \tag{4.9}$$

(cf. (iv) and (v)). Then the following properties are equivalent:

(α) $\Phi(\mathcal{M}(T)) = \mathcal{M}_G(H)$.
(β) For each $\nu \in \mathcal{M}^1_G(H)$ *the induced probability measure* $\nu_* \in \mathcal{M}^1(T, \mathcal{B}_0(T))$ *has an extension* $\tau_\nu \in \mathcal{M}^1(T)$.
(γ) For each $\nu \in \mathcal{M}_G(H)$ *the induced measure* $\nu_* \in \mathcal{M}^\sigma(T, \mathcal{B}_0(T))$ *has an extension* $\tau_\nu \in \mathcal{M}^+(T)$.

Any of these conditions implies $\Phi^{-1}(\mathcal{M}_G(H)) = \mathcal{M}(T)$.
(vii) *Let* $\nu \in \mathcal{M}_G(H)$, $f \geq 0$ *a* G*-invariant numerical function and* $\nu = \left(\mu_{(\mathrm{S})} \otimes \tau_\nu\right)^\varphi$. *Then* $\eta := f \cdot \nu \in \mathcal{M}^+_G(H)$. *If* f *is locally* ν*-integrable then* $\eta \in \mathcal{M}_G(H)$ *and*

$$\eta = \left(\mu_{(\mathrm{S})} \otimes f_T \cdot \tau_\nu\right)^\varphi \qquad \text{with} \qquad f_T(t) := f\left(\varphi(s_0, t)\right). \tag{4.10}$$

If $T \subseteq H$ *and* $E_T(\{t\}) = \varphi(S \times \{t\}) \cap T$ *(i.e.* $= E_G(\{\varphi(s_0, t)\}) \cap T$*) then* $f_T(t) = f(t)$, *i.e.* f_T *is given by the restriction* $f_{|T}$ *of* $f\colon G \to [0, \infty]$ *to* T.

Proof. Theorem 2.26 □

The statements 4.8(iii) and (vi) are very similar. The additional conditions in 4.8(vi) ensure that Φ maps Borel measures onto Borel measures and that the pre-images of Borel measures are also Borel measures. Statement 4.8(vii) is very useful for concrete computations. An explicit expression for a particular pre-image τ_ν (typically given by its density with respect to a particular measure $\kappa \in \mathcal{M}(T)$) gives explicit expressions for a whole class of pre-images, namely pre-images of G-invariant Borel measures η on H with $\eta \ll \nu$. (Note that by 4.7(iii) there exists a G-invariant density $f^{(*)}$ with $\eta = f^{(*)} \cdot \nu$.)

By Theorem 4.8 the existence of a pre-image $\tau_\nu \in \Phi^{-1}(\nu)$ is equivalent to the existence of a measure extension, namely whether $\nu_* \in \mathcal{M}(T, \mathcal{B}_0)$ can be

extended to $\mathcal{B}(T)$. Depending on the concrete situation this measure extension problem may also be very difficult, at least if it shall be solved for all $\nu \in \mathcal{M}_G(H)$ simultaneously. For continuous φ 4.9(i) answers the existence problem positively. In 4.9(ii) the continuity assumption is relaxed (cf. Example 2.38). In all applications treated in the following sections the mapping φ will be continuous.

Theorem 4.9. *Suppose that the 5-tuple (G, S, T, H, φ) has Property (*). Then*
(i) *If φ is continuous then $\Phi(\mathcal{M}^1(T)) = \mathcal{M}_G^1(H)$.*
If, additionally, each $h \in H$ has a G-invariant neighbourhood U_h with relatively compact pre-image $(\varphi \circ \iota)^{-1}(U_h) \subseteq T$ then also $\Phi(\mathcal{M}(T)) = \mathcal{M}_G(H)$ and $\Phi^{-1}(\mathcal{M}_G(H)) = \mathcal{M}(T)$.
(ii) *Suppose that there exist countably many compact subsets $K_1, K_2, \ldots \subseteq T$ which meet the following conditions.*

(α) The restriction $\varphi_{|S\times K_j}: S \times K_j \to H$ is continuous for each $j \in \mathbb{N}$.
(β) $T = \bigcup_{j\in N} K_j$.

Then $\Phi(\mathcal{M}^1(T)) = \mathcal{M}_G^1(H)$.
If, additionally, each $h \in H$ has a G-invariant neighbourhood U_h with relatively compact pre-image $(\varphi \circ \iota)^{-1}(U_h) \subseteq T$ and if $\Phi^{-1}(\mathcal{M}_G(H)) \subseteq \mathcal{M}(T)$ then also $\Phi(\mathcal{M}(T)) = \mathcal{M}_G(H)$ and $\Phi^{-1}(\mathcal{M}_G(H)) = \mathcal{M}(T)$.

Proof. Statement (i) coincides with 2.32(ii) and 2.51, resp. Statement (ii) equals 2.34(iv). □

So far, we have essentially considered the pure existence problem, i.e. whether $\Phi(\mathcal{M}^1(T)) = \mathcal{M}_G^1(H)$ and $\Phi(\mathcal{M}(T)) = \mathcal{M}_G(H)$. Clearly, a solution of the existence problem is interesting for itself, and this knowledge may be useful for specific applications (e.g. for specific statistical problems). For integration problems or stochastic simulations we yet need an explicit expression for a pre-image $\tau_\nu \in \Phi^{-1}(\nu)$. It is often helpful to have a represention τ_ν as an image measure, i.e. $\tau_\nu = \nu^\psi$, under a suitable mapping $\psi: H \to T$. Theorems 4.10 and 4.11 illustrate the situation. Before, we return to the field of measure extensions.

Assume for the moment that $\mathcal{A}$ is a σ-algebra over a non-empty set Ω, $\eta \in \mathcal{M}^+(\Omega, \mathcal{A})$ and $\mathcal{N}_\eta := \{N \subseteq \Omega \mid N \subseteq A \text{ for an } A \in \mathcal{A} \text{ with } \eta(A) = 0\}$. There exists a unique extension η_v of η to the σ-algebra $\mathcal{A}_\eta := \{A + N \mid A \in \mathcal{A}, N \in \mathcal{N}_\eta\}$. In particular, $\eta_v(A + N) := \eta(A)$, and η_v is called the *completion* of η (cf. 2.20(iii)).

Note that each G-invariant Borel measure $\eta \in \mathcal{M}_G(H)$ can be uniquely extended to the σ-algebra $\mathcal{B}(H)_F := \bigcap_{\nu\in\mathcal{M}_G^1(H)} \mathcal{B}(H)_\nu$ (cf. Remark 2.30).

Theorem 4.10. *Suppose that the 5-tupel (G, S, T, H, φ) has Property (*). Then*

(i) *Assume that* $\psi\colon H \to T$ *is a* $(\mathcal{B}(H)_F, \mathcal{B}(T))$*-measurable mapping with* $\varphi(s_0, \psi(h)) \in E_G(\{h\})$ *for all* $h \in H$*. Then*

$$\nu = \left(\mu_{(\mathrm{S})} \otimes {\nu_v}^{\psi}\right)^{\varphi} \qquad \text{for each } \nu \in \mathcal{M}^1_G(H). \tag{4.11}$$

(ii) *If* $\psi\colon H \to T$ *has the properties demanded in (i) then there also exists a mapping* $\psi'\colon H \to T$ *with the same properties which additionally is constant on each* G*-orbit.*
(iii) *If* φ *is continuous then there exists a measurable mapping* $\psi\colon H \to T$ *with* $\varphi(s_0, \psi(h)) \in E_G(\{h\})$ *for all* $h \in H$*.*

Proof. Theorem 2.32(i), (iv), (ii). □

If $\varphi\colon H \to T$ meets the assumptions from 4.10(i) then ${\nu_v}^{\psi} \in \Phi^{-1}(\nu)$. Clearly, if $\psi\colon H \to T$ is measurable (i.e. $(\mathcal{B}(H), \mathcal{B}(T))$-measurable) then ${\nu_v}^{\psi} = \nu^{\psi}$. The existence of a mapping $\psi\colon H \to T$ with the properties demanded in 4.10(i) implies $\Phi(\mathcal{M}^1(T)) = \mathcal{M}^1_G(H)$. If the additional conditions (4.9) are fulfilled then also $\Phi(\mathcal{M}(T)) = \mathcal{M}_G(H)$ and $\Phi^{-1}(\mathcal{M}_G(H)) = \mathcal{M}(T)$. The chances and advantages implied by the property that each $\nu \in \mathcal{M}^1_G(H)$ (or even each $\nu \in \mathcal{M}_G(H)$) admits a representation $\nu = (\mu_{(\mathrm{S})} \otimes \tau_\nu)^{\varphi}$ with $\tau_\nu \in \mathcal{M}(T)$, have intensively been discussed in the introduction and in the preceding chapter.

Of course, if ψ is constant on each G-orbit then $\psi(H) \subseteq T$ is a *section* with respect to the equivalence relation $\sim$, i.e. $\psi(H)$ contains exactly one element of each $E_T(\cdot)$-orbit. However, the image $\psi(H)$ need not be measurable. Theorem 4.11 is the pendant to 4.10.

Theorem 4.11. *Suppose that the 5-tupel* (G, S, T, H, φ) *has Property* $(*)$ *and that* $R_G \subseteq T$ *is a measurable section with respect to* $\sim$*. To this section* R_G *there corresponds a unique mapping*

$$\psi\colon H \to R_G \subseteq T, \qquad \psi(h) := t_h \tag{4.12}$$

where $t_h \in R_G$ *denotes the unique element with* $\varphi \circ \iota(t_h) \in E_G(\{h\})$*. Finally, let* $\nu \in \mathcal{M}_G(H)$*.*
(i) $\psi^{-1}(C) = \varphi(S \times (C \cap R_G))$ *for each* $C \in \mathcal{B}(T)$*.*
(ii) *Suppose that* $\varphi(S \times (C \cap R_G)) \in \mathcal{B}(H)_F$ *for all* $C \in \mathcal{B}(T)$*. Then*

$$\tau_\nu(C) := \nu_v(\varphi(S \times (C \cap R_G))) \qquad \text{for each } C \in \mathcal{B}(T) \tag{4.13}$$

defines a σ*-finite measure* $\tau_\nu \in \Phi^{-1}(\nu)$ *with* $\tau_\nu(T \setminus R_G) = 0$*. If* $\nu = \left(\mu_{(\mathrm{S})} \otimes \tau_1\right)^{\varphi}$ *and* $\tau_1(T \setminus R_G) = 0$ *then* $\tau_1 = \tau_\nu$*.*
(iii) *The following properties are equivalent:*

(α) $\varphi(S \times (C \cap R_G)) \in \mathcal{B}(H)_F$ *for all* $C \in \mathcal{B}(T)$*.*
(β) The mapping $\psi\colon H \to T$ *is* $(\mathcal{B}(H)_F, \mathcal{B}(T))$*-measurable.*

Each of these properties implies ${\nu_v}^{\psi} = \tau_\nu$.
(iv) *Let* φ *be continuous and* R_G *be locally compact in the induced topology. Then* $\varphi(S \times (C \cap R_G)) \in \mathcal{B}(H)$ *for each* $C \in \mathcal{B}(T)$, *i.e. the mapping* $\psi\colon H \to T$ *is* $(\mathcal{B}(H), \mathcal{B}(T))$*-measurable.*

Proof. Theorem 2.36. □

The uniqueness property from 4.11(i) supports the computation of a preimage τ_ν. For concrete computations Theorem 4.11 usually is more suitable than Theorem 4.10. In typical applications it is not difficult to determine a measurable section $R_G \subseteq T$. In Sections 4.3 to 4.11 we will determine locally compact sections. By 4.11(iv) the corresponding mapping $\psi\colon H \to R_G \subseteq T$ is measurable then.

Next, we touch another field of possible applications of the symmetry concept besides integration and simulation problems, namely statistics. We begin with some definitions.

Definition 4.12. *A* test *on the measurable space* $(V, \mathcal{V})$ *is a measurable mapping* $\Psi\colon V \to [0,1]$. *The terms* Γ, $\Gamma_0 \subseteq \Gamma$ *and* $\Gamma \setminus \Gamma_0$ *denote non-empty parameter sets. In the sequel we identify* Γ_0 *and* $\Gamma \setminus \Gamma_0$ *with disjoint sets of probability measures* $\mathrm{P}_{\Gamma_0}, \mathrm{P}_{\Gamma\setminus\Gamma_0} \subseteq \mathcal{M}^1(V, \mathcal{V})$, *i.e.* $\gamma \mapsto p_\gamma$ *defines a bijection* $\Gamma \to \mathrm{P}_\Gamma := \mathrm{P}_{\Gamma_0} \cup \mathrm{P}_{\Gamma\setminus\Gamma_0}$. *A* test problem *on* $(V, \mathcal{V})$ *is given by a tupel* $\left(\mathrm{P}_{\Gamma_0}, \mathrm{P}_{\Gamma\setminus\Gamma_0}\right)$. *We call* P_Γ *the* admissible hypotheses, P_{Γ_0} *the* null hypothesis *and* $\mathrm{P}_{\Gamma\setminus\Gamma_0}$ *the* alternative hypothesis. *The mapping* $\mathrm{pow}_\Psi\colon \mathrm{P}_\Gamma \to [0,1]$, $\mathrm{pow}_\Psi(p_\gamma) := \int_V \Psi \,\mathrm{d}p_\gamma$ *is called* power function *of* Ψ. *Let* $(W, \mathcal{W})$ *denote a measurable space. A* $(\mathcal{V}, \mathcal{W})$*-measurable mapping* $\chi\colon V \to W$ *is called a* sufficient statistic *if for each* $B \in \mathcal{V}$ *exists a* $(\mathcal{W}, \mathcal{B}(\mathbb{R}))$*-measurable mapping* $h_B\colon W \to \mathbb{R}$ *with* $\int_V 1_B \, dp_\gamma = \int_W h_B \,\mathrm{d}{p_\gamma}^{\chi}$ *for all* $\gamma \in \Gamma$.

Roughly speaking, the statistician observes a value $v \in V$ which he interprets as a realization of a random variable with unknown distribution η. His goal is to decide whether η belongs to the null or the alternative hypothesis. Applying a test $\Psi\colon V \to \mathbb{R}$ means that the null hypothesis is rejected iff the test value $\Psi(v) \in [0,1]$ lies in a specified rejection area. If $\Psi(V) \in \{0,1\}$ the test is called *deterministic*. An elementary example was discussed in 2.38. If $\chi\colon V \to W$ is a sufficient statistic then for each test $\Psi\colon V \to [0,1]$ there exists a mapping $h_\Psi\colon W \to \mathbb{R}$ with $\int_V \Psi \,\mathrm{d}p_\gamma = \int_W h_\Psi \,\mathrm{d}{p_\gamma}^{\chi}$ for all $\gamma \in \Gamma$ (cf. Section 2.2). In particular, without loss of information one may consider the transformed test problem $\left(\overline{\mathrm{P}}_{\Gamma_0}, \overline{\mathrm{P}}_{\Gamma\setminus\Gamma_0}\right)$ on W with sample $w_0 := \chi(v_0)$ instead of v_0 and $\bar{p}_\gamma := {p_\gamma}^{\chi}$. In particular, the determination of a powerful test and the computation of its distribution under the admissible hypotheses can completely be transferred to the transformed test problem. Usually, this simplifies the necessary computations. A number of examples of sufficient statistics can be found in [51] and [82], for instance. Readers who are completely unfamiliar with statistics are referred to introductory works ([51, 82] etc.).

Now suppose that the statistician observes a sample $z_1, \dots, z_m \in H$ which he (motivated by the specific random experiment) interprets as realizations of independent but not necessarily identically distributed random variables $Z_1, \dots, Z_m$ with unknown G-invariant distributions $\nu_1, \dots, \nu_m$. Of course, the cases of most practical relevance are $\nu_1 = \cdots = \nu_m$ (one-sample problem) and $\nu_1 = \cdots = \nu_k$, $\nu_{k+1} = \cdots = \nu_m$ (two-sample problem). It will turn out that under weak additional assumptions our symmetry concept can be used to transform the test problem on $H \times \cdots \times H$ into a test problem on $T \times \cdots \times T$ without loss of any information. This confirms our intuitive argumentation from the introduction. Theorem 4.14 gives a precise formulation. For the sake of readability we first introduce some abbreviations.

Definition 4.13. *The m-fold cartesian product $D \times \cdots \times D$ of a set D will be abbreviated with D^m. Similarly, $\mathcal{V}^m := \mathcal{V} \otimes \cdots \otimes \mathcal{V}$ if $\mathcal{V}$ is a σ-algebra and $\mathcal{M}^1(V,\mathcal{V})^m := \{\tau_1 \otimes \cdots \otimes \tau_m \mid \tau_j \in \mathcal{M}^1(V,\mathcal{V})\}$. If $\tau_j = \tau$ for all $j \le m$ then τ^m stands for $\tau_1 \otimes \cdots \otimes \tau_m$. The m-fold product $\chi \times \cdots \times \chi \colon V_1^m \to V_2^m$ of a mapping $\chi \colon V_1 \to V_2$ is denoted with χ^m.*

Theorem 4.14. (i) *Suppose that $\psi \colon H \to T$ is a $(B(H)_F, \mathcal{B}(T))$-measurable mapping with $\varphi \circ \iota \circ \psi(h) \in E_G(\{h\})$ for all $h \in H$. Then*

$$\psi^m \colon H^m \to T^m, \quad \psi^m(h_1, \dots, h_m) := (\psi(h_1), \dots, \psi(h_m)) \tag{4.14}$$

is a sufficient statistic for all test problems $\left(\mathrm{P}_{\Gamma_0}, \mathrm{P}_{\Gamma \setminus \Gamma_0}\right)$ on H^m with $\mathrm{P}_\Gamma \subseteq \mathcal{M}^1_G(H)^m$.
Consequently, for each $(\mathcal{B}(H^m), \mathcal{B}(\mathbb{R}))$-measurable function $\Psi \colon H^m \to [0,1]$ there exists a $(\mathcal{B}(T^m), \mathcal{B}(\mathbb{R}))$-measurable mapping $h_\Psi \colon T^m \to \mathbb{R}$ with

$$\int_{H^m} \Psi(\mathbf{h})\, p_\gamma(\mathrm{d}\mathbf{h}) = \int_{H^m} h_\Psi \circ \psi^m(\mathbf{h})\, p_\gamma(\mathrm{d}\mathbf{h}) = \int_{T^m} h_\Psi(\mathbf{t})\, {p_{\gamma_v}}^{\psi^m}(\mathrm{d}\mathbf{t}) \tag{4.15}$$

for all $\gamma \in \Gamma$.
(ii) Suppose that $R_G \subseteq T$ is a measurable section and that the corresponding mapping $\psi \colon H \to T$ is $(\mathcal{B}(H)_F, \mathcal{B}(T))$-measurable (cf. Theorem 4.11). Assume further that $p_\gamma = \nu_1 \otimes \cdots \otimes \nu_m$ and $\nu_j = (\mu_{(\mathrm{S})} \otimes \tau_j)^\varphi$ with $\tau_j(R_G) = 1$ for all $j \le m$. Then

$$(p_{\gamma_v})^{\psi^m} = \tau_1 \otimes \cdots \otimes \tau_m. \tag{4.16}$$

Proof. Theorem 2.44, the definition of sufficiency and Lemma 2.43(ii) □

Theorem 2.44 provides a concrete expression for h_Ψ. For many applications this is of subordinate meaning as the determination of a powerful test and the computation of its distribution under the admissible hypotheses can completely be moved from H^m to T^m or even to a measurable section ${R_G}^m \subseteq T^m$. In Sec. 4.5 we will illustrate the situation at two examples.

The product group G^m acts on H^m by $(g_1, \dots, g_m), (h_1, \dots, h_m) \mapsto (g_1 h_1, \dots, g_m h_m)$. The G^m-orbits equal $E_G(\{h_1\}) \times \cdots \times E_G(\{h_m\})$ with

$h_1, \ldots, h_m \in H$. If $\psi: H \to T$ is constant on each G^m-orbit (cf. 4.10(ii)) the mapping $\psi^m: H^m \to T^m$ is maximal invariant for each G^m-invariant test problem $\left(\mathrm{P}_{\Gamma_0}, \mathrm{P}_{\Gamma\setminus\Gamma_0}\right)$ on H^m. Note that the G-invariance of a test problem (cf. Definition 2.40) is weaker than the condition $p_\gamma \in \mathcal{M}^1_G(H)^m$ for all $p_\gamma \in \mathrm{P}_{\Gamma_0}$. Although we will not deepen this aspect in the following we point out that maximal invariant mappings are of particular importance in statistics (see [51], pp. 285 ff., and [20, 21, 36, 80, 81] for instance). Note that a sufficient statistic $\psi^m: H^m \to T^m$ does always exist if φ is continuous.

The Remarks 4.15, 4.16 and 4.17 below summarize significant properties of Haar measures, differentiable manifolds and Lie groups as far as these properties will be needed in the following sections. In some proofs and preparatory lemmata also less elementary properties of Lie groups will be exploited. The facts are explained in the sections where they are needed, and usually the interested reader is pointed to suitable references.

Remark 4.15. (Haar measures) For each locally compact group G' there exists a positive left-invariant measure $\mu_{G';l} \in \mathcal{M}(G')$, i.e. $\mu_{G';l}(B) = \mu_{G';l}(gB)$ for all $(g, B) \in G' \times \mathcal{B}(G')$. The *left invariant Haar measure* $\mu_{G';l}$ is unique up to a scalar factor. Analogously, there exists a *right invariant Haar measure* $\mu_{G';r}$ with corresponding properties. In general, left invariant Haar measures are not right invariant and, vice versa, right invariant Haar measures are not left invariant. If $\mu_{G';l} = \mathrm{const}\cdot\mu_{G';r}$ the group G' is said to be *unimodular*, and we simply speak of the *Haar measure* which we denote with $\mu_{G'}$. In particular, abelian groups and compact groups are unimodular. The most familiar unimodular group is $(\mathbb{R}, +)$, the real line equipped with the usual addition of real numbers. Its Haar measure is the Lebesgue measure which is invariant under translation. If G' is not compact then $\mu_{G';l}(G') = \mu_{G';r}(G') = \infty$. If G' is compact all Haar measures are finite, and $\mu_{G'}$ denotes the unique Haar probability measure. Due to their outstanding symmetry Haar measures are of particular importance. For a comprehensive treatment of Haar measures we refer the interested reader to [54].

Remark 4.16. (Differentiable manifolds) An n-dimensional *differentiable manifold* N is a second countable Hausdorff space with a maximal *differentiable atlas*. This atlas consists of a family of *charts* $(U_\alpha, \chi_\alpha)_{\alpha\in I}$ where $U_\alpha \subseteq N$ and $\chi_\alpha(U_\alpha) \subseteq \mathbb{R}^n$ are open subsets, and $\chi_\alpha: U_\alpha \to \chi_\alpha(U_\alpha)$ is a homeomorphism between open subsets for each $\alpha \in I$ where I denotes an index set. Moreover, $\bigcup_{\alpha\in I} U_\alpha = N$, and whenever $U_\alpha \cap U_\beta \neq \emptyset$ the composition

$$\chi_\beta \circ \chi_\alpha^{-1}: \chi_\alpha(U_\alpha \cap U_\beta) \to \chi_\beta(U_\alpha \cap U_\beta) \tag{4.17}$$

is a diffeomorphism, i.e. $\chi_\beta \circ \chi_\alpha^{-1}$ and its inverse $\chi_\beta \circ \chi_\alpha^{-1}$ are smooth, that is, infinitely often differentiable. A continuous map $f: N_1 \to N_2$ between two manifolds is called *differentiable* at $n_1 \in N_1$ if there exist charts (U_α, χ_α) at n_1 and (U'_β, χ'_β) at $f(n_1)$ so that the composition ${\chi'}_\beta \circ f \circ \chi_\alpha^{-1}: \chi_\alpha(U_\alpha) \to \chi'_\beta(U'_\beta)$ is infinitely often differentiable. The mapping f is called differentiable if it is

differentiable at each $n_1 \in N_1$. A differentiable bijective mapping $f\colon N_1 \to N_2$ between n-dimensional manifolds N_1 and N_2 is called a *diffeomorphism* if its inverse $f^{-1}\colon N_2 \to N_1$ is also differentiable.

Assume that $f\colon N_1 \to N_2$ has a positive Jacobian at n_1, i.e. that $|\det D(\chi'_\beta \circ f \circ \chi_\alpha^{-1})|(\chi_\alpha(n_1)) > 0$ for some charts (U_α, χ_α) on N_1 and (U_β, χ'_β) on N_2 where 'D' stands for the differential. Then there exists a neighbourhood U of n_1 for which the restriction $f_{|U}$ is a diffeomorphism ([81], Theorem 3.1.1). Note that this is equivalent to saying that the differential $Df(n_1)$ induces a bijective linear mapping between the tangential spaces at n_1 and $f(n_1)$ ([81], Theorem 3.3.1). If $f\colon N_1 \to N_2$ is bijective and if these equivalent conditions are fulfilled for each $n_1 \in N_1$ then f is a diffeomorphism.

Remark 4.17. (Lie groups) Roughly speaking, a *Lie group* is a topological group with differentiable group multiplication and inversion. To each Lie group G' there corresponds a *Lie algebra* LG' which is unique up to an isomorphism. A Lie algebra is a vector space equipped with a skew-symmetric bilinear mapping $[\cdot,\cdot]\colon LG' \times LG' \to LG'$ the so-called *Lie bracket*. The Lie bracket fulfils the Jacobi identity, i.e.

$$[[x,y],z] + [[z,x],y] + [[y,z],x] = 0 \qquad \text{for all } x,y,z \in LG' \tag{4.18}$$

(see, e.g. [40]). Assume that $V' \subseteq LG'$ is a neighbourhood of $0 \in LG'$. If V' is sufficiently small the *exponential map* $\exp_{G'}\colon LG' \to G'$ maps V' diffeomorphic onto a neighbourhood of the identity element $e_{G'} \in G'$. Using the exponential function many problems on G' can be transferred to equivalent problems on LG' where the latter can be treated with techniques from linear algebra. The best-known Lie groups are $\mathbb{R}$ and diverse matrix groups, in particular the *general linear group* $\mathrm{GL}(n)$ or closed subgroups of $\mathrm{GL}(n)$ as the *orthogonal group* $\mathrm{O}(n)$ and the *special orthogonal group* $\mathrm{SO}(n)$. We recall that the $\mathrm{GL}(n)$ consists of all real invertible $(n \times n)$-matrices while $\mathrm{O}(n) = \{\boldsymbol{M} \in \mathrm{GL}(n) \mid \boldsymbol{M}^{-1} = \boldsymbol{M}^t\}$ and $\mathrm{SO}(n) = \{\boldsymbol{T} \in \mathrm{O}(n) \mid \det(\boldsymbol{T}) = 1\}$. The Lie algebra of $\mathrm{GL}(n)$ can be identified with $\mathrm{Mat}(n,n)$, the vector space of all real $(n \times n)$-matrices, where the Lie bracket is given by the commutator $[x,y] = xy - yx$. The Lie algebras of $\mathrm{O}(n)$ and $\mathrm{SO}(n)$ are isomorphic to the Lie subalgebra $\mathrm{so}(n) := \{\boldsymbol{M} \in \mathrm{Mat}(n,n) \mid \boldsymbol{M}^t = -\boldsymbol{M}\}$ of $\mathrm{Mat}(n,n)$. For these matrix groups the exponential map is defined by the power series $\exp_{G'}(x) = \sum_{j=0}^{\infty} x^j/j!$ (cf. Remark 4.35 and Section 4.4).

It follows immediately from the definition that Borel measures on compact spaces are finite. Since each finite (non-zero) measure is a scalar multiple of a unique probability measure results on probability measures can immediately be transferred to finite measures. Probability measures are of particular significance for stochastic simulations and statistical applications. Therefore we will only consider probability measures in the following sections if the respective spaces are compact. For non-compact spaces we will consider all Borel measures. In the context of stochastic simulations we also use the term 'distribution' in place of 'probability measure'.

Definition 4.18. *The terms λ, λ_n and λ_C denote the* Lebesgue measure *on $\mathbb{R}$, the n-dimensional Lebesgue measure on $\mathbb{R}^n$ and the restriction of λ_n to a Borel subset $C \subseteq \mathbb{R}^n$, i.e. $\lambda_C(B) = \lambda_n(B)$ for $B \in \mathcal{B}(C) = \mathcal{B}(\mathbb{R}^n) \cap C$. Further, $N(\mu, \sigma^2)$ denotes the* normal distribution *(or* Gauss distribution*) with mean μ and variance σ^2 while $\varepsilon_m \in \mathcal{M}^1(M)$ stands for the* Dirac measure *with total mass on the single element $m \in M$.*
If G is a group (resp. a vector space) then $G' \leq G$ means that G' is a subgroup (resp. a vector subspace) of G. Matrices are denoted with bold capital letters, its components with the same small non-bold letters (e.g. $\mathbf{M} = (m_{ij})_{1 \leq i,j \leq n}$). Similarly, the components of $\mathbf{v} \in \mathbb{R}^n$ are denoted with $v_1, \ldots, v_n$. As usual, $\mathrm{Mat}(n, m)$ denotes the vector space of all real $(n \times m)$-matrices, and the trace of $\boldsymbol{M} \in \mathrm{Mat}(n, n)$ is given by $\mathrm{tr}(\boldsymbol{M}) := \sum_{j=1}^{n} m_{jj}$. Its determinant is denoted with $\det(\boldsymbol{M})$. Further, 1_n stands for the n-dimensional identity matrix.

4.2 Equidistribution on the Grassmannian Manifold and Chirotopes

In this section we elaborate an example which is not of type (G, S, T, H, φ). This example is yet considered as it also exploits equivariant mappings and uses the transport of invariant measures. It meets all requirements on appropriate simulation algorithms which we have worked out in Chapter 3. In a first step we derive an efficient algorithm for the simulation of the equidistribution on the Grassmannian manifold, that is, for the simulation of the unique $\mathrm{O}(n)$-invariant probability measure. We will briefly point out the advantages of this approach and show how this result can be used to confirm a conjecture of Goodman and Pollack from the field of combinatorial geometry. In fact, this goal had been the reason to work out efficient algorithms for the generation of equidistributed pseudorandom elements on the Grassmannian manifold. For further details concerning the simulation part we refer the interested reader to [64], Chaps. F and G, and [68]. Combinatorial and geometrical aspects are dealt in [8], for example.

For $0 < m < n$ let $\mathcal{G}_{n,m}^{\mathbb{R}}$ denote the set of all m-dimensional vector subspaces of the $\mathbb{R}^n$. For the moment let $H_m := \{(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m) \in \mathbb{R}^n \times \cdots \times \mathbb{R}^n \mid \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m \text{ are linear independent}\}$ while $sp\colon H_m \to \mathcal{G}_{n,m}^{\mathbb{R}}$ maps each m-tupel $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m) \in H_m$ onto the m-dimensional subspace of $\mathbb{R}^n$ which is spanned by its components. Clearly, the mapping sp is surjective. Equipped with the quotient topology, that is, equipped with the finest (i.e. largest) topology for which the mapping $sp\colon H_m \to \mathcal{G}_{n,m}^{\mathbb{R}}$ is continuous, $\mathcal{G}_{n,m}^{\mathbb{R}}$ is a real manifold.

Definition 4.19. *As usual, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ denote the* unit vectors *in $\mathbb{R}^n$. Further, $\mathcal{G}_{n,m}^{\mathbb{R}}$ stands for the* Grassmannian manifold *whereas $W_0 \in \mathcal{G}_{n,m}^{\mathbb{R}}$ is the m-dimensional vector subspace of $\mathbb{R}^n$ which is spanned by $\{\mathbf{e}_1, \ldots, \mathbf{e}_j\}$.*

The mapping

$$\Theta\colon \mathrm{O}(n) \times \mathcal{G}_{n,m}^{\mathbb{R}} \to \mathcal{G}_{n,m}^{\mathbb{R}} \qquad \Theta(\boldsymbol{T}, W) := \boldsymbol{T}W \tag{4.19}$$

defines a transitive group action on $\mathcal{G}_{n,m}^{\mathbb{R}}$ since for any $W', W'' \in \mathcal{G}_{n,m}^{\mathbb{R}}$ an orthonormal basis of W' can be transformed into an orthonormal basis of W'' by the left multiplication with a suitable orthogonal matrix. Hence $\mathcal{G}_{n,m}^{\mathbb{R}}$ is a homogeneous $\mathrm{O}(n)$-space, and by 4.6(iii) there exists a unique $\mathrm{O}(n)$-invariant probability measure $\mu_{n,m}$ on $\mathcal{G}_{n,m}^{\mathbb{R}}$. We point out that the product group $\mathrm{O}(m) \times \mathrm{O}(n-m) \leq \mathrm{O}(n)$ is the isotropy group of $W_0 \in \mathcal{G}_{n,m}^{\mathbb{R}}$. Viewed as a $\mathrm{O}(n)$-space $\mathcal{G}_{n,m}^{\mathbb{R}}$ is isomorphic to the factor space $\mathrm{O}(n)/\left(\mathrm{O}(m) \times \mathrm{O}(n-m)\right)$ on which $\mathrm{O}(n)$ acts by left multiplication (cf. [54], p. 128). In particular, its dimension equals

$$\begin{aligned} \dim\left(\mathcal{G}_{n,m}^{\mathbb{R}}\right) &= \dim\left(\mathrm{O}(n)\right) - \dim\left(\mathrm{O}(m)\right) - \dim\left(\mathrm{O}(n-m)\right) \\ &= n(n-1)/2 - m(m-1)/2 - (n-m)(n-m-1)/2 = m(n-m). \end{aligned} \tag{4.20}$$

Even for small parameters n and m the dimension of $\mathcal{G}_{n,m}^{\mathbb{R}}$ is rather large. This rules simulation algorithms out which are based on chart mappings. Instead, we intend to use Theorem 4.6 with $G = \mathrm{O}(n)$ and $M_2 = \mathcal{G}_{n,m}^{\mathbb{R}}$. We need a suitable $\mathrm{O}(n)$-space M_1 and an equivariant mapping $\pi\colon M_1 \to M_2$. Then we have to choose an $\mathrm{O}(n)$-invariant distribution on M_1 which is easy to simulate. However, this plan requires some preparatory work.

Definition 4.20. *If V is a real vector space with scalar product $(\cdot,\cdot)\colon V \times V \to \mathbb{R}$ then $\|v\| := \sqrt{(v,v)}$ denotes the* norm *of v. A mapping $\chi\colon V \to V$ is* orthogonal *if $(\chi(v), \chi(v)) = (v,v) =: \|v\|^2$ for all $v \in V$. Further, the group of all orthogonal mappings is denoted with $\mathrm{O}(V)$.*

A function $h\colon V \to \mathbb{R}$ is said to be radially symmetric *if $h(v)$ does only depend on the norm of its argument. Further,* $\mathrm{Mat}(n,m)_* := \{\boldsymbol{M} \in \mathrm{Mat}(n,m) \mid \mathrm{rank}(\boldsymbol{M}) = m\}$.

Note that $\mathrm{O}(V)$ is a compact group which acts on V via $(\chi, v) \mapsto \chi(v)$ and that $(\chi(u), \chi(v)) = (u,v)$ for all $u, v \in V$. Any finite-dimensional real vector space V is in particular a locally compact abelian group and hence there exists a measure λ_V on V which is invariant under translation (Remark 4.15). Clearly, the vector space $V = \mathrm{Mat}(n,m)$ is isomorphic to $\mathbb{R}^{nm}$ and hence

$$\lambda_{\mathrm{Mat}(n,m)}\left(\{\boldsymbol{M} \mid a_{ij} \leq m_{ij} \leq b_{ij}\}\right) = \prod_{i,j}(b_{ij} - a_{ij}) \quad \text{if } a_{ij} \leq b_{ij} \text{ for all } (i,j) \tag{4.21}$$

defines a Borel measure on $\mathrm{Mat}(n,m)$ which is invariant under translation. We call $\lambda_{\mathrm{Mat}(n,m)}$ the *Lebesgue measure* on $\mathrm{Mat}(n,m)$.

Lemma 4.21. *(i)* $(\boldsymbol{M},\boldsymbol{N}) := \operatorname{tr}\left(\boldsymbol{M}^t\boldsymbol{N}\right)$ *defines a scalar product on the vector space* $\mathrm{Mat}(n,m)$.
(ii) The mapping

$$\Theta\colon \mathrm{O}(n)\times\mathrm{Mat}(n,m)\to\mathrm{Mat}(n,m),\qquad \Theta(\boldsymbol{T},\boldsymbol{M}) := \boldsymbol{T}\boldsymbol{M} \tag{4.22}$$

defines an $\mathrm{O}(n)$*-action on* $\mathrm{Mat}(n,m)$. *In particular,* $\Theta_{\mathbf{T}}\colon\mathrm{Mat}(n,m)\to\mathrm{Mat}(n,m)$, $\Theta_{\mathbf{T}}(\boldsymbol{M}) := \boldsymbol{T}\boldsymbol{M}$ *induces a linear mapping. More accurately*

$$\{\Theta_{\mathbf{T}} \mid \boldsymbol{T}\in\mathrm{O}(n)\} \le \mathrm{O}(\mathrm{Mat}(n,m)). \tag{4.23}$$

(iii) Suppose that $\nu = h\cdot\lambda_{\mathrm{Mat}(n,m)}\in\mathcal{M}(\mathrm{Mat}(n,m))$ *with a radially symmetric density* $h\colon\mathrm{Mat}(n,m)\to\mathbb{R}$. *Then* $\nu^{\chi}=\nu$ *for all* $\chi\in\mathrm{O}(\mathrm{Mat}(n,m))$. *In other words:* $\nu\in\mathcal{M}_{\mathrm{O}(n)}(\mathrm{Mat}(n,m))$.
(iv) As $\Theta_{\mathbf{T}}(\mathrm{Mat}(n,m)_*)\subseteq\mathrm{Mat}(n,m)_*$ *for all* $\boldsymbol{T}\in\mathrm{O}(n)$ *the subset* $\mathrm{Mat}(n,m)_*$ *is an* $\mathrm{O}(n)$*-space, too.*
(v) Suppose that $\mathrm{pol}\colon\mathbb{R}^k\to\mathbb{R}$ *is a polynomial function. If* $\mathrm{pol}\not\equiv 0$ *then* $\lambda_k(\mathrm{pol}^{-1}(\{0\}))=0$.
(vi) $\lambda_{\mathrm{Mat}(n,m)}(\mathrm{Mat}(n,m)\setminus\mathrm{Mat}(n,m)_*)=0$.

Proof. The statements (i)-(iv) and (vi) are shown in [68], Lemma 3.5 and Theorem 3.6 (i). For (v) see [64], p. 163 or [60], p. 171 (Aufgabe 3). □

Lemma 4.21 provides the preparatory work for the first two main results of this section which correspond with Theorem 3.6 (ii), Theorem 3.7 und Corollary 3.8 in [68].

Theorem 4.22. (i) *The orthogonal group* $\mathrm{O}(n)$ *acts on* $\mathrm{Mat}(n,m)$, *on its* $\mathrm{O}(n)$*-invariant subset* $\mathrm{Mat}(n,m)_*$ *and on* $\mathcal{G}_{n,m}^{\mathbb{R}}$ *via* $(\boldsymbol{T},\boldsymbol{A})\mapsto\boldsymbol{T}\boldsymbol{A}$ *and* $(\boldsymbol{T},\boldsymbol{A})\mapsto\boldsymbol{T}\boldsymbol{A}$ *and* $(\boldsymbol{T},W)\mapsto\boldsymbol{T}W$, *resp. If* $\mathbf{a}_1,\dots,\mathbf{a}_m$ *denote the columns of the matrix* $\boldsymbol{A}$ *and* $span\{\mathbf{a}_1,\dots,\mathbf{a}_m\}$ *the vector subspace spanned by* $\mathbf{a}_1,\dots,\mathbf{a}_m$ *then the mapping*

$$p\colon\mathrm{Mat}(n,m)_*\to\mathcal{G}_{n,m}^{\mathbb{R}}\qquad p(\boldsymbol{A}) := span\{a_1,\dots,a_m\} \tag{4.24}$$

is $\mathrm{O}(n)$*-equivariant. For all* $\eta\in\mathcal{M}^1_{\mathrm{O}(n)}(\mathrm{Mat}(n,m)_*)$ *we have* $\eta^p=\mu_{m,n}$.
(ii) *Let*

$$\bar{p}\colon\mathrm{Mat}(n,m)\to\mathcal{G}_{n,m}^{\mathbb{R}},\qquad \bar{p}(\boldsymbol{A}) := \begin{cases} p(\boldsymbol{A}) & \text{if } \boldsymbol{A}\in\mathrm{Mat}(n,m)_* \\ W_0 & \text{else}\end{cases}. \tag{4.25}$$

Then $\bar{p}$ *is a measurable extension of* p, *and* $\nu^{\bar{p}}=\mu_{n,m}$ *for each* $\nu\in\mathcal{M}^1_{\mathrm{O}(n)}(\mathrm{Mat}(n,m))$ *with* $\nu\big(\mathrm{Mat}(n,m)\setminus\mathrm{Mat}(n,m)_*\big)=0$.

Proof. The assertions in (i) concerning the $\mathrm{O}(n)$-actions have already been verified in 4.21(ii), (iv) and in the preceding paragraphs. As $\boldsymbol{T}p(\boldsymbol{A}) = \boldsymbol{T}\,span\{a_1,\dots,a_m\} = span\{\boldsymbol{T}a_1,\dots,\boldsymbol{T}a_m\} = p(\boldsymbol{T}\boldsymbol{A})$ the mapping p is

$\mathrm{O}(n)$-equivariant, and applying 4.6(iii) completes the proof of (i). By assumption, the restriction $\nu_{|\mathrm{Mat}(n,m)_*} \in \mathcal{M}^1_{\mathrm{O}(n)}(\mathrm{Mat}(n,m)_*)$, and further $\nu\big(\mathrm{Mat}(n,m) \setminus \mathrm{Mat}(n,m)_*\big) = 0$. Consequently, (ii) is an immediate consequence of (i). □

Corollary 4.23. *If the random variables $X_{11}, X_{12}, \ldots, X_{nm}$ are iid $N(0,1)$-distributed then the image measure $\mathrm{N}I(0,1)^n_m$ of*

$$X = \begin{pmatrix} X_{11} & X_{12} & \ldots & X_{1m} \\ X_{21} & X_{22} & \ldots & X_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \ldots & X_{nm} \end{pmatrix}$$

meets the assumptions of 4.22(ii), and the random variable $\bar{p}(X)$ is $\mu_{n,m}$-distributed.

Proof. The image measure of X has radially symmetric Lebesgue density

$$\begin{aligned} h(\boldsymbol{A}) &= \prod_{i=1}^{n}\prod_{j=1}^{m} c\mathrm{e}^{-\frac{1}{2}a_{ij}^2} = c^{nm}\mathrm{e}^{-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}^2} \\ &= c^{nm}\mathrm{e}^{-\frac{1}{2}(\boldsymbol{A},\boldsymbol{A})} = c^{nm}\mathrm{e}^{-\frac{1}{2}\|\boldsymbol{A}\|^2}. \end{aligned}$$

Hence the corollary follows from 4.21(iii), (v) and 4.22(ii). □

We point out that nearly any algorithm which is based on Theorem 4.22 should be appropriate for simulation purposes. In fact, the mapping $\bar{p}$ retracts the essential part of a simulation of $\mu_{n,m}$ from $\mathcal{G}^{\mathbb{R}}_{n,m}$ to the well-known vector space $\mathrm{Mat}(n,m) \cong \mathbb{R}^{nm}$. There are no chart boundaries which could cause problems (cf. Chapter 3), and p does not distinguish particular directions or regions of $\mathrm{Mat}(n,m)_*$. Corollary 4.23 is of particular significance for applications as it decomposes a simulation problem on a high-dimensional manifold into nm independent identical one-dimensional simulation problems for which many well-tried algorithms are known (see e.g. [18]). Clearly, $\dim(\mathcal{G}^{\mathbb{R}}_{n,m}) = m(n-m)$ constitutes a lower bound for the average number of standard random numbers required per pseudorandom element on $\mathcal{G}^{\mathbb{R}}_{n,m}$. Marsaglia's method ([18], p. 235f.), for instance, needs about 1.27 standard random numbers per $N(0,1)$-distributed pseudorandom number in average. Consequently, the simulation algorithm suggested by Corollary 4.23 requires about $1.27nm$ standard random numbers per pseudorandom element $\bar{p}(\widetilde{X})$. This is an excellent value since acceptance-rejection algorithms applied in high-dimensional vector spaces or manifolds (here: $\dim(\mathcal{G}^{\mathbb{R}}_{n,m}) = m(n-m)$) usually are much less efficient. A further advantage of Theorem 4.22 is that one can choose any radially symmetric distribution on $\mathrm{Mat}(n,m)_*$. For example, it is possible to generate the columns of the pseudorandom matrices on $\mathrm{Mat}(n,m)$ iid equidistributed on $\mathrm{S}^{n-1} := \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1\}$ ([68], Remark 3.9 (i)). Note that Theorem 4.22 can also be formulated coordinate-free (cf. [68]) which underlines the universality of the equivariance concept.

Remark 4.24. Although $\mathrm{Prob}(X \notin \mathrm{Mat}(n,m)_*) = 0$ it cannot be excluded that any pseudorandom element may assume values in $\mathrm{Mat}(n,m) \backslash \mathrm{Mat}(n,m)_*$. Instead of mapping them onto W_0 one may alternatively reject those pseudorandom elements as it is common practice in similar situations.

We have reached our first goal. As announced in the introductory paragraph of this section we will sketch an application from the field of combinatorial geometry. Next, we provide some definitions.

Definition 4.25. *For $A \in \mathrm{Mat}(n,m)$ let $[j_1, j_2, \ldots, j_m]_A$ denote the determinant of that $(m \times m)$-submatrix $A_{j_1,\ldots,j_m}$ which consists of the rows $j_1, \ldots, j_m$. Further,* $\mathrm{GF}(3)$ *is the Galois field over $\{-1, 0, 1\}$, and $\mathbb{R}^* := \mathbb{R} \setminus \{0\}$. Finally,* PR^k *and* $\mathrm{PGF}(3)^k$ *denote the projective spaces $(\mathbb{R}^k \setminus \{0\})/\mathbb{R}^*$ and $(\mathrm{GF}(3)^k \setminus \{0\})/\{1, -1\}$, resp. For the remainder of this section we set $F := \{1, \ldots, n\}$. A skew-symmetric mapping $\chi\colon F^m \to \{-1, 0, 1\}$ is called a* chirotope *if for all increasing sequences $j_1, j_2, \ldots, j_{m+1} \in F$ and $k_1, k_2, \ldots, k_{m-1} \in F$ the set*
$\{(-1)^{i-1}\chi(j_1, \ldots, j_{i-1}, j_{i+1}, \ldots, j_{m+1})\chi(j_i, k_1, \ldots, k_{m-1}) \mid 1 \le i \le m+1\}$
either equals $\{0\}$ or is a superset of $\{1, -1\}$.

The *Grassmann-Plücker relations* ([9], pp. 10f.) imply that the mapping

$$\chi_A\colon F^m \to \{-1, 0, 1\} \quad \chi_A(j_1, j_2, \ldots, j_m) := sgn[j_1, j_2, \ldots, j_m]_A \tag{4.26}$$

is a chirotope for each $A \in \mathrm{Mat}(n,m)$.

Definition 4.26. *A chirotope χ is said to be* realizable *if there exists a matrix $A \in \mathrm{Mat}(n,m)$ with $\chi = \chi_A$. An m-tupel $(j_1, j_2, \ldots, j_m) \in F^m$ is* ordered *if $j_1 < j_2 < \cdots < j_m$. We denote the set of all ordered m-tuples in F^m with $\Lambda(F,m)$ and call a chirotope χ* simplicial *if $\chi(J) \in \{1, -1\}$ for all $J \in \Lambda(F,m)$. The set of all realizable chirotopes is denoted by $\mathrm{Re}_\chi(n,m)$ while $\mathrm{Si}_\chi(n,m)$ denotes the set of all simplicial chirotopes. As usually, $sgn\colon \mathbb{R} \to \{1, 0, -1\}$ stands for the signum function.*

Since chirotopes are skew-symmetric functions they are completely determined by their restriction to $\Lambda(F,m)$. In the following we hence identify the chirotope χ with the $\binom{n}{m}$-tuple $\big(\chi(1,2,\ldots,m), \chi(1,\ldots,m-1,m+1), \ldots, \chi(n-m+1, n-m+2, \ldots, n)\big) \in \mathrm{GF}(3)^{\binom{n}{m}}$. In fact, there is a natural mapping which assigns pairs of chiroptopes, that is, elements in $\mathrm{PGF}(3)^{\binom{n}{m}}$ to the m-dimensional subspaces $\mathbb{R}^n$ (cf. 4.28 below). This mapping induces an image measure P_χ of $\mu_{n,m}$ on this projective space. The goal of [8] was to determine those elements where P_χ attains its maximum.

Remark 4.27. (i) Chirotopes are a useful tool to describe the combinatorial structure of geometrical configurations. If we identify the row vectors of $\boldsymbol{A} \in \mathrm{Mat}(n,m)$ with points $Q_1, Q_2, \ldots, Q_n \in \mathbb{R}^m$ then $\chi_A(j_1, j_2, \ldots, j_m)$ equals

the orientation of $Q_{j_1}, Q_{j_2}, \ldots, Q_{j_m}$. Moreover, definitions and objects from classical geometry can be transferred and generalized to chirotopes. Their combinatorial properties can be used to give alternate proofs for well-known theorems from elementary geometry. We point out that chirotopes can be identified with oriented matroids (see e.g. [61, 50, 15], [64], pp. 134f.).
(ii) In fact, P_χ describes the random combinatorial structure of n points which are randomly (equidistributed) and independently thrown on an m-sphere, namely by the orientation of all subsets with m elements ([68], Remark 4.5).

Theorem 4.28. *Let*

$$\begin{aligned}\Psi_{\mathcal{G}}\colon \mathrm{Mat}(n,m) \to \mathbb{R}^{\binom{n}{m}} \qquad \Psi_{\mathcal{G}}(A) =: ([1,2,\ldots,m]_A,\\ [1,2,\ldots,m-1,m+1]_A,\ldots,[n-m+1,n-m+2,\ldots,n]_A)\end{aligned} \tag{4.27}$$

where the subdeterminants of A are ordered lexicographically. Further, let

$$\mathrm{pr}\colon \mathbb{R}^{\binom{n}{m}} \setminus \{0\} \to \mathrm{PR}^{\binom{n}{m}} \qquad \mathrm{pr}(x) := x\mathbb{R}^* \tag{4.28}$$

while p and the $\mathrm{O}(n)$-actions on $\mathrm{Mat}(n,m)_*$ and $\mathcal{G}_{n,m}^{\mathbb{R}}$ are defined as in 4.22. Then
(i) *For each $\boldsymbol{T} \in \mathrm{O}(n)$ there exists a unique orthogonal matrix $\Gamma_\wedge(\boldsymbol{T}) \in \mathrm{O}\left(\binom{n}{m}\right)$ with*

$$\Gamma_\wedge(\boldsymbol{T})(\Psi_{\mathcal{G}}(\boldsymbol{A})) = \Psi_{\mathcal{G}}(\boldsymbol{TA}) \qquad \text{for each } \boldsymbol{A} \in \mathrm{Mat}(n,m). \tag{4.29}$$

The orthogonal group $\mathrm{O}(n)$ acts on $\mathbb{R}^{\binom{n}{m}} \setminus \{0\}$ via $(\boldsymbol{T}, \mathbf{x}) \mapsto \Gamma_\wedge(\boldsymbol{T})\mathbf{x}$, and $(\boldsymbol{T}, \mathbf{x}\mathbb{R}^*) \mapsto \boldsymbol{T}.\mathbf{x}\mathbb{R}^* := (\Gamma_\wedge(\boldsymbol{T})\,\mathbf{x})\mathbb{R}^*$ defines an $\mathrm{O}(n)$-action on $\mathrm{PR}^{\binom{n}{m}}$.
(ii) *Let $\bar{\Psi}_{\mathcal{G}}\colon \mathcal{G}_{n,m}^{\mathbb{R}} \to \mathrm{PR}^{\binom{n}{m}}$ be defined by $\bar{\Psi}_{\mathcal{G}}(W') := \Psi_{\mathcal{G}}(\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m)\mathbb{R}^*$ for $W' \in \mathcal{G}_{n,m}^{\mathbb{R}}$ where $\mathbf{w_1}, \ldots, \mathbf{w_m}$ denotes any basis of W'. Then $\bar{\Psi}_{\mathcal{G}}$ is a well-defined injective mapping. With respect to the $\mathrm{O}(n)$-actions defined above the mappings $p, \Psi_{\mathcal{G}}, \bar{\Psi}_{\mathcal{G}}$ and pr are $\mathrm{O}(n)$-equivariant.*
(iii) *Let the mappings $\Upsilon\colon \mathbb{R}^{\binom{n}{m}} \setminus \{0\} \to \mathrm{GF}(3)^{\binom{n}{m}} \setminus \{0\}$, $\bar{\Upsilon}\colon \mathrm{PR}^{\binom{n}{m}} \to \mathrm{PGF}(3)^{\binom{n}{m}}$ and $\mathrm{pr}_3\colon \mathrm{GF}(3)^{\binom{n}{m}} \setminus \{0\} \to \mathrm{PGF}(3)^{\binom{n}{m}}$ be given by*

$$\begin{aligned}&\Upsilon(\mathbf{x_1}, \ldots, \mathbf{x}_{\binom{\mathbf{n}}{\mathbf{m}}}) := (sgn(\mathbf{x_1}), \ldots, sgn(\mathbf{x}_{\binom{\mathbf{n}}{\mathbf{m}}}),\\ &\bar{\Upsilon}(\mathbf{x}\mathbb{R}^*) := \Upsilon(\mathbf{x})\mathrm{GF}(3)^* \quad \text{and} \quad \mathrm{pr}_3(q) := \{q, -q\}\end{aligned} \tag{4.30}$$

where $(\mathbf{x_1}, \ldots, \mathbf{x}_{\binom{\mathbf{n}}{\mathbf{m}}}), \mathbf{x}$ and q denote elements of the respective domains. Then the following diagram is commutative.

$$\begin{array}{ccc} \mathrm{Mat}(n,m)_* & \xrightarrow{p} & \mathcal{G}_{n,m}^{\mathbb{R}} \\ \Psi_{\mathcal{G}} \big\downarrow & & \big\downarrow \bar{\Psi}_{\mathcal{G}} \\ \mathbb{R}^{\binom{n}{m}} \setminus \{0\} & \xrightarrow{\mathrm{pr}} & \mathrm{PR}^{\binom{n}{m}} \\ \Upsilon \big\downarrow & & \big\downarrow \bar{\Upsilon} \\ \mathrm{GF}(3)^{\binom{n}{m}} \setminus \{0\} & \xrightarrow{\mathrm{pr}_3} & \mathrm{PGF}(3)^{\binom{n}{m}} \end{array} \tag{4.31}$$

(iv) *Let* $P_\chi := (\mu_{n,m})^{\bar{\Upsilon}\circ\bar{\Psi}_{\mathcal{G}}}$. *Then*

$$P_\chi = \nu^{\mathrm{pr}_3\circ\Upsilon\circ\Psi_{\mathcal{G}}} \qquad \text{for each } \nu \in \mathcal{M}^1_{\mathrm{O}(n)}(\mathrm{Mat}(n,m)_*) \tag{4.32}$$

(v) *In particular,*

$$P_\chi(\{q,-q\}) \begin{cases} > 0 & \text{if } \{q,-q\} \in (\mathrm{Re}_\chi(n,m) \cap \mathrm{Si}_\chi(n,m)) \cdot \{1,-1\} \\ = 0 & \text{else.} \end{cases} \tag{4.33}$$

Proof. Theorem 4.28 is an extract of Theorem 4.4 in [68]. The essential part of its proof deals with the verification of the statements concerning the $\mathrm{O}(n)$-action on $\mathbb{R}^{\binom{n}{m}} \setminus \{0\}$ and the equivariance of $\Psi_{\mathcal{G}}$. As it is apart from the scope of this book we do not give the complete proof. We merely remark that the pair $\left(\mathbb{R}^{\binom{n}{m}}, \Psi_{\mathcal{G}}\right)$ can be identified with the m-th exterior power of $\mathbb{R}^n$ ([68], Theorem 3.12). □

Based on geometrical considerations Goodman and Pollack conjectured that P_χ attains its maximum at $\{(1,1,\ldots,1),(-1,-1,\ldots,-1)\} \in \mathrm{PGF}(3)^{\binom{n}{m}}$. Theorem 4.28 suggests the following procedure to check this conjecture: Simulate any distribution $\nu \in \mathcal{M}^1_{\mathrm{O}(n)}(\mathrm{Mat}(n,m)_*)$ and apply the mapping $\mathrm{pr}_3 \circ \Upsilon \circ \Psi_{\mathcal{G}}$ to the generated pseudorandom elements. Exploiting the commutativity of diagram (4.31) avoids time-consuming arithmetic operations on the high-dimensional Grassmannian manifold $\mathcal{G}^{\mathbb{R}}_{n,m}$ and the projective space $\mathrm{PR}^{\binom{n}{m}}$. Moreover, due to 4.28(v) one may restrict his attention to $\operatorname{supp} P_\chi = (\mathrm{Re}_\chi(n,m) \cap \mathrm{Si}_\chi(n,m)) \cdot \{1,-1\}$. However, this approach is hardly practically feasible for $(n,m) = (8,4)$ which was the case of particular interest ([64, 8]) since $|\operatorname{supp} P_\chi| \approx 12 \cdot 10^9$ which is very large. Without further insights one had to generate a giantic number of pseudorandom elements to obtain a reliable estimator for P_χ.

Applying Theorem 4.6 again yields the decisive breakthrough. Since non-trivial $\mathrm{O}(n)$-actions on $\mathrm{GF}(3)^{\binom{n}{m}} \setminus \{0\}$ and $\mathrm{PGF}(3)^{\binom{n}{m}}$ do not exist which could supply further information on P_χ we have to search for another group which acts on all spaces appearing in diagram (4.31).

Definition 4.29. *We call a matrix* $\boldsymbol{M} \in \mathrm{O}(k)$ monoidal *if it maps the set* $\{\pm e_1, \ldots, \pm e_k\}$ *onto itself. The set of all monoidal matrices of rank* k *is denoted with* $\mathrm{Mon}(k)$.

In each row and each column of a monoidal matrix there is exactly one non-zero element which equals 1 or -1. It can easily be checked that $\mathrm{Mon}(k) \leq \mathrm{O}(k)$ and $\Gamma_\wedge(\mathrm{Mon}(n)) \leq \mathrm{Mon}\left(\binom{n}{m}\right)$. In particular, $\Gamma_\wedge(\boldsymbol{M})$ maps hyperquadrants in $\mathbb{R}^{\binom{n}{m}}$ onto hyperquadrants. This induces $\mathrm{Mon}(n)$-actions on $\mathrm{GF}(3)^{\binom{n}{m}} \setminus \{0\}$ and $\mathrm{PGF}(3)^{\binom{n}{m}}$. Clearly, as $\mathrm{Mon}(n) \leq \mathrm{O}(n)$ each $\mathrm{O}(n)$-space is also a $\mathrm{Mon}(n)$-space, and each $\mathrm{O}(n)$-equivariant mapping is also $\mathrm{Mon}(n)$-equivariant. Theorem 4.30 coincides with Theorem 4.7 in [68].

Theorem 4.30. (i) $\mathrm{Mon}(n) \leq \mathrm{O}(n)$ *and* $\Gamma_\wedge(\mathrm{Mon}(n)) \leq \mathrm{Mon}(\binom{n}{m})$.
(ii) *The spaces* $\mathrm{GF}(3)^{\binom{n}{m}} \setminus \{0\}$ *and* $\mathrm{PGF}(3)^{\binom{n}{m}}$ *become* $\mathrm{Mon}(n)$*-spaces via*

$$(M, q) \mapsto M.q := \Upsilon(\Gamma_\wedge(M)x_q) \quad \textit{and} \quad (M, \{q, -q\}) \mapsto \mathrm{pr}_3(M.q) \tag{4.34}$$

for $(M, q) \in \mathrm{Mon}(n) \times \left(\mathrm{GF}(3)^{\binom{n}{m}} \setminus \{0\}\right)$ *and* $x_q \in \Upsilon^{-1}(\{q\})$.
(iii) *The diagram (4.31) is commutative and all mappings are* $\mathrm{Mon}(n)$*-equivariant.*
(iv) *The* $\mathrm{Mon}(n)$*-action divides* $\mathrm{PGF}(3)^{\binom{n}{m}}$ *into orbits which are called* reorientation classes*. The probability measure* P_χ *is equidistributed on each orbit. This statement is also true for* $\operatorname{supp} P_\chi = (\mathrm{Re}_\chi(n, m) \cap \mathrm{Si}_\chi(n, m)) \cdot \{1, -1\}$ *instead of* $\mathrm{PGF}(3)^{\binom{n}{m}}$.

Proof. The first assertion of (i) is obvious. To see the second let temporarily $\mathcal{E} := \{\boldsymbol{T} \in \mathrm{Mat}(n, m) \mid \boldsymbol{T} = (\mathbf{e_{j_1}}, \ldots, \mathbf{e_{j_m}}), 1 \leq j_1 < \cdots, j_m \leq n\}$. We first note that $\mathcal{E}' := \Psi_\mathcal{G}(\mathcal{E})$ equals the standard vector basis of $\mathbb{R}^{\binom{n}{m}}$. In particular, $\Gamma_\wedge(\boldsymbol{M})(\Psi_\mathcal{G}(\mathcal{E})) = \Psi_\mathcal{G}(\boldsymbol{M}\mathcal{E}) \subseteq \mathcal{E}' \cup -\mathcal{E}'$ for each $\boldsymbol{M} \in \mathrm{Mon}(n)$ which proves $\Gamma_\wedge(\mathrm{Mon}(n)) \subseteq \mathrm{Mon}(\binom{n}{m})$. The $\mathrm{Mon}(n)$-invariance of $\Psi_\mathcal{G}$ implies that $\Gamma_\wedge(\mathrm{Mon}(n))$ is a group which completes the proof of (i). As $\Gamma_\wedge(\boldsymbol{M})$ is monoidal the left multiplication with $\Gamma_\wedge(\boldsymbol{M})$ maps hyperquadrants bijectively onto hyperquadrants. As a consequence the $\mathrm{Mon}(n)$-action on $\mathrm{GF}(3)^{\binom{n}{m}} \setminus \{0\}$ is well-defined. As $\Gamma_\wedge(\boldsymbol{M})(-\mathbf{x}) = -\Gamma_\wedge(\boldsymbol{M})(\mathbf{x})$ the $\mathrm{Mon}(n)$-action on $\mathrm{PGF}(3)^{\binom{n}{m}}$ is well-defined, too. Recall that the mappings from the upper half of diagram (4.31) are $\mathrm{O}(n)$-equivariant. These mappings are in particular $\mathrm{Mon}(n)$-equivariant, and it follows immediately from their definition that the mappings from the lower half are also $\mathrm{Mon}(n)$-equivariant. To prove (iv) it is sufficient to show that $\mathrm{Re}_\chi(n, m) \cap \mathrm{Si}_\chi(n, m)$ is $\mathrm{Mon}(n)$-invariant: If $q \in \mathrm{Si}_\chi(n, m)$ each $x_q \in \Upsilon^{-1}(q)$ has only non-zero components. As $\Gamma_\wedge(\boldsymbol{M})$ is monoidal the same is true for $\Gamma_\wedge(\boldsymbol{M})(x_q)$, and $\boldsymbol{M}.q$ is simplicial. If q is realizable $q = \Upsilon \circ \Psi_\mathcal{G}(\boldsymbol{A})$ for a suitable matrix $\boldsymbol{A} \in \mathrm{Mat}(n, m)$. From (iii) we conclude $\boldsymbol{M}.q = \Upsilon \circ \Psi_\mathcal{G}(\boldsymbol{MA}) \in \mathrm{Re}_\chi(n, m)$ which completes the proof of Theorem 4.30. □

Theorem 4.30 yields the desired improvement as it reduces the cardinality of the simulation problem drastically. In a first step the cardinality of the realizable reorientation classes $K_1, K_2, \ldots, K_s \subseteq (\mathrm{Re}_\chi(n, m) \cap \mathrm{Si}_\chi(n, m)) \cdot \{1, -1\}$ have to be determined which is a pure combinatorial problem. Then one has to generate pseudorandom elements on $\mathrm{Mat}(n, m)_*$ to obtain estimators for the class probabilities $P_\chi(K_1), \ldots, P_\chi(K_s)$. For $(n, m) = (8, 4)$ we have $s = 2604$, that is, we have to find a maximum of 2604 values instead of $12 \cdot 10^9$. This is a reduction by the factor $|\operatorname{supp} P_\chi|/s \approx 5 \cdot 10^6$. We ran three simulations where we altogether generated 1.2 million random chirotopes and used two different algorithms ([64], pp. 152 ff.). In all three cases the maximal relative frequency was attained at the pair $(\{1, \ldots, 1\}, \{-1, \ldots, -1\})$, or more precisely, at each

element of this reorientation class. The simulation results may be viewed as a serious indicator that Goodman and Pollack's conjecture concerning the location of the maximum of P_χ is true for $(n, m) = (8, 4)$ ([64], pp. 151 ff.). To be precise, we obtained

$$P_\chi(\{(1, 1, \ldots, 1), (-1, -1, \ldots, -1)\}) \approx 7.0 \cdot 10^{-10}. \tag{4.35}$$

For details the interested reader is referred to [64], pp. 149–158 and [8].

Remark 4.31. Exploiting specific properties of normally distributed random variables and Theorem 3.2 of [77] one can find a coset T of an $(m(n-m))$-dimensional vector subspace T' of $\mathrm{Mat}(n, m)$ and a probability measure τ on T with $\tau^{\Upsilon \circ \Psi_\mathcal{G}}(\{q, -q\}) = P_\chi(\{q, -q\})$. We point out that τ is not $\mathrm{O}(n)$-invariant. Its simulation, however, needs less standard random numbers per pseudorandom matrix than that of $\mathrm{N}I(0, 1)_m^n$. We used this observation for our simulations. As it does not belong to the scope of this monograph we refer the interested reader to [64], pp. 145f., or [8] for a detailed treatment. We merely remark that $\bar{\Upsilon} \circ \bar{\Psi}_\mathcal{G} \circ p(\boldsymbol{A}) = \bar{\Upsilon} \circ \bar{\Psi}_\mathcal{G} \circ p(\boldsymbol{AS})$ for all $\boldsymbol{A} \in \mathrm{Mat}(n, m)$ and $\boldsymbol{S} \in \mathrm{GL}(m)$. Also the multiplication of single rows of A with positive scalars does not change its value under $\bar{\Upsilon} \circ \bar{\Psi}_\mathcal{G} \circ p$.

4.3 Conjugation-invariant Probability Measures on Compact Connected Lie Groups

Section 4.3 covers a whole class of important examples: the special orthogonal group $\mathrm{SO}(n)$, the group of all unitary matrices $\mathrm{U}(n)$ and all closed connected subgroups of $\mathrm{SO}(n)$ und $\mathrm{U}(n)$, for instance. The special cases $G = \mathrm{SO}(n)$ and in particular $G = \mathrm{SO}(3)$ will be investigated in the Sections 4.4 and 4.5. In this section we treat the most general case where G is a compact connected Lie group. At first sight this may appear to be unnecessarily laborious and far from being useful for applications. However, the Theorems 4.39 and 4.42 reveal useful insights which do not need to be proved for each special case separately. Of course, as G is compact each Borel measure on G is finite and a scalar multiple of a probability measure. We hence restrict our attention to the probability measures on G. Remark 4.37 provides some useful information on compact connected Lie groups.

Definition 4.32. *A topological space M is called* connected *if it cannot be represented as a disjoint union of two non-empty open subsets. A Lie group T' which is isomorphic to $(\mathbb{R}/\mathbb{Z})^k \cong (\mathrm{S}^1)^k$ (cf. Example 2.13(iii)) is said to be a k-*dimensional torus. *A torus subgroup $T \leq G$ is* maximal *if there is no other torus subgroup T'' of G which is a proper supergroup of T.*

Example 4.33. The spaces $\mathbb{R}$, $\mathrm{Mat}(n, m)$, $\mathrm{GL}(n)$ and $\mathrm{SO}(n)$ are connected. The orthogonal group $\mathrm{O}(n)$ is not connected. It falls into the components

$\mathrm{SO}(n)$ and $\mathrm{O}(n) \setminus \mathrm{SO}(n)$. Any discrete space with more than one element is not connected.

We mention that each compact connected Lie group G has a maximal torus and that this maximal torus is unique up to conjugation ([13], pp. 156 and 159 (Theorem (1.6))). Thus all maximal tori have the same dimension. In the following we fix a maximal torus T. In particular $k = \dim(T) \leq \dim(G) = n$.

We equip the factor space $G/T := \{gT \mid g \in G\}$ with the quotient topology , that is, with the finest (i.e. largest) topology for which the mapping $G \to G/T$, $g \mapsto gT$ is continuous. As T is closed the factor space G/T is an $(n-k)$-dimensional manifold ([13], p. 33 (Theorem 4.3)). The group G acts transitively on G/T by left multiplication, i.e. by $(g', gT) \mapsto g'gT$. The group G acts on itself by conjugation, i.e. by $(g', g) \mapsto g'gg'^{-1}$.

Definition 4.34. *We suggestively call* $\nu \in \mathcal{M}^1(G)$ conjugation-invariant *if* $\nu(B) = \nu(gBg^{-1})$ *for all* $(g, B) \in G \times \mathcal{B}(G)$. *Similarly, a function* $f\colon G \to \mathbb{R}$ *is said to be* conjugation-invariant *if* $f(g'gg'^{-1}) = f(g)$ *for all* $g, g' \in G$. *Further,*

$$q\colon G/T \times T \to G \qquad\qquad q(gT, t) := gtg^{-1}. \tag{4.36}$$

We point out that the mapping q is well-defined: As T is commutative $(gt')t(gt')^{-1} = gt'tt'^{-1}g^{-1} = gtg^{-1}$ for each $t' \in T$. In other words: $gtg^{-1} = g'tg'^{-1}$ for all $g' \in gT$, that is, the value $q(gT, t)$ does not depend on the representative of the coset gT. Clearly, $\mathcal{M}^1_G(G)$ consists of all conjugation-invariant probability measures on G.

Remark 4.35. In [66] we used the less suggestive expression 'con-invariant' in place of 'conjugation-invariant'.

Lemma 4.36. *The* 5*-tupel* $(G, G/T, T, G, q)$ *has Property* $(*)$.

Proof. The spaces G and G/T are manifolds and hence are second countable. In particular, G/T is a Hausdorff space. The subgroup $T \leq G$ is closed and hence compact. Further, T also has a countable base, and G acts transitively on G/T. The mapping q is continuous and hence in particular measurable. Above all, q is surjective ([13], p. 159 (Lemma 1.7)). Finally, $g'q(gT, t)g'^{-1} = g'(gtg^{-1})g'^{-1} = q(g'(gT), t)$, i.e. q is equivariant. □

Consequently, we can apply our symmetry concept from Chapter 2 to the 5-tupel $(G, G/T, T, G, q)$. As the subgroup T is compact there exists a unique Haar probability measure μ_T on T. Weyl's famous integral formula (Theorem 4.39(ii)) provides a smooth μ_T-density h_G for which $\mu_G = (\mu_{(G/T)} \otimes h_G \cdot \mu_T)^q$. The density $h_G\colon G \to \mathbb{R}$ can be expressed as a function of the adjoint representation defined below. The theoretical background is briefly explained in the following remark.

Remark 4.37. The mapping $c(g)\colon G \to G$ denotes the conjugation with $g \in G$, i.e. $(c(g))(g') = gg'g^{-1}$ for all $g, g' \in G$. As in 4.17 the Lie algebra of G is denoted with LG while $\mathrm{Aut}(LG)$ is the vector space of all linear automorphisms on LG. Further, $\mathrm{Ad}(g)$ denotes the differential of $c(g)$ at the identity element e_G. Since LG can canonically be identified with the tangential space at e_G the assignment $(g, v) \mapsto \mathrm{Ad}(g)(v)$ induces a linear action $G \times LG \to LG$. The Lie algebra LT of T can be identified with a Lie subalgebra of LG. As T is commutative we have $c(t)_{|T} = \mathrm{id}_T$ and hence $\mathrm{Ad}(t)_{|LT} = \mathrm{id}_{LT}$ for all $t \in T$. There exists a scalar product $\langle \cdot, \cdot \rangle$ on LG such that $\mathrm{Ad}(t)(LT^\perp) = LT^\perp$. (If $(\cdot, \cdot)$ is any scalar product on LG then $\langle v, w \rangle := \int_G (\mathrm{Ad}(g)(v), \mathrm{Ad}(g)(w))\, \mu_G(\mathrm{d}g)$ defines a scalar product with the desired property (Weyl's trick). To see this note that $c(g_2 g_1) = c(g_2) \circ c(g_1)$ implies $\mathrm{Ad}(g_2 g_1) = \mathrm{Ad}(g_2)\mathrm{Ad}(g_1)$ and recall that $\mathrm{Ad}(t)_{|LT} = \mathrm{id}_{LT}$.) Then Ad induces a linear T-action on $LT^\perp$ via

$$\mathrm{Ad}_{G/T}\colon T \to \mathrm{Aut}(LT^\perp), \quad \mathrm{Ad}_{G/T}(t)(w) := \mathrm{Ad}(t)(w) \quad \text{for all } w \in LT^\perp. \tag{4.37}$$

This definition of $\mathrm{Ad}_{G/T}$ is suitable for theoretical considerations. Formula (4.39) below provides a representation of $\mathrm{Ad}_{G/T}(t)$ which is more suitable for concrete computations (cf. Section 4.4). The terms ad, $\exp_G$ and exp stand for *adjoint representation* on LG (i.e. $\mathrm{ad}(x)(y) := [x, y]$), and the exponential mappings on LG and $L(\mathrm{Aut}(LG))$. The latter is given by the power series $\exp(x) = \sum_{j=0}^{\infty} x^j/j!$. As usual, $(\mathrm{ad}(x))^j$ denotes the j-fold composition $(\mathrm{ad} \circ \cdots \circ \mathrm{ad})$. From [13], p. 23 (2.3) we obtain $\mathrm{Ad} \circ \exp_G = \exp \circ \mathrm{ad}$ (with $f = \mathrm{Ad}$ and $Lf = \mathrm{ad}$; cf. [13], p. 18 (2.10)). As $\exp_G\colon LG \to G$ is surjective ([13], p. 165 (2.2)) for each $g \in G$ there is a $v_g \in LG$ with $g = \exp_G(v_g)$. Hence

$$\mathrm{Ad}(g) = \mathrm{Ad}\,(\exp_G(v_g)) = \exp\,(\mathrm{ad}(v_g)) = \sum_{j=0}^{\infty} \frac{1}{j!}\,(\mathrm{ad}(v_g))^j \tag{4.38}$$

which finally yields

$$\mathrm{Ad}_{G/T}(t^{-1}) = \exp(\mathrm{ad}(-v_t))_{|LT^\perp} \qquad \text{for } t \in T. \tag{4.39}$$

A more comprehensive treatment of the preceding can be found in special literature on Lie groups (e.g. [13])

To derive a matrix representation for $\mathrm{Ad}_{G/T}(t^{-1}) - \mathrm{id}_{|LT^\perp}$ one chooses a basis of the vector space $LT^\perp$. The computation of the density h_G then requires no more than linear algebra. At the example of $G = \mathrm{SO}(n)$ this approach will be demonstrated in the following section. To obtain a closed expression for $h_{SO(n)}$ for all $n \in \mathbb{N}$, however, this basis has to be chosen in a suitable manner (see Section 4.4 and [66], pp. 253–257).

Definition 4.38. *Let* $N(T) := \{n \in G \mid nT = Tn\} \leq G$ *be the* normalizer *of* T. *The factor group* $W(G) := N(T)/T$ *is called the* Weyl group. *The unique* G*-invariant probability measure on* G/T *is denoted with* $\mu_{(G/T)}$.

Recall that the Haar measure μ_G on G is left- and right invariant (Remark 4.15) and is in particular invariant under conjugation. We point out that the Weyl group $W(G)$ is finite ([13], p. 158 (Theorem 1.5)).

Theorem 4.39. *For $\Phi\colon \mathcal{M}^1(T) \to \mathcal{M}^1_G(G)$, $\Phi(\tau) := (\mu_{(G/T)} \otimes \tau)^q$ the following statements are valid:*
(i) $\Phi(\mathcal{M}^1(T)) = \mathcal{M}^1_G(G)$.
(ii) *(Weyl's integral formula) For*

$$h_G\colon T \to [0,\infty), \qquad h_G(t) := \frac{1}{|W(G)|} \det(\mathrm{Ad}_{G/T}(t^{-1}) - \mathrm{id}_{LT^\perp}) \tag{4.40}$$

we have

$$\mu_G = \left(\mu_{(G/T)} \otimes h_G \cdot \mu_T\right)^q. \tag{4.41}$$

In particular,

$$h_G(t) := \frac{1}{|W(G)|} \det\left(\exp(\mathrm{ad}(-v_t))_{|LT^\perp} - \mathrm{id}_{|LT^\perp}\right) \tag{4.42}$$

with $v_t \in LT$ and $\exp_G(v_t) = t$.
(iii) *Let $\nu \in \mathcal{M}^1_G(G)$ and $\eta = f \cdot \nu \in \mathcal{M}^1(G)$ with conjugation-invariant ν-density f. Then $\eta \in \mathcal{M}^1_G(G)$, and $\Phi(\tau_\nu) = \nu$ implies $\Phi\left(f_{|T} \cdot \tau_\nu\right) = \eta$.*
(iv) *$(\mathcal{M}^1_G(G), \odot)$ is a subsemigroup of the convolution semigroup $(\mathcal{M}^1(G), \odot)$. In particular, $\varepsilon_{e_G} \in \mathcal{M}^1_G(G)$.*

Proof. Statement (i) follows from the continuity of q (Theorem 4.9(i)), and (ii) is shown in [13], pp. 157–163, resp. follows from (4.39). Clearly, $(t_1 \sim t_2) \iff (q(G/T \times \{t_1\}) = q(G/T \times \{t_2\})) \iff (t_1 = q(e_G T, t_2) \in E_G(q(e_G T, t_2)) = E_G(t_2))$, i.e. the equivalence relation $\sim$ is given by the intersection of the G-orbits $E_G(\cdot)$ with T. Theorem 4.8(vii) implies (iii). The mapping $\Theta_g\colon G \to G$, $\Theta_g(g_1) := g g_1 g^{-1}$ is a homomorphism for each $g \in G$. In fact: $\Theta_g(e_G) := e_G$ and $\Theta_g(g_1 g_2) := g(g_1 g_2)g^{-1} = g g_1 g^{-1} g g_2 g^{-1} = \Theta_g(g_1)\Theta_g(g_2)$. Trivially, $e_G^{\Theta_g} = e_G$ and hence (iv) is an immediate consequence of Theorem 4.7(iv) $\square$

Remark 4.40. The proof of Theorem 4.39(ii) given in [13], pp. 157 ff., exploits the fact that μ_G can be interpreted as a differential form and uses techniques from the field of differential geometry. This proof cannot be extended to arbitrary conjugation-invariant measures on G.

Theorem 4.39 says that Φ is surjective. That is, for each $\nu \in \mathcal{M}^1_G(G)$ there exists a $\tau_\nu \in \mathcal{M}^1(T)$ with $(\mu_{(G/T)} \otimes \tau_\nu)^q = \nu$. For $\nu = f \cdot \mu_G$ with conjugation-invariant density f Theorem 4.39 provides an explicit expression for τ_ν. Clearly, the torus subgroup T represents an optimal domain of integration and it favours stochastic simulations (cf. Remark 4.43). In view of statistical applications we will construct a measurable section $R_G \subseteq T$

for which the corresponding mapping $\psi: G \to T$ (cf. Theorem 4.11) is measurable. The uniqueness property of $\tau \in \Phi^{-1}(\nu)$ ensured by the condition $\tau(R_G^c) = 0$ may also be useful for concrete computations. As already shown in the proof of 4.39 the equivalence relation $\sim$ on T is given by the intersection of the G-orbits on G with T. Lemma 4.41 says that the equivalence relation $\sim$ can also be expressed by the action of the Weyl group $W(G)$ on T.

Lemma 4.41. *(i) The Weyl group $W(G)$ acts on T via $(nT, t) \mapsto ntn^{-1}$. In particular, $W(G)t = E_T(\{t\})$ for all $t \in T$.*
(ii) $\mu_T(\{t \in T \mid |E_T(\{t\})| = |W(G)|\}) = 1$.
(iii) $q(G/T \times C) \in \mathcal{B}(G)_F$ for all $C \in \mathcal{B}_E(T) = \{C \in \mathcal{B}(T) \mid C = E_T(C)\}$.

Proof. The assertions (i) and (ii) are shown in [13], p. 166 (Lemma (2.5)), p. 162 (Lemma 1.9(i)) together with p. 38 (equation (4.13)). Assertion (iii) is proved in in [67] (Lemma 4.8). Its core is a deep theorem of Bierlein ([7]) which was generalized (in a straight-forward manner) from the unit interval to arbitrary compact spaces with countable bases (cf. [67], Lemma 4.7). □

As announced above we are now going to construct a measurable section $R_G \subseteq T$. Therefore, we identify the Weyl group $W(G)$ with a finite group W of continuous automorphisms on T which acts on T by $(\alpha, t) \mapsto \alpha(t)$. Now let $W_1 := W, W_2, \ldots, W_s$ denote a maximal collection of mutually non-conjugate subgroups of W while W_t denotes the isotropy group of $t \in T$. Further, for $j \leq s$ let temporarily $S_j := \{t \in T \mid W_t = W_j\}$. We first note that

$$S_j := \bigcap_{\alpha \in W_j} \{t \in T \mid \alpha(t) = t\} \cap \bigcap_{\beta \in W \setminus W_j} \{t \in T \mid \beta(t) \neq t\} \tag{4.43}$$

is contained in $\mathcal{B}(T)$. By 2.4(ii) S_j is locally compact in the induced topology, and by 2.4(i) S_j is second countable. Note that for any $t' \in T$ there exists an automorphism $\beta \in W$ and an index $j \leq s$ such that $W_{t'} = \beta^{-1} W_j \beta$. Clearly, $\beta(t')$ has isotropy group W_j and hence is contained in S_j. Consequently, the disjoint union $\sum_{j=1}^{s} S_j \subseteq T$ intersects each W-orbit. However, in general it is no section. In the next step we determine a subset of $\sum_{j=1}^{s} S_j$ which is a section. Since any two elements which are contained in the same orbit have conjugate isotropy groups the summands can be treated independently. Clearly, the orbit of each $t \in S_1$ is singleton and hence S_1 is the first subset of our section. Let $W_j \neq W$. As $\alpha(t) = t$ for all $(t, \alpha) \in S_j \times W_j$ and since the automorphisms are continuous for each $y \in S_j$ there exists an open neighbourhood $U_{j;y} \subseteq S_j$ (in the induced topology on S_j) such that $U_{j;y} \cap \beta(U_{j;y}) = \emptyset$ for all $\beta \in W \setminus W_j$. As S_j is second countable there exist countably many $y_1, y_2, \ldots \in S_j$ for which the sets $U_{j;y_1}, U_{j;y_2}, \ldots$ cover S_j. Now define

$$V_{j;1} := U_{j;y_1}, \qquad V_{j;i+1} := U_{j;y_{i+1}} \setminus \bigcup_{k=1}^{i} \left(\bigcup_{\alpha \in W} \alpha(U_{j;y_k}) \right) \quad \text{for } i \geq 1. \tag{4.44}$$

By construction, $E_T(\bigcup_i V_{j;i}) \supseteq S_j$, and $\bigcup_i V_{j;i}$ intersects each W-orbit in T which has non-empty intersection with S_j in exactly one point. Further, $V_{j;i} \in \mathcal{B}(S_j) = S_j \cap \mathcal{B}(T) \subseteq B(T)$. By 2.4(ii) each $V_{j;i}$ is locally compact and second countable. Altogether, we have constructed a measurable section R_G which can be expressed as a countable disjoint union of subsets $E_1, E_2, \ldots \in \mathcal{B}(T)$ (relabel the sets $V_{j;i}$) which are locally compact and second countable in their induced topology.

Theorem 4.42. (i) *Let $R_G = \sum_j E_j \subseteq T$ denote the section constructed above. As in Theorem 4.11 let the corresponding mapping given by*

$$\psi\colon G \to T, \qquad \psi(g) := t_g \tag{4.45}$$

where t_g is the unique element in $E_G(\{g\}) \cap R_G$. Then ψ is measurable.
(ii) *For each measurable section $R'_G \subseteq T$ we have*

$$q(G/T \times (C \cap R'_G)) \in \mathcal{B}(G)_F \qquad \text{for all } C \in \mathcal{B}(T). \tag{4.46}$$

That is, the corresponding mapping $\psi'\colon G \to T$ is $(\mathcal{B}(G)_F$-$\mathcal{B}(T))$-measurable.
(iii) *Suppose that $\tau_\nu \in \Phi^{-1}(\nu)$, and let*

$$\tau^*_\nu(B) := \frac{1}{|W(G)|} \sum_{nT \in N(T)} \tau_\nu(nBn^{-1}) \qquad \text{for each } B \in \mathcal{B}(T). \tag{4.47}$$

Then $\tau^*_\nu \in \Phi^{-1}(\nu)$. If $\tau_\nu = f \cdot \mu_T$ then

$$\tau^*_\nu := f^*_W \cdot \mu_T \qquad \text{with} \qquad f^*_W(t) := \frac{1}{|W(G)|} \sum_{nT \in N(T)} f(ntn^{-1}). \tag{4.48}$$

(iv) *The density h_G is constant on each $W(G)$-orbit.*
(v) *For $\tau_\nu := f \cdot \mu_T \in \Phi^{-1}(\nu)$ let*

$$f_{R_G}\colon T \to [0,\infty], \qquad f_{R_G}(t) := 1_{R_G}(t) \cdot \sum_{t' \in E_T(\{t\})} f(t'). \tag{4.49}$$

Then $f_{R_G} \cdot \mu_T \in \Phi^{-1}(\nu)$, and $f_{R_G} \cdot \mu_T$ has its total mass on R_G. In particular, $(h_G)_{R_G}(t) = 1_{R_G}(t)|W(G)| \cdot h_G(t)$.

Proof. As the restriction $q_{|G/T \times E_j}\colon G/T \times E_j \to G$ is continuous the pre-image $\psi^{-1}(K) = q(G/T \times K)$ is compact for each compact subset $K \subseteq E_j$. As E_j is locally compact and second countable (and hence σ-compact) the compact subsets generate $\mathcal{B}(E_j)$ which implies $\psi^{-1}(B') \in B(G)$ for all $B' \in \mathcal{B}(E_j)$. For any $B \in \mathcal{B}(T)$ we have $\psi^{-1}(B) = \sum_j \psi^{-1}(B \cap E_j)$ which completes the proof of (i). As all $\alpha \in W$ are homeomorphisms $E_T(C \cap R'_G) = \bigcup_{\alpha \in W} \alpha\,(C \cap R'_G) \in \mathcal{B}_E(T)$ for each $C \in \mathcal{B}(T)$. From 4.41(iii) we obtain $\psi^{-1}(C) = q(G/T \times (C \cap R'_G)) = q(G/T \times E_T(C \cap R'_G)) \in$

$\mathcal{B}(G)_F$ which completes the proof of (ii). As $W(G)\{t\} = E_T(\{t\})$ in particular $\mathcal{B}_E(T) = \mathcal{B}_{W(G)}(T)$, and 4.7(ii) implies $\tau_{\nu|\mathcal{B}_E(T)} = \tau^*_{\nu|\mathcal{B}_E(T)}$. As $\mathcal{B}_0(T) \subseteq B_E(T)$ the first assertion of (iii) is an immediate consequence of 4.7(ii) since the Haar probability measure on the finite group $W(G)$ equals the equidistribution. The second assertion of (iii) follows from 4.7(iii); it merely remains to prove that μ_T is $W(G)$-invariant. As $|N(t)/T|$ is finite the normalizer $N(T)$ is a compact group with Haar probability measure $\mu_{N(t)}$. In particular $\mu_{N(t)}(T) = 1/|W(G)|$ and $\mu_{N(T)}(B) = \mu_{N(T)}(tB)$ for all $(t, B) \in T \times \mathcal{B}(T)$. Hence the restriction $\mu_{N(t)|T}$ is a Haar measure on T. More precisely, $\mu_{N(t)|T} = \mu_T/|W(G)|$, and μ_T is invariant under the action of the Weyl group. Note that $\mathrm{Ad}(ntn^{-1}) = \mathrm{Ad}(n)\mathrm{Ad}(t)\mathrm{Ad}(n)^{-1}$ which proves (iv). Finally, (v) follows from 4.41(ii), 4.11 and assertion (iv) of this theorem. □

We point out that 4.42(i) is stronger than Lemma 3.5 in [66] as there the mapping ψ was only shown to be $\mathcal{B}(G)_F$-$\mathcal{B}(T)$-measurable.

Consider the test problem $(\mathrm{P}_{\Gamma_0}, \mathrm{P}_{\Gamma\setminus\Gamma_0})$ with $\mathrm{P}_\Gamma \subseteq \mathcal{M}^1_G(G)^m$. From 4.14 $\psi^m\colon G^m \to T^m$ is a sufficient statistic, and we have $(\nu_1 \otimes \cdots \otimes \nu_m)^{\psi^m} = (\tau_1 \otimes \cdots \otimes \tau_m)$ with $\tau_j \in \Phi^{-1}(\nu_j)$ and $\tau_j(R_G) = 1$. Suppose that $p_\gamma := \nu_{\gamma;1} \otimes \cdots \otimes \nu_{\gamma;m}$ for each $\gamma \in \Gamma$ and, additionally, $\nu_{\gamma;j} \ll \mu_G$. Then for each pair $(\gamma, j) \in \Gamma \times \{1, \ldots, m\}$ the probability measure $\nu_{\gamma;j}$ has a conjugation-invariant μ_T-density $f_{\gamma;j}$ (4.7(iii)). In particular $\tau_{\gamma;j} = (f_{\gamma;j})_{|T}(h_G)_{R_G}\cdot\mu_T$. By this, we have determined the transformed test problem $(\overline{\mathrm{P}}_{\Gamma_0}, \overline{\mathrm{P}}_{\Gamma\setminus\Gamma_0})$ without any computation. In Section 4.5 we will consider two examples.

Remark 4.43. (i) There is a group isomorphism between the torus subgroup $T \leq G$ and $(\mathrm{S}^1)^k \cong ([0, r), \oplus)^k$ where '$\oplus$' stands for the addition modulo r. Up to a scalar factor μ_T is the image measure of the k-dimensional Lebesgue measure on $[0, r)^k$ under this isomorphism. Consequently, the essential part of the simulation of a random variable with values in T can be transferred to $[0, r)^k$. More precisely, to simulate the distribution $f \cdot \mu_T$ on T one generates pseudorandom vectors on $[0, r)^k$ with Lebesgue density f/r^k. For many distribution classes in $\mathbb{R}^k$ efficient simulation algorithms are known (see, e.g. [18]). The domain of simulation is a cube. Due to the properties of the standard random numbers (cf. Chapter 3) this favours simulation aspects additionally.
(ii) The mapping $\mathrm{pr}_G\colon G \to G/T$, $\mathrm{pr}_G(g) := gT$ transforms the Haar measure μ_G into $\mu_{(G/T)}$. Let $q(\mathrm{pr}_G(g), t) = gtg^{-1} =: q_0(g, t)$ for all $(g, t) \in G \times T$. In particular $q_0 = q \circ (\mathrm{pr}_G \times \mathrm{id}_T)$, and hence $(\mu_{(G/T)} \otimes \tau)^q = (\mu_G \otimes \tau)^{q_0}$ for all $\tau \in \mathcal{M}^1(T)$. In particular, the 5-tupel (G, G, T, G, q_0) also has Property $(*)$.

4.4 Conjugation-invariant Probability Measures on $\mathrm{SO}(n)$

For all $n \in \mathbb{N}$ the special orthogonal group $\mathrm{SO}(n)$ is a compact connected Lie group. As in Section 4.3 we consider the *conjugation-invariant* probabil-

ity measures on SO(n), i.e. those $\nu \in \mathcal{M}^1(\mathrm{SO}(n))$ with $\nu(B) = \nu(\boldsymbol{S}B\boldsymbol{S}^t)$ for all $(\boldsymbol{S}, B) \in \mathrm{SO}(n) \times \mathcal{B}(\mathrm{SO}(n))$. In particular, all statements from the preceding section are valid and can almost literally be transferred. In this section we determine an explicit formula for the density $\mathrm{h}_{SO(n)}$, i.e. we compute the determinant $(\mathrm{Ad}_{SO(n)/T}(t^{-1}) - \mathrm{id}_{LT^\perp})$. Further, we determine a section $R_{\mathrm{SO}(n)} \subseteq T$ which can simply be described. This supports the efficient computation of a sufficient statistic $\psi_n^m(\cdot)$. Interestingly, we have to distinguish between even and odd n. We recall that $\boldsymbol{S}^{-1} = \boldsymbol{S}^t$ for all $\boldsymbol{S} \in \mathrm{SO}(n)$ which facilitates the computation of matrix products $\boldsymbol{S}\boldsymbol{T}\boldsymbol{S}^{-1}$, e.g. within stochastic simulations.

Definition 4.44. *For $\theta \in \mathbb{R}$ we define*

$$\mathbf{T}_{(\theta)} := \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{4.50}$$

For $n = 2k$ the term $\mathbf{T}_{(\theta_1,\theta_2,\ldots,\theta_k)}$ denotes the block diagonal matrix consisting of the blocks $\boldsymbol{T}_{(\theta_1)}, \boldsymbol{T}_{(\theta_2)}, \ldots, \boldsymbol{T}_{(\theta_k)}$. If $n = 2k+1$ we additionally attach the (1×1)-block '1' as component $(2k+1, 2k+1)$.

Note that for $n \in \{2k, 2k+1\}$ the subgroup

$$T := \{\mathbf{T}_{(\theta_1,\theta_2,\ldots,\theta_k)} \mid \theta_1, \theta_2, \ldots, \theta_k \in \mathbb{R}\} \leq SO(n) \tag{4.51}$$

is a maximal torus of SO(n) ([13], p. 171; cf. also Definition 4.32). Clearly, $\mathbf{T}_{(\theta_1,\theta_2,\ldots,\theta_k)} \mapsto (\theta_1 (\mathrm{mod}\, 2\pi), \ldots, \theta_k (\mathrm{mod}\, 2\pi))$ induces a group isomorphism between T and $([0, 2\pi), \oplus)^k \cong (\mathrm{S}^1)^k$. In particular we define

$$q_n \colon \mathrm{SO}(n)/T \times T \to \mathrm{SO}(n) \qquad q_n(\boldsymbol{S}T, \boldsymbol{T}) := \boldsymbol{S}\boldsymbol{T}\boldsymbol{S}^t, \tag{4.52}$$

and we obtain

Lemma 4.45. *The 5-tuple $(\mathrm{SO}(n), \mathrm{SO}(n)/T, T, \mathrm{SO}(n), q_n)$ has Property (∗).*

Proof. special case of Lemma 4.36 □

Next, we are going to prove a pendant to Theorem 4.39. In particular, we will derive an explicit formula for the density $\mathrm{h}_{SO(n)}$. This needs some preparatory work.

Definition 4.46. *As usual, μ_T denotes the Haar probability measure on T and $\mathrm{span}\{v_1, v_2, \ldots, v_p\}$ the linear span of the vectors $v_1, v_2, \ldots, v_p$. For $l \neq p$ we define the $(n \times n)$-matrix $\boldsymbol{E}^{(lp)}$ by $e_{lp}^{(lp)} := 1$, $e_{pl}^{(lp)} := -1$ and $e_{ij}^{(lp)} := 0$ if $\{i, j\} \neq \{l, p\}$. Further, $G(k) := \{\pi \mid \pi$ is a permutation on $\{-k, -k+1, \ldots, -1, 1, \ldots, k\}$ with $\pi(-s) = -\pi(s)$ for $1 \leq s \leq k\}$ and $SG(k) := \{\pi \in G(k) \mid \pi$ is even$\}$.*

Remark 4.47. We mention that $L(\mathrm{SO}(n)) = \mathrm{so}(n)$, the Lie algebra of all skew-symmetric $(n \times n)$-matrices. The Lie bracket on $\mathrm{so}(n)$ is given by the commutator $[\boldsymbol{K}, \boldsymbol{L}] := \boldsymbol{KL} - \boldsymbol{LK}$ (cf. Remark 4.17), and $\mathcal{E} := \{\boldsymbol{E}^{(lp)} \mid 1 \leq l < p \leq n\}$ is a vector basis $\mathrm{so}(n)$. The exponential map is given by the power series

$$\exp_{SO(n)}\colon \mathrm{so}(n) \to \mathrm{SO}(n), \qquad \exp_{SO(n)}(\boldsymbol{L}) := \sum_{j=0}^{\infty} \boldsymbol{L}^j/j!, \tag{4.53}$$

and the adjoint representation by

$$\mathrm{ad}\colon \mathrm{so}(n) \to \mathrm{so}(n) \qquad \mathrm{ad}(\boldsymbol{K})\boldsymbol{M} = \boldsymbol{KM} - \boldsymbol{MK}. \tag{4.54}$$

As in Section 4.3 $N(T) := \{\boldsymbol{S} \in \mathrm{SO}(n) \mid \boldsymbol{S}T = T\boldsymbol{S}\}$. The Weyl group $W(\mathrm{SO}(n)) := N(T)/T$ is finite, and $(\boldsymbol{N}T, \boldsymbol{T}) \mapsto \boldsymbol{NTN}^t$ defines a $W(\mathrm{SO}(n))$-action on T. By 4.41(i) $E_T(\{\boldsymbol{T}\})$ equals the $W(\mathrm{SO}(n))$-orbit of $\boldsymbol{T}$.

Lemma 4.48. (i) $\exp_{SO(n)}\left(\sum_{r=1}^{k} \theta_r \boldsymbol{E}^{(2r-1,2r)}\right) = \mathbf{T}_{(\theta_\mathbf{1},\theta_\mathbf{2},\ldots,\theta_\mathbf{k})}$.
(ii) *Let* $\boldsymbol{K}, \boldsymbol{L} \in \mathrm{so}(n)$ *with* $[\boldsymbol{K}, \boldsymbol{L}] = 0$. *Then* $\mathrm{ad}(\boldsymbol{K})$ *and* $\mathrm{ad}(\boldsymbol{L})$ *commute, too, and*

$$\exp_{SO(n)}(\mathrm{ad}(\boldsymbol{K} + \boldsymbol{L})) = \exp_{SO(n)}(\mathrm{ad}(\boldsymbol{K})) \circ \exp_{SO(n)}(\mathrm{ad}(\boldsymbol{L})). \tag{4.55}$$

(iii) For $n \in \{2k, 2k+1\}$

$$W(\mathrm{SO}(2k+1)) \cong G(k), \qquad W(\mathrm{SO}(2k)) \cong SG(k). \tag{4.56}$$

This induces a $G(k)$*-action, resp. a* $SG(k)$*-action, on* T *which is given by*

$$(\alpha, \mathbf{T}_{(\theta_\mathbf{1},\theta_\mathbf{2},\ldots,\theta_\mathbf{k})}) \mapsto \mathbf{T}_{(\theta_{\alpha^{-1}(\mathbf{1})},\theta_{\alpha^{-1}(\mathbf{2})},\ldots,\theta_{\alpha^{-1}(\mathbf{k})})} \qquad \text{with } \theta_{-j} = -\theta_j \tag{4.57}$$

for $\left(\alpha, \mathbf{T}_{(\theta_\mathbf{1},\theta_\mathbf{2},\ldots,\theta_\mathbf{k})}\right) \in G(k) \times T$ *or* $\left(\alpha, \mathbf{T}_{(\theta_\mathbf{1},\theta_\mathbf{2},\ldots,\theta_\mathbf{k})}\right) \in SG(k) \times T$, *resp. In particular* $|W(\mathrm{SO}(2k))| = 2^{k-1}k!$ *and* $|W(\mathrm{SO}(2k+1))| = 2^k k!$.

Proof. As any pair of summands in $\sum_{r=1}^{k} \theta_r \boldsymbol{E}^{(2r-1,2r)}$ commutes we conclude $\exp_{SO(n)}(\sum_{r=1}^{k} \theta_r \boldsymbol{E}^{(2r-1,2r)}) = \prod_{r=1}^{k} \exp_{SO(n)}(\theta_r \boldsymbol{E}^{(2r-1,2r)})$ which reduces the computation of the exponential image essentially to that of 2×2-matrices. Since the powers of $\boldsymbol{E}^{(2r-1,2r)}$ have period 4 one verifies easily that $\exp_{SO(n)}(\theta_r \boldsymbol{E}^{(2r-1,2r)}) = \mathbf{T}_{0,\ldots,0,\theta_r,0,\ldots,0}$. For commuting matrices $\boldsymbol{K}, \boldsymbol{L} \in \mathrm{so}(n)$ the Jacobi identity implies $\mathrm{ad}(\boldsymbol{K})\mathrm{ad}(\boldsymbol{L})\boldsymbol{M} - \mathrm{ad}(\boldsymbol{L})\mathrm{ad}(\boldsymbol{K})\boldsymbol{M} = -[\boldsymbol{M}, [\boldsymbol{K}, \boldsymbol{L}]] = 0$, that is, $\mathrm{ad}(\boldsymbol{K})$ and $\mathrm{ad}(\boldsymbol{L})$ commute. Together with $\mathrm{ad}(\boldsymbol{K} + \boldsymbol{L}) = \mathrm{ad}(\boldsymbol{K}) + \mathrm{ad}(\boldsymbol{L})$ this proves the functional equation (4.55). The first two assertions of (iii) are shown in ([13] pp. 171f.). Straight-forward combinatorial considerations yield $|G(k)| = 2^k k!$ and $|SG(k)| = |G(k)|/2$ which completes the proof of (iii). □

Theorem 4.49. *Let* $\Phi\colon \mathcal{M}^1(T) \to \mathcal{M}^1_{\mathrm{SO}(n)}(\mathrm{SO}(n))$,
$\Phi(\tau) := (\mu_{(SO(n)/T)} \otimes \tau)^{q_n}$. *Then the following statements are valid:*
(i) $\Phi(\mathcal{M}^1(T)) = \mathcal{M}^1_{SO(n)}(\mathrm{SO}(n))$.
(ii) *(Weyl's integral formula) For* $\mathrm{h}_{SO(n)}\colon T \to [0,\infty)$,

$$\mathrm{h}_{SO(n)}(t) := \det(\mathrm{Ad}_{SO(n)/T}(t^{-1}) - \mathrm{id}_{LT^\perp})/|W(\mathrm{SO}(n))| \tag{4.58}$$

we have

$$\mu_{SO(n)} = \left(\mu_{(SO(n)/T)} \otimes \mathrm{h}_{SO(n)} \cdot \mu_T\right)^{q_n}. \tag{4.59}$$

More precisely: For $n \geq 2$ *it is* $h_{\mathrm{SO}(n)}(\mathbf{T}_{(\theta_1,\theta_2,\ldots,\theta_k)}) =$

$$\begin{cases} \frac{2^{k^2-2k+1}}{k!} \prod_{1\leq s<t\leq k} (\cos\theta_s - \cos\theta_t)^2 & \text{if } n = 2k \\ \frac{2^{k^2-k}}{k!} \prod_{1\leq s<t\leq k} (\cos\theta_s - \cos\theta_t)^2 \prod_{1\leq r\leq k} (1-\cos\theta_r) & \text{if } n = 2k+1. \end{cases} \tag{4.60}$$

(iii) *Let* $\nu \in \mathcal{M}^1_{SO(n)}(\mathrm{SO}(n))$ *and* $\eta = f \cdot \nu \in \mathcal{M}^1(\mathrm{SO}(n))$ *with conjugation-invariant* ν*-density* f. *Then* $\eta \in \mathcal{M}^1_{SO(n)}(\mathrm{SO}(n))$, *and* $\Phi(\tau_\nu) = \nu$ *implies* $\Phi(f_{|T} \cdot \tau_\nu) = \eta$.
(iv) $(\mathcal{M}^1_{SO(n)}(\mathrm{SO}(n)), \odot)$ *is a subsemigroup of the convolution semigroup* $(\mathcal{M}^1(\mathrm{SO}(n)), \odot)$. *In particular* $\varepsilon_{1_n} \in \mathcal{M}^1_{SO(n)}(\mathrm{SO}(n))$.

Proof. Assertions (i), (iii), (iv), (4.58) and (4.59) are special cases of Theorem 4.39. Only (4.60) remains to be verified. We point out that its proof applies techniques from linear algebra. As it is rather technical (cf. [66], Sect. 4) we merely sketch its central ideas. From 4.48(i) we conclude $LT = span\,\mathcal{E}_1 := span\{\boldsymbol{E}^{1,2}, \ldots, \boldsymbol{E}^{2k-1,2k}\}$. Let further $\mathcal{E}_2 := \mathcal{E} \setminus \mathcal{E}_1$. Careful but elementary computations verify that $span\,\mathcal{E}_2 = LT^\perp$ with respect to the scalar product $(\boldsymbol{K}, \boldsymbol{L}) \mapsto \mathrm{tr}(\boldsymbol{K}^t \boldsymbol{L})$. If the basis $\mathcal{E}_2$ is ordered suitably then the matrix representation of $(\mathrm{Ad}_{SO(n)/T}(\mathbf{T}_{(\theta_1,\theta_2,\ldots,\theta_k)}{}^{-1}) - \mathrm{id}_{LT^\perp})$ is a block diagonal matrix ([66], Lemma 4.5 (iii)). More concrete: Its diagonal blocks are (4×4)-matrices which depend on two parameters but are essentially equal. If n is odd additionally k (2×2) diagonal blocks appear which depend on one parameter and are essentially equal. The structure of the matrix blocks does not depend on n, and straight-forward computations yield (4.60). □

Next, we are going to determine a section $\mathrm{R}_{SO(n)} \subseteq T$ and the corresponding mapping $\psi_n\colon \mathrm{SO}(n) \to \mathrm{R}_{SO(n)}$. We recall that for each $\boldsymbol{S} \in \mathrm{SO}(2k)$ there exist real numbers $0 \leq \rho_1(\boldsymbol{S}) \leq \rho_2(\boldsymbol{S}) \leq \cdots \leq \rho_k(\boldsymbol{S}) \leq \pi$ such that the eigenvalues of $\boldsymbol{S}$ (counted with their multiplicity) are given by $\mathrm{e}^{i\rho_1(\boldsymbol{S})}, \mathrm{e}^{-i\rho_1(\boldsymbol{S})}, \ldots, \mathrm{e}^{i\rho_k(\boldsymbol{S})}, \mathrm{e}^{-i\rho_k(\boldsymbol{S})}$. For $n = 2k+1$ the matrix $\boldsymbol{S}$ additionally has the eigenvalue 1. Clearly, for $\mathbf{T}_{(\boldsymbol{\theta_1},\boldsymbol{\theta_2},\ldots,\boldsymbol{\theta_k})} \in T$ the numbers $\rho_1(\mathbf{T}_{(\boldsymbol{\theta_1},\boldsymbol{\theta_2},\ldots,\boldsymbol{\theta_k})}), \ldots, \rho_k(\mathbf{T}_{(\boldsymbol{\theta_1},\boldsymbol{\theta_2},\ldots,\boldsymbol{\theta_k})})$ are a permutation of the values $\min\{\theta_1, 2\pi - \theta_1\}, \ldots, \min\{\theta_k, 2\pi - \theta_k\}$. The proof of Theorem 4.50 exploits the action of the Weyl group $W(\mathrm{SO}(n))$ (cf. Lemma 4.48).

Theorem 4.50. *Let $k \geq 1$. Then*
(i)

$$\mathrm{R}_{SO(2k+1)} := \{\mathbf{T}_{(\theta_\mathbf{1},\theta_\mathbf{2},\ldots,\theta_\mathbf{k})} \mid 0 \leq \theta_1 \leq \ldots \leq \theta_k \leq \pi\} \subseteq T \subseteq \mathrm{SO}(2k+1) \tag{4.61}$$

is a measurable section.
(ii)

$$\begin{aligned} &\mathrm{R}_{SO(2k)} := R_{2k;1} + R_{2k;2} \subseteq T \subseteq \mathrm{SO}(2k) \qquad \textit{with} \\ &R_{2k;1} := \{\mathbf{T}_{(\theta_\mathbf{1},\theta_\mathbf{2},\ldots,\theta_\mathbf{k})} \mid 0 < \theta_1 \leq \ldots \leq \theta_{k-1} \leq \pi;\ \theta_{k-1} \leq \theta_k \leq 2\pi - \theta_{k-1}\} \\ &R_{2k;2} := \{\mathbf{T}_{(\theta_\mathbf{1},\theta_\mathbf{2},\ldots,\theta_\mathbf{k})} \mid 0 = \theta_1 \leq \ldots \leq \theta_k \leq \pi\} \end{aligned} \tag{4.62}$$

is a measurable section.
(iii) *Let $\boldsymbol{S} \in \mathrm{SO}(2k)$ with $0 < \rho_1(\boldsymbol{S}) \leq \rho_2(\boldsymbol{S}) \leq \cdots \leq \rho_k(\boldsymbol{S}) < \pi$. Then there exists a set of orthonormal eigenvectors $v_1, v_2, \ldots, v_k \in \mathbb{C}^{2k}$ corresponding to the eigenvalues $\mathrm{e}^{i\rho_1(\boldsymbol{S})}, \mathrm{e}^{i\rho_2(\boldsymbol{S})}, \ldots, \mathrm{e}^{i\rho_k(\boldsymbol{S})}$. For $j \leq k$ let $\overline{v}_j$ denote the vector which is componentwise conjugate complex to v_j. If $\det(v_1, \overline{v}_1, v_2, \ldots, v_k, \overline{v}_k) = (-i)^k$ then $\boldsymbol{S} \in E_G(\{\mathbf{T}_{(\rho_\mathbf{1}(\mathbf{S}),\rho_\mathbf{2}(\mathbf{S}),\ldots,\rho_\mathbf{k}(\mathbf{S}))}\})$ while $\boldsymbol{S} \in E_G(\{\mathbf{T}_{(\rho_\mathbf{1}(\mathbf{S}),\rho_\mathbf{2}(\mathbf{S}),\ldots,\rho_\mathbf{k-1}(\mathbf{S}),\mathbf{2\pi}-\rho_\mathbf{k}(\mathbf{S}))}\})$ else (namely, if $\det(v_1, \overline{v}_1, v_2, \ldots, v_k, \overline{v}_k) = -(-i)^k$).*
(iv) *For the sections $R_{\mathrm{SO}(2k+1)}$ and $R_{\mathrm{SO}(2k)}$ the corresponding mapping $\psi_n\colon \mathrm{SO}(n) \to R_{\mathrm{SO}(n)}$ is given by $\psi_n(\boldsymbol{S}) :=$*

$$\begin{cases} \mathbf{T}_{(\rho_\mathbf{1}(\mathbf{S}),\rho_\mathbf{2}(\mathbf{S}),\ldots,\rho_\mathbf{k}(\mathbf{S}))} & \textit{if } n = 2k+1 \\ \mathbf{T}_{(\rho_\mathbf{1}(\mathbf{S}),\rho_\mathbf{2}(\mathbf{S}),\ldots,\rho_\mathbf{k}(\mathbf{S}))} & \textit{if } n = 2k \textit{ and } \big[\rho_1(\boldsymbol{S}) = 0 \textit{ or } \rho_k(\boldsymbol{S}) = \pi \\ & \qquad \textit{or } \det(v_1, \overline{v}_1, \ldots, \overline{v}_k) = (-i)^k\big] \\ \mathbf{T}_{(\rho_\mathbf{1}(\mathbf{S}),\rho_\mathbf{2}(\mathbf{S}),\ldots,\rho_\mathbf{k-1}(\mathbf{S}),\mathbf{2\pi}-\rho_\mathbf{k}(\mathbf{S}))} & \textit{else.} \end{cases} \tag{4.63}$$

In particular, ψ_n is measurable.

Proof. For the integers $1 \leq i < j \leq k$ let $\alpha_{i,j}$ and β_j denote permutations on the set $\{-k, \ldots, -1, 1, \ldots, k\}$ which are given by $\alpha_{i,j}(i) = j$, $\alpha_{i,j}(j) = i$, $\alpha_{i,j}(-i) = -j$, $\alpha_{i,j}(-j) = -i$, and $\alpha_{i,j}(r) = r$ for $r \notin \{-i, -j, i, j\}$ while $\beta_j(j) = -j$, $\beta_j(-j) = j$, and $\beta_j(r) = r$ for $r \notin \{-j, j\}$. Clearly, $\alpha_{i,j}, \beta_j \in G(k)$. Assume $n = 2k+1$, $0 \leq \widetilde{\theta}_j \leq 2\pi$, and let $\theta_1 \leq \cdots \leq \theta_k$ be a permutation of the values $\min\{\widetilde{\theta}_1, 2\pi - \widetilde{\theta}_1\}, \ldots, \min\{\widetilde{\theta}_k, 2\pi - \widetilde{\theta}_k\}$. Using suitable elements $\alpha_{i,j}, \beta_j \in G(k)$ the matrix $\boldsymbol{T}_{\widetilde{\theta}_1,\ldots,\widetilde{\theta}_k} \in T$ can successively be transformed to $\boldsymbol{T}_{\theta_1,\ldots,\theta_k} \in R_{2k+1}$ by the action of $G(k)$. This means that R_{2k+1} intersects each $W(\mathrm{SO}(2k+1)) = G(k)$-orbit on T. As the eigenvalues are invariant under conjugation R_{2k+1} is a section. For $n = 2k$ the situation is more complicated since the permutations β_j are not even. However, there is a $\gamma \in SG(k)$ which maps $\boldsymbol{T}_{\widetilde{\theta}_1,\ldots,\widetilde{\theta}_k}$ onto $\boldsymbol{T}_{\theta_1,\ldots,\theta_k}$ with $0 \leq \theta_1 \leq \cdots \theta_{k-1} \leq \pi$ and $\theta_{k-1} \leq \theta_k \leq 2\pi - \theta_{k-1}$. If $\pi < \theta_k$ the matrix $\boldsymbol{T}_{\theta_1,\ldots,2\pi-\theta_k}$ is not contained in the same $SG(k)$-orbit as $\boldsymbol{T}_{\theta_1,\ldots,\theta_k}$ unless $\widetilde{\theta}_1 = 0$ (Apply $\beta_k \circ \beta_1$ in that case.)

The existence of vectors $v_1, v_2, \ldots, v_k$ with the properties claimed in (iii) is guaranteed since the eigenspaces of distinct eigenvalues of $\boldsymbol{S}$ are orthogonal. Further, $\boldsymbol{S}\overline{v}_j = e^{-i\rho_j(\boldsymbol{S})}\overline{v}_j$ and $v_1, \overline{v}_1, v_2, \ldots, \overline{v}_k$ is an orthonormal basis of $\mathbb{C}^{2k}$ ([74], pp. 334 ff.). For $j \leq k$ let $w_j := (v_j + \overline{v}_j)/\sqrt{2}$ and $w'_j := (v_j - \overline{v}_j)/\sqrt{2}\,i$ while $\boldsymbol{W}$ denotes the matrix whose columns are given by $w_1, w'_1, w_2, \ldots, w'_k$. Then $\boldsymbol{W}$ is orthogonal and $\boldsymbol{W}^t\boldsymbol{S}\boldsymbol{W} = \mathbf{T}_{(\rho_{\mathbf{1}}(\mathbf{S}),\rho_{\mathbf{2}}(\mathbf{S}),\ldots,\rho_{\mathbf{k}}(\mathbf{S}))}$ ([74], pp. 334 ff.). Clearly, $\boldsymbol{S} \in E_G(\{\mathbf{T}_{(\rho_{\mathbf{1}}(\mathbf{S}),\rho_{\mathbf{2}}(\mathbf{S}),\ldots,\rho_{\mathbf{k}}(\mathbf{S}))}\})$ if and only if $1 = \det \boldsymbol{W} = i \cdot \det(v_1, \overline{v}_1, w_2, w'_2, w_3, \ldots, w'_k) = \cdots = i^k \det(v_1, \overline{v}_1, v_2, \ldots, \overline{v}_k)$ which completes the proof of (iii). Assertion (iv) is an immediate consequence from (i), (ii) and (iii). □

Theorem 4.51. (i) $\mu_T(\{\boldsymbol{T} \in T \mid |E_T(\{\boldsymbol{T}\})| = |W(\mathrm{SO}(n))|\}) = 1$.
(ii) *For each measurable section $R'_G \subseteq T$ the corresponding mapping $\psi' : \mathrm{SO}(n) \to T$ is $(\mathcal{B}(\mathrm{SO}(n))_F$-$\mathcal{B}(T))$-measurable.*
(iii) *Suppose that $\tau_\nu \in \Phi^{-1}(\nu)$, and let*

$$\tau^*_\nu(B) := \frac{1}{|W(\mathrm{SO}(n))|} \sum_{nT \in N(T)} \tau_\nu(nBn^{-1}) \qquad \text{for each } B \in \mathcal{B}(T). \quad (4.64)$$

Then $\tau^*_\nu \in \Phi^{-1}(\nu)$. If $\tau_\nu = f \cdot \mu_T$ then

$$\tau^*_\nu := f^*_W \cdot \mu_T \qquad \text{with} \qquad f^*_W(\boldsymbol{T}) := \frac{1}{|W(\mathrm{SO}(n))|} \sum_{\mathbf{T}' \in E_T(\{\mathbf{T}\})} f(\mathbf{T}'). \quad (4.65)$$

(iv) *The density $\mathrm{h}_{SO(n)}$ is constant on each $W(\mathrm{SO}(n))$-orbit.*
(v) *For $\tau_\nu := f \cdot \mu_T \in \Phi^{-1}(\nu)$ let*

$$f_{R_{\mathrm{SO}(n)}} : T \to [0, \infty], \qquad f_{R_{\mathrm{SO}(n)}}(\boldsymbol{T}) := 1_{R_{\mathrm{SO}(n)}}(\boldsymbol{T}) \sum_{\mathbf{T}' \in E_T(\{\mathbf{T}\})} f(\mathbf{T}') \quad (4.66)$$

Then $f_{R_{\mathrm{SO}(n)}} \cdot \mu_T \in \Phi^{-1}(\nu)$, and $f_{R_{\mathrm{SO}(n)}} \cdot \mu_T$ has its total mass on $R_{\mathrm{SO}(n)}$. In particular

$$(h_{\mathrm{SO}(n)})_{R_{\mathrm{SO}(n)}}(\boldsymbol{T}) = 1_{R_{\mathrm{SO}(n)}}(\boldsymbol{T}) h_{R_{\mathrm{SO}(n)}}(\boldsymbol{T}) |W(\mathrm{SO}(n))|. \quad (4.67)$$

Proof. As the SO(n) is a compact connected Lie group statement (i) follows from 4.41(ii), and the statements (ii) to (v) of Theorem 4.42. □

The Theorems 4.49, 4.50 and 4.51 provide useful information for the efficient evaluation of integrals with respect to conjugation-invariant measures, for the simulation of conjugation-invariant distributions and the handling of test problems on SO(n) where the admissible hypotheses P_Γ are products of conjugation-invariant probability measures. In the sequel we give some advice which may be useful for a practical carrying through. For $n = 3$ examples of stochastic simulations and sufficient statistics are discussed in Section 4.5. We abstained from examples for $n > 3$ for two reasons: First of all, $n = 3$

surely constitutes the case of most practical relevance for stochastic simulations. Further, the essential effects and mechanisms should not be covered by complicated analytical or numerical computations and setup operations.

Suppose that $\widetilde{X}$ and $\widetilde{Y}$ are independent $\mu_{(\mathrm{SO}(n)/T)}$-distributed, resp. τ_ν-distributed, pseudorandom elements elements on $\mathrm{SO}(n)/T$ and T, resp. Then $\widetilde{Z} := q_n(\widetilde{X}, \widetilde{Y})$ may be viewed as a $\mu_{\mathrm{SO}(n)}$-distributed pseudorandom element. The use of the mapping q_n decomposes conjugation-invariant simulation problems on $\mathrm{SO}(n)$ into two independent simulation problems on spaces of smaller dimension. Clearly, in regard to the dimensions this decomposition is optimal since $\dim(\mathrm{SO}(n)) = \dim(\mathrm{SO}(n)/T) + \dim(T)$. (The same is true for the general case treated in Section 4.3 were G is a compact connected Lie group.) For the simulation of $\mu_{(\mathrm{SO}(n)/T)}$ we need a suitable representation of the factor space $\mathrm{SO}(n)/T$. For $n = 3$, for example, the homogeneous space $\mathrm{SO}(3)/T$ is isomorphic to S^2 (cf. Section 4.5). If n is large it may cost some effort to obtain a suitable coordinate representation of the factor space $\mathrm{SO}(n)/T$. At least if a large number of pseudorandom matrices on $\mathrm{SO}(n)$ have to be generated this might be worth doing. However, based on Remark 4.43(ii) also another approach is possible which may require more computation time per generated pseudorandom matrix on $\mathrm{SO}(n)$ but saves this preparatory work.

In place of $\widetilde{X} \in \mathrm{SO}(n)/T$ one generates a $\mu_{\mathrm{SO}(n)}$-distributed pseudorandom matrix $\widetilde{X}_0 \in \mathrm{SO}(n)$ and computes $\widetilde{Z} = \widetilde{X}_0 \widetilde{Y} \widetilde{X}_0^t$ instead of $q_n(\widetilde{X}_0 \widetilde{Y})$. At first sight it may appear disconcerting to generate pseudorandom matrices on $\mathrm{SO}(n)$ and T, resp., in order to obtain a single one on $\mathrm{SO}(n)$. However, this approach should also provide enourmous advantages. A direct simulation of a conjugation-invariant distribution $\nu \neq \mu_{\mathrm{SO}(n)}$ generally requires costly acceptance-rejection steps on the $(n(n-1)/2)$-dimensional manifold $\mathrm{SO}(n)$. As will be outlined in Section 4.5 even for $n = 3$ this may be labourious. Using $\widetilde{Z} = \widetilde{X}_0 \widetilde{Y} \widetilde{X}_0^t$ the ν-specific part of the simulation is transferred from $\mathrm{SO}(n)$ to the maximal torus T. Clearly, the dimension of T is much smaller than that of $\mathrm{SO}(n)$ and its domain is very suitable for simulation purposes (cf. Remark 4.43(i)). The $\mathrm{SO}(7)$, for example, is a 21-dimensional manifold which can be embedded in $\mathbb{R}^{49}$ whereas the maximal torus T is isomorphic to the three-dimensional cube $[0, 2\pi)^3$ (equipped with the componentwise addition modulo 2π). Due to its outstanding symmetry there exist efficient algorithms for the simulation of the Haar probability measure $\mu_{\mathrm{O}(n)}$ on the orthogonal group $\mathrm{O}(n)$. Useful algorithms can be found in [77] and [34]. For the latter one should consider the corrections in [79]. Heiberger's algorithm ([34]) needs no more than $n(n+1)/2$ independent $N(0,1)$-distributed pseudorandom numbers to generate one $\mu_{\mathrm{O}(n)}$-distributed pseudorandom orthogonal matrix. Clearly, these algorithms can also be used to simulate $\mu_{\mathrm{SO}(n)}$. If the determinant of the generated pseudorandom matrix is negative then one just multiplies its first column with -1.

For nearly all conjugation-invariant $\nu \in \mathcal{M}^1_{\mathrm{SO}(n)}(\mathrm{SO}(n))$ the simulation algorithm derived from Theorem 4.49(ii) and its variant described above reduce the average computation time and the average number of standard random numbers needed per generated ν-distributed pseudorandom matrix considerably. In particular, the symmetry concept supports simulations on conjugation-invariant subsets of SO(n). In the next section we will quantify these advantages for the special case $G = \mathrm{SO}(3)$.

Theorem 4.51 is relevant for test problems $(\mathrm{P}_{\Gamma_0}, \mathrm{P}_{\Gamma\setminus\Gamma_0})$ with $\mathrm{P}_\Gamma \subseteq \mathcal{M}^1_{\mathrm{SO}(n)}(\mathrm{SO}(n))^m$. As already pointed out in the previous chapters we may consider the transformed test problem $(\overline{\mathrm{P}}_{\Gamma_0}, \overline{\mathrm{P}}_{\Gamma\setminus\Gamma_0})$ instead of the original one without loss of information. As in Section 4.3 $\psi^m\colon G^m \to T^m$ is a sufficient statistic, and we have $(\nu_1 \otimes \cdots \otimes \nu_m)^{\psi^m} = (\tau_1 \otimes \cdots \otimes \tau_m)$ with $\tau_j \in \Phi^{-1}(\nu_j)$ and $\tau_j(R_{\mathrm{SO}(n)}) = 1$. Suppose that $p_\gamma := \nu_{\gamma;1} \otimes \cdots \otimes \nu_{\gamma;m}$ for each pair $(\gamma, j) \in \Gamma \times \{1, \ldots, m\}$. By 4.7(iii) each $\nu_{\gamma;j}$ has a conjugation-invariant μ_T-density $f_{\gamma;j}$. In particular $\tau_{\gamma;j} = (f_{\gamma;j})_{|T}(\mathrm{h}_{\mathrm{SO}(n)})_{\mathrm{R}_{\mathrm{SO}(n)}} \cdot \mu_T$. By this, we have determined the transformed test problem $(\overline{\mathrm{P}}_{\Gamma_0}, \overline{\mathrm{P}}_{\Gamma\setminus\Gamma_0})$ without any computation. Instead of the sample $\boldsymbol{S}_1, \ldots, \boldsymbol{S}_m \in \mathrm{SO}(n)$ only the images $\psi_n(\boldsymbol{S}_1), \ldots, \psi_n(\boldsymbol{S}_m) \in \mathrm{R}_{\mathrm{SO}(n)} \subseteq T$ have to considered in the sequel. The expositions subsequent to Theorem 4.42 can almost literally be transferred. Moreover, Theorem 4.49 provides an explicit expression for the density $\mathrm{h}_{\mathrm{SO}(n)}$.

For a practical carrying through it may be recommendable to move the remaining (=transformed) test problem from the torus subgroup $T \leq \mathrm{SO}(n)$ to a k-dimensional cube ($n \in \{2k, 2k+1\}$) using the canonical group isomorphism

$$is_n\colon T \to [0, 2\pi)^k, \qquad is_n(\mathbf{T}_{(\theta_\mathbf{1}, \theta_\mathbf{2}, \ldots, \theta_\mathbf{k})}) := (\theta_1, \theta_2, \ldots, \theta_k). \tag{4.68}$$

where we equip $[0, 2\pi)^k$ with the componentwise addition modulo 2π. For $\tau \in \mathcal{M}(T)$ and $B \in \mathcal{B}([0, 2\pi)^k)$ clearly $\tau^{is_n}(B) = \tau(\{\mathbf{T}_{(\theta_\mathbf{1}, \theta_\mathbf{2}, \ldots, \theta_\mathbf{k})} \in T \mid (\theta_1, \ldots, \theta_k) \in B\})$. The Haar probability measure μ_T on T is mapped onto $(2\pi)^{-k} \cdot \lambda_{k\,[0,2\pi)^k}$. Moreover, $(f \cdot \tau)^{is_n} = (2\pi)^{-k}(f \circ is_n^{-1}) \cdot \lambda_{[0,2\pi)^k}$. Note that $is_n(\mathrm{R}_{\mathrm{SO}(n)})$ is convex. For odd n it is even a simplex.

Remark 4.52. To compute $is_n \circ \psi_n(\boldsymbol{S}_1), \ldots, is_n \circ \psi_n(\boldsymbol{S}_m) \in is_n(\mathrm{R}_{\mathrm{SO}(n)})$ for a concrete sample $\boldsymbol{S}_1, \ldots, \boldsymbol{S}_m \in \mathrm{SO}(n)$ one at least has to determine their eigenvalues. For $n \geq 4$ numerical methods have to be applied. As all eigenvalues of $\boldsymbol{S}$ have absolute value 1 the well-known LR- and QR-algorithms ([78], pp. 361 ff.) break down. We hence suggest to add the n-dimensional identity matrix 1_n twice in order to separate the absolute values of the distinct eigenvalues unless they are conjugate complex. The eigenvalues of $(\boldsymbol{S} + 2 \cdot 1_n)$ can be computed with an LR- or QR-algorithm, respectively (see [78], p. 367 and p. 373). Subtracting 2 then gives the eigenvalues of $\boldsymbol{S}$ and finally wanted numbers $\rho_1(\boldsymbol{S}), \rho_2(\boldsymbol{S}), \ldots, \rho_k(\boldsymbol{S})$. As a by-product the LR- and the QR-algorithm supply the eigenvectors of $\boldsymbol{S}$. Therefore and since

the determinant $\det(v_1, \overline{v}_1, \ldots, v_k, \overline{v}_k)$ from Theorem 4.50 can only attain the values $(-1)^k$ and $-(-1)^k$ the apparent handicap for even n does not cause additional serious numerical problems. For further details we refer the reader to special literature treating numerical aspects of matrix eigenvalue problems.

We close this section with an example.

Example 4.53. We evaluate $\int_{SO(2k)} \operatorname{tr}(\boldsymbol{S})\, \mu_{SO(2k)}(\mathrm{d}\boldsymbol{S})$ for all $k \in \mathbb{N}$.
As the trace function is conjugation-invariant

$$\int_{\mathrm{SO}(2k)} \operatorname{tr}(\boldsymbol{S})\, \mu_{\mathrm{SO}(2k)}(\mathrm{d}\boldsymbol{S}) = \int_T \operatorname{tr}(\boldsymbol{T}) h_{\mathrm{SO}(2k)}(\boldsymbol{T})\, \mu_T(\mathrm{d}\boldsymbol{T})$$

$$= \int_{[0,2\pi)^k} \underbrace{\sum_{j=1}^{k} 2\cos\theta_j \prod_{1\le s<t\le k} (\cos\theta_s - \cos\theta_t)^2 \frac{2^{k^2-2k+1}}{k!} (2\pi)^{-k}}_{:=f_0(\theta_1,\ldots,\theta_k)}\, \mathrm{d}\theta_1 \cdots \mathrm{d}\theta_k.$$

As $\cos(\theta_j+\pi) = -\cos\theta_j$ we conclude $f_0(\theta_1+\pi, \ldots, \theta_k+\pi) = -f_0(\theta_1, \ldots, \theta_k)$. As the cosine function is 2π-periodic

$$\int_{[0,2\pi)^k} f_0(\theta_1, \ldots, \theta_k)\, \mathrm{d}\theta_1 \cdots \mathrm{d}\theta_k = \int_{[-\pi,\pi)^k} f_0(\theta_1 + \pi, \ldots, \theta_k + \pi)\, \mathrm{d}\theta_1 \cdots \mathrm{d}\theta_k$$

$$= -\int_{[0,2\pi)^k} f_0(\theta_1, \ldots, \theta_k)\, \mathrm{d}\theta_1 \cdots \mathrm{d}\theta_k.$$

Finally,

$$0 = \int_{[0,2\pi)^k} f_0(\theta_1, \ldots, \theta_k)\, \mathrm{d}\theta_1 \cdots \mathrm{d}\theta_k = \int_{\mathrm{SO}(2k)} \operatorname{tr}(\boldsymbol{S})\, \mu_{\mathrm{SO}(2k)}(\mathrm{d}\boldsymbol{S}). \quad (4.69)$$

4.5 Conjugation-invariant Probability Measures on SO(3)

The special orthogonal group SO(3) consists of all three-dimensional rotation matrices. A probability measure $\nu \in \mathcal{M}^1(\mathrm{SO}(3))$ with $\nu(B) = \nu(\boldsymbol{S}B\boldsymbol{S}^t)$ for all $(\boldsymbol{S}, B) \in \mathrm{SO}(3) \times \mathcal{B}(\mathrm{SO}(3))$ is called *conjugation-invariant*. The SO(3) constitutes the most important representative of the more general examples treated in the preceding sections. In a number of examples we evaluate integrals on SO(3) and consider statistical applications. Our emphasis, however, lies on simulations aspects. We discuss several examples which are partly motivated by applications from the field of computer aided graphical processing. The convolution product of two conjugation-invariant distributions is itself conjugation-invariant. We will derive algorithms and formulas which support both, the efficient simulation of the product of conjugation-invariantly

distributed random variables on SO(3) and the computation of their distribution. Besides theoretical aspects we also consider implementation aspects. In a first step we determine a representation of the factor space $\mathrm{SO}(3)/T$ which is more suitable for concrete computations.

We use the definitions from Section 4.4. In particular, $\mathbf{T}_{(\theta)}$ induces a rotation around the z-axis by angle θ, and a maximal torus subgroup of SO(3) is given by

$$T = \{\mathbf{T}_{(\theta)} \mid 0 \leq \theta < 2\pi\}. \tag{4.70}$$

Lemma 4.54. *(i) The matrix group* SO(3) *acts on* $\mathrm{SO}(3)/T$ *and* S^2 *by left multiplication. The factor space* $\mathrm{SO}(3)/T$ *and the sphere* S^2 *are homogeneous* SO(3)*-spaces.*
(ii) Let s_j *denote the* j^{th} *column of* $\boldsymbol{S} \in \mathrm{SO}(3)$*. The mapping*

$$\mathrm{pr}_{3,T}\colon \mathrm{SO}(3)/T \to \mathrm{S}^2 \qquad \mathrm{pr}_{3,T}(\boldsymbol{S}T) := s_3 \tag{4.71}$$

is an isomorphism of SO(3)*-spaces.*
(iii) Let the mapping $\varphi_3\colon \mathrm{S}^2 \times [0, 2\pi) \to \mathrm{SO}(3)$ *be given by* $\varphi_3(\mathbf{p}, \theta) :=$

$$(1-\cos\theta)\begin{pmatrix} p_1^2 & p_1p_2 & p_1p_3 \\ p_1p_2 & p_2^2 & p_2p_3 \\ p_1p_3 & p_2p_3 & p_3^2 \end{pmatrix} + \begin{pmatrix} \cos\theta & -p_3\sin\theta & p_2\sin\theta \\ p_3\sin\theta & \cos\theta & -p_1\sin\theta \\ -p_2\sin\theta & p_1\sin\theta & \cos\theta \end{pmatrix}. \tag{4.72}$$

Then $\varphi_3(\mathbf{p}, \theta) = q_3({pr_{3,T}}^{-1}(\mathbf{p}), \mathbf{T}_{(\theta)})$ *for all* $(\mathbf{p}, \theta) \in \mathrm{S}^2 \times [0, 2\pi)$ *with* $\mathbf{p} := (p_1, p_2, p_3)$*. In particular,* $\varphi(\mathbf{p}, \theta) \neq 1_3$ *induces a rotation by* θ *around the axis* $\mathbf{p}$ *or, equivalently, a rotation by* $2\pi - \theta$ *around the axis* $-\mathbf{p}$*.*
(iv) The 5*-Tupel* $(\mathrm{SO}(3), \mathrm{S}^2, [0, 2\pi), \mathrm{SO}(3), \varphi_3)$ *has Property (*). In particular*

$$\left(\mu_{(\mathrm{S}^2)} \otimes \tau\right)^{\varphi_3} = \left(\mu_{(SO(3)/T)} \otimes \tau^{is_3^{-1}}\right)^{q_3} \qquad \text{for each } \tau \in \mathcal{M}^1([0, 2\pi)) \tag{4.73}$$

where $\mu_{(\mathrm{S}^2)}$ *denotes the normed geometric surface measure on* S^2*.*

Proof. As the SO(3)-actions on $\mathrm{SO}(3)/T$ and S^2 are transitive both spaces are homogeneous SO(3)-spaces (2.14(v)). For each $\boldsymbol{S}' \in \boldsymbol{S}T$ there exists a number $\theta \in [0, 2\pi)$ with $\boldsymbol{S}' = \boldsymbol{S}\,\mathbf{T}_{(\theta)}$. In particular $s_3 = s_3'$, and hence the mapping $\mathrm{pr}_{3,T}$ is well-defined. The equation $\mathrm{pr}_{3,T}(\boldsymbol{S}_aT) = \mathrm{pr}_{3,T}(\boldsymbol{S}_bT)$ implies $(s_a)_3 = (s_b)_3$, and hence the component $v_{33} = (s_a)_3 \cdot (s_b)_3$ of the matrix product $\mathbf{V} := \boldsymbol{S}_a^{-1}\boldsymbol{S}_b = \boldsymbol{S}_a^t\boldsymbol{S}_b \in \mathrm{SO}(3)$ equals 1. Consequently, $v_{13} = v_{23} = v_{32} = v_{31} = 0$, i.e. $V = \mathbf{T}_{(\beta)}$ for a suitable $\beta \in [0, 2\pi)$. Hence $\boldsymbol{S}_b = \boldsymbol{S}_a\mathbf{T}_{(\beta)} \in \boldsymbol{S}_aT$, i.e. $\mathrm{pr}_{3,T}$ is injective. Let $\mathbf{p} \in \mathrm{S}^2$ while '$\times$' denotes the vector product in $\mathbb{R}^3$ for the moment. Choose a vector $\mathbf{v} \in \mathrm{S}^2$ with $\mathbf{p} \cdot \mathbf{v} = 0$. In particular, $\boldsymbol{S}_v := (\mathbf{v} \times \mathbf{p}, \mathbf{v}, \mathbf{p}) \in \mathrm{SO}(3)$ with $\mathrm{pr}_{3,T}(\mathbf{v} \times \mathbf{p}, \mathbf{v}, \mathbf{p}) = \mathbf{p}$. That is, $\mathrm{pr}_{3,T}$ is also surjective and hence bijective. Finally, for each $\boldsymbol{S}_0 \in \mathrm{SO}(3)$ we have

$$\mathrm{pr}_{3,T}(\boldsymbol{S}_0(\boldsymbol{S}T)) = \mathrm{pr}_{3,T}((\boldsymbol{S}_0\boldsymbol{S})T)) = (\boldsymbol{S}_0\boldsymbol{S})_3 = \boldsymbol{S}_0\mathrm{pr}_{3,T}(\boldsymbol{S}T),$$

i.e. $\mathrm{pr}_{3,T}$ is equivariant which completes the proof of (ii) as the continuity of $\mathrm{pr}_{3,T}$ and its inverse is obvious. The proof of (iii) requires elementary but careful calculations. To improve its readability we use the functions $\delta, \vartheta\colon \{1,2,3\}\times\{1,2,3\} \to \{0,1,-1\}$ for the remainder of this proof. The term $\delta(\cdot,\cdot)$ denotes Kronecker's symbol while

$$\vartheta(i,j) := \begin{cases} 1 & \text{if } (i,j) \in \{(2,1),(1,3),(3,2)\} \\ -1 & \text{if } (i,j) \in \{(1,2),(3,1),(2,3)\} \\ 0 & \text{else.} \end{cases}$$

Moreover, s_{ij} and t_{ij} stand for the components of $\boldsymbol{S}_v$ and $\mathbf{T}_{(\theta)}$, resp. For each pair (i,j) we obtain

$$\begin{aligned} \left(\boldsymbol{S}_v\,\mathbf{T}_{(\theta)}\,\boldsymbol{S}_v^t\right)_{ij} &= \sum_{l=1}^{3} \left(\boldsymbol{S}_v\,\mathbf{T}_{(\theta)}\right)_{il} \left(\boldsymbol{S}_v^t\right)_{lj} = \sum_{k,l=1}^{3} s_{ik}s_{jl}t_{kl} = \sum_{k,l=1}^{2} s_{ik}s_{jl}t_{kl} + s_{i3}s_{j3} \\ &= \cos\alpha \underbrace{\left(s_{i1}s_{j1} + s_{i2}s_{j2}\right)}_{\delta(i,j)-s_{i3}s_{j3}} + \sin\alpha \underbrace{\left(s_{i2}s_{j1} - s_{i1}s_{j2}\right)}_{\vartheta(i,j)s_{m3}} + s_{i3}s_{j3} \\ &= p_i p_j(1-\cos\alpha) + \delta(i,j)\cos\alpha + \vartheta(i,j)p_m\sin\alpha \end{aligned}$$

where $m := \min(\{1,2,3\} \setminus \{i,j\})$. We mention that for $i = j$ the second underbrace is trivially true. If $i \neq j$ it follows from the properties of the vector product and the orthogonality relations between the columns of a rotation matrix. In fact, $(v \times p) \times v = p$, $p \times (v \times p) = v$ and $v \times p = (v \times p)$. The last statement of (iii) follows immediately from $\varphi_3(\mathbf{p},\theta) = \boldsymbol{S}_v\,\mathbf{T}_{(\theta)}\,\boldsymbol{S}_v^t$. As the 5-tupel $(\mathrm{SO}(3), \mathrm{SO}(3)/T, T, \mathrm{SO}(3), q_3)$ has Property $(*)$ statement (iv) is an immediate consequence of (ii) and (iii). □

We point out that the normed geometric surface measure $\mu_{(\mathrm{S}^2)}$ is the unique SO(3)-invariant probability measure on S^2. We recall that $\lambda_{[0,2\pi)}$ and $\lambda_{[-1,3]}$ denote the restriction of the Lebesgue measure on $[0,2\pi)$ or $[-1,3]$, resp.

Theorem 4.55. *Let* $\Phi\colon \mathcal{M}^1([0,2\pi)) \to \mathcal{M}^1_{\mathrm{SO}(3)}(\mathrm{SO}(3))$, $\Phi(\tau) := (\mu_{\mathrm{S}^2} \otimes \tau)^{\varphi_3}$. *Then the following statements are valid:*
(i) $\Phi(\mathcal{M}^1([0,2\pi))) = \mathcal{M}^1_{\mathrm{SO}(3)}(\mathrm{SO}(3))$.
(ii) *(Weyl's integral formula) For*

$$h_3\colon [0,2\pi) \to [0,\infty), \qquad h_3(t) := \frac{1}{2\pi}(1-\cos\theta) \tag{4.74}$$

we have

$$\mu_{\mathrm{SO}(3)} = (\mu_{(\mathrm{S}^2)} \otimes h_3 \cdot \lambda_{[0,2\pi)})^{\varphi_3}. \tag{4.75}$$

(iii) *Let* $\nu \in \mathcal{M}^1_{\mathrm{SO}(3)}(\mathrm{SO}(3))$ *and* $\eta = f \cdot \nu \in \mathcal{M}^1(\mathrm{SO}(3))$ *with conjugation-invariant* ν*-density* f. *Then* $\eta \in \mathcal{M}^1_{\mathrm{SO}(3)}(\mathrm{SO}(3))$, *and* $\Phi(\tau_\nu) = \nu$ *implies* $\Phi(f_{[0,2\pi)} \cdot \tau_\nu) = \eta$ *where* $f_{[0,2\pi)}\colon [0,2\pi) \to \mathbb{R}$ *is given by* $f_{[0,2\pi)}(\theta) := f(\mathbf{T}_{(\theta)})$.
(iv) $(\mathcal{M}^1_{\mathrm{SO}(3)}(\mathrm{SO}(3)), \odot)$ *is a subsemigroup of the convolution semigroup* $(\mathcal{M}^1(\mathrm{SO}(3)), \odot)$. *In particular* $\varepsilon_{1_3} \in \mathcal{M}^1_{\mathrm{SO}(3)}(\mathrm{SO}(3))$.

Proof. Theorem 4.55 is an immediate consequence of Theorem 4.49 and Lemma 4.54(iv) as $\mu_T^{is_3} = (2\pi)^{-1}\,\lambda_{[0,2\pi)}$. □

Theorem 4.56. (i) $E_T(\{t\}) = \{t, 2\pi - t\}$ *for* $t \in (0, 2\pi)$ *and* $E_T(\{0\}) = \{0\}$. *In particular, the density* $\mathrm{h}_{SO(3)}$ *is constant on* $E_T(\{t\})$ *for each* $t \in [0, 2\pi)$.
(ii) *The subset*

$$\mathrm{R}_{SO(3)} := [0, \pi] \subseteq [0, 2\pi) \tag{4.76}$$

is a measurable section. The corresponding mapping

$$\psi_3 \colon \mathrm{SO}(3) \to R_{\mathrm{SO}(3)}, \qquad \psi_3(\boldsymbol{S}) := \rho_1(\boldsymbol{S}) \tag{4.77}$$

is continuous.
(iii) *Suppose that* $\tau_\nu \in \Phi^{-1}(\nu)$, *and let*

$$\tau_\nu^*(B) := \frac{1}{2}\left(\tau_\nu(B) + \tau_\nu(2\pi - B)\right) \qquad \textit{for each } B \in \mathcal{B}([0, 2\pi)) \tag{4.78}$$

where we identify 2π *with* 0. *Then* $\tau_\nu^* \in \Phi^{-1}(\nu)$. *If* $\tau_\nu = f \cdot \lambda_{[0,2\pi)} \in \Phi^{-1}(\nu)$ *then*

$$\tau_\nu^* := f_W^* \cdot \lambda_{[0,2\pi)} \qquad \textit{with} \qquad f_W^*(t) := \frac{1}{2}\left(f(t) + f(2\pi - t)\right). \tag{4.79}$$

(iv) *Let* $\nu = (\mu_{SO(3)} \otimes f \cdot \lambda_{[0,2\pi)})^{\varphi_3}$ *and let* $f_{\mathrm{R}_{SO(3)}} \colon [0, 2\pi) \to [0, \infty]$ *be given by* $f_{\mathrm{R}_{SO(3)}}(t) := 1_{[0,\pi]}(t)(f(t) + f(2\pi - t))$. *Then*

$$\nu = \left(\mu_{\mathrm{S}^2} \otimes f_{\mathrm{R}_{SO(3)}} \cdot \lambda_{[0,2\pi)}\right)^{\varphi_3}. \tag{4.80}$$

In particular, $\tau_\nu := f_{\mathrm{R}_{SO(3)}} \cdot \lambda_{[0,2\pi)}$ *has its total mass on* $[0, \pi]$, *and* $(h_3)_{R_{\mathrm{SO}(3)}}(t) := 1_{[0,\pi]} \frac{1}{\pi}(1 - \cos t)$.
(v) *The mapping* $\rho_1^m \colon \mathrm{SO}(3)^m \to [0, \pi]^m \subseteq [0, 2\pi)^m$ *is a sufficient statistic for each test problem* $(\mathrm{P}_{\Gamma_0}, \mathrm{P}_{\Gamma \setminus \Gamma_0})$ *with* $\mathrm{P}_\Gamma \subseteq \mathcal{M}^1_{SO(3)}(\mathrm{SO}(3))^m$. *The mapping* $\mathrm{tr}^m \colon \mathrm{SO}(3)^m \to [-1, 3]^m$ *is also a sufficient statistic. If* $\nu, \eta \in \mathcal{M}^1_{SO(3)}(\mathrm{SO}(3))$ *with* $\eta \ll \nu$ *then there exists a measurable function* $f_0 \colon [-1, 3] \to [0, \infty]$ *with* $\eta = (f_0 \circ \mathrm{tr}) \cdot \nu$. *Further,* $\eta^{\mathrm{tr}} = f_0 \cdot \nu^{\mathrm{tr}}$ *and*

$$\mu_{SO(3)}^{\mathrm{tr}} = h_{0;3} \cdot \lambda_{[-1,3]} \qquad \textit{with } h_{0;3}(x) := \sqrt{3 - x}/2\pi\sqrt{1 + x}. \tag{4.81}$$

Proof. Statement (i) is an immediate consequence from (ii), 4.54 and the fact that eigenvalues are invariant under conjugation. The statements (ii), (iii) and (iv) follow from 4.50(i),(iv), 4.51(iii), 4.51(v), and the fact that $\rho_1(\boldsymbol{S}) = \arccos((\mathrm{tr}(\boldsymbol{S}) - 1)/2)$. The first assertion of (v) follows from (ii) and 4.14(i). Since $g \colon [0, \pi] \to [-1, 3]$, $g(y) := 2\cos y + 1$, is a homeomorphism the mapping $(g \circ \rho_1)^m = \mathrm{tr}^m$ is a sufficient statistic for P_Γ as ρ_1^m is sufficient. Since $\eta \ll \nu$ there exists a conjugation-invariant ν-density f with $\eta = f \cdot \nu$. As the trace function is constant on the SO(3)-orbits $f_0 := f \circ is_3^{-1} \circ g^{-1}$ is

such a density. (Note that $is_3^{-1} \circ g^{-1} \circ \mathrm{tr}(\boldsymbol{S}) = \boldsymbol{T}_{\rho_1(\mathbf{S})}$.) Assertion (iv) and the transformation theorem in $\mathbb{R}$ finally imply

$$h_{0;3}(x) = 2h_3(g(x)) \cdot |g'(x)| = \frac{2h_3(\arccos((x-1)/2))}{2\sqrt{1-((x-1)/2)^2}} = \frac{\sqrt{3-x}}{2\pi\sqrt{1+x}}. \qquad \square$$

Besides $\rho_1(\boldsymbol{S})^m \colon \mathrm{SO}(3)^m \to [0,\pi]^m$ also $\mathrm{tr}^m \colon \mathrm{SO}(3)^m \to [-1,3]^m$ is a sufficient statistic for all test problems $(\mathrm{P}_{\Gamma_0}, \mathrm{P}_{\Gamma\setminus\Gamma_0})$ with $\mathrm{P}_\Gamma \subseteq \mathcal{M}^1_{SO(3)}(\mathrm{SO}(3))^m$. Of course, from the theoretical point of view both mappings are equivalent but the second is clearly more favourable for concrete applications. The computation of $\mathrm{tr}(\boldsymbol{S})$ requires no more than two additions whereas the computation of $\rho_1(\boldsymbol{S})$ needs a subtraction, a division by two and, above all, the calculation of an arcus function.

Remark 4.57. (i) At first sight it may be surprising that the Haar measure $\mu_{\mathrm{SO}(3)} \neq (\mu_{(SO(3)/T)} \otimes (2\pi)^{-1} \cdot \lambda_{[0,2\pi)})^{\varphi_3}$. However, this becomes clear when one recalls the following: If Z is a $\mu_{\mathrm{SO}(3)}$-distributed random rotation then the random vector $Z\mathbf{v}$ is equidistributed on S^2 for each $\mathbf{v} \in \mathrm{S}^2$. If $\psi(\mathbf{p},\theta)\mathbf{v} \in \mathrm{S}^2$ is contained in a small neighbourhood of the antipodal point $-\mathbf{v}$ then θ must be close to π and $\mathbf{p}$ must 'almost' be perpendicular to $\mathbf{v}$. On the other hand there exist axes $\mathbf{p}$ for which no angle θ exists which maps $\mathbf{v}$ into this neighbourhood. As $Z\mathbf{v}$ is equidistributed on S^2 this implies a preference for angles close to π - just as it is the case for the density $\mathrm{h}_{\mathrm{SO}(3)}$.
(ii) Once again we remind on the convention saying that pseudorandom elements (pseudorandom vectors, pseudorandom numbers) $\widetilde{V}_1, \widetilde{V}_2, \ldots$ are 'τ-distributed'. More precisely, one should say that $\widetilde{V}_1, \widetilde{V}_2, \ldots$ are values which have similar statistical properties as 'true' realizations of independent τ-distributed random variables (cf. Remark 3.1).
(iii) A rotation around the axis $\mathbf{p}$ by angle θ can also be represented as a conjugation with a unit quaternion $qu(\mathbf{p},\theta) := (\cos(\theta/2), \sin(\theta/2) \cdot \mathbf{p})$ (cf. Remark 4.61). In particular, Theorem 4.55 can also be used to generate pseudorandom unit quaternions.

The following example treats integration problems on SO(3) with conjugation-invariant integrands. Applying (3.3) these integrals can be reduced to integrals on $[0, 2\pi)$. Note that these results are used to determine normalization constants for the simulation problems considered in Example 4.59.

Example 4.58.

$$\int_{\mathrm{SO}(3)} \mathrm{tr}(\boldsymbol{S})^2 \, \mu_{\mathrm{SO}(3)}(\mathrm{d}\boldsymbol{S}) = \int_0^{2\pi} \frac{(1+2\cos\theta)^2(1-\cos\theta)}{2\pi} \, \mathrm{d}\theta \qquad (4.82)$$

$$= \int_0^{2\pi} \frac{(1+3\cos\theta-4\cos^3\theta)}{2\pi} \, \mathrm{d}\theta = 1.$$

$$\int_{\mathrm{SO}(3)} |\mathrm{tr}(\boldsymbol{S})| \, \mu_{\mathrm{SO}(3)}(\mathrm{d}\boldsymbol{S}) = \int_0^{2\pi} \frac{|1+2\cos\theta|(1-\cos\theta)}{2\pi} \, \mathrm{d}\theta \qquad (4.83)$$

$$= \frac{\sqrt{27}}{2\pi}.$$

$$\int_{SO(3)} \frac{1}{3 - \operatorname{tr}(\boldsymbol{S})}\, \mu_{SO(3)}(\mathrm{d}\boldsymbol{S}) = \int_0^{2\pi} \frac{1 - \cos\theta}{2\pi(3 - (2\cos\theta + 1))}\, \mathrm{d}\theta \qquad (4.84)$$

$$= \int_0^{2\pi} \frac{1}{4\pi}\, \mathrm{d}\theta = \frac{1}{2}.$$

Applying Theorem 4.55 enables a decomposition of conjugation-invariant simulation problems on SO(3) into independent simulation problems on S^2 and $[0, 2\pi)$. The identification of $\mathrm{SO}(3)/T$ with S^2 favours concrete computations.

In Example 4.59 we compare the efficiency of two classes of simulation algorithms for conjugation-invariant distributions on SO(3). As the first class exploits Theorem 4.55 we suggestively call them *Torus-algorithms*. The algorithms contained in the second class use Euler's angles ([47], pp. 215). Already Leonard Euler realized that each rotation in $\mathbb{R}^3$ can be expressed as a composition of three 'elementary' rotations around the z-, the x-, and again around the z-axis. Using Euler angles the simulation problem can be transferred from SO(3) to the quader $Q := [0, 2\pi) \times [0, \pi] \times [0, 2\pi)$. Algorithms based on Euler's angles are suggestively called *Euler algorithms*. In the following $\mathrm{Eu}(\alpha, \beta, \gamma)$ denotes the composition of three rotations around the z-axis, the x-axis and again around the z-axis by the angles γ, β and α, resp. In particular, for each (not necessarily conjugation-invariant) $\mu_{SO(3)}$-density f we conclude from [13], p. 53 (Exercise 6) that

$$f \cdot \mu_{SO(3)} = (f_Q \cdot \lambda_{3;Q})^{\mathrm{Eu}} \quad \text{with } f_Q(\alpha, \beta, \gamma) := \frac{[(f \circ \mathrm{Eu})(\alpha, \beta, \gamma)] \cdot \sin\beta}{8\pi^2}. \qquad (4.85)$$

Torus- and Euler-algorithms simplify conjugation-invariant simulation problems on SO(3) to independent simulation problems on S^2 and $[0, 2\pi)$, or they transfer simulation problems from SO(3) to $Q \subseteq \mathbb{R}^3$, resp. What we essentially need are hence efficient simulation algorithms on $[0, 2\pi)$ and Q, resp. In many applications the transformed distributions have Lebesgue densities. Unless these densities have very specific properties at least on Q only the well-known acceptance-rejection algorithm is at disposal ([18], pp. 40 ff.) or modifications thereof ([18], pp. 47 ff.). The simplest form of an acceptance-rejection algorithm for the simulation of a Lebesgue density $f\colon A \subseteq \mathbb{R}^k \to [0, \infty]$ is given below. Recall that standard random numbers are viewed as realizations of iid $\lambda_{[0,1)}$-distributed random variables.

Acceptance-rejection algorithm

0. Choose a Lebesgue density $g\colon A \to [0, \infty]$ and compute the value $c := \sup_{\mathbf{x} \in A} f(\mathbf{x})/g(\mathbf{x})$.

1. Generate independently a $g \cdot \lambda_k$-distributed pseudorandom vector $\widetilde{V}$ on A and a standard random number $\widetilde{U} \in [0,1)$.
2. If $\widetilde{U} \cdot c \cdot g(\widetilde{V}) > f(\widetilde{V})$ goto 1. Otherwise accept $\widetilde{V}$ as a $f \cdot \lambda_k$-distributed pseudorandom vector.
3. END.

In average the generation of a $f \cdot \lambda_k$-distributed pseudorandom vector requires c many $g \cdot \lambda_k$-distributed pseudorandom vectors and c standard random numbers. Consequently, the number c should be small whereas the density g should be chosen that $g \cdot \lambda_k$-distributed pseudorandom vectors can be generated efficiently. Unfortunately, these criteria usually counteract each other. If A and f are bounded often the constant density $g \equiv \lambda_k(A)^{-1}$ is used which meets the second criterion perfectly. On the negative side the constant c may be very large.

Example 4.59. In this example we compare the efficiency of Torus- and Euler algorithms at hand of five conjugation-invariant distributions $\nu_1, \ldots, \nu_5$. In a first step we determine probability measures $\tau_j \in \mathcal{M}^1([0,2\pi))$ with $\nu_j = (\mu_{(\mathrm{S}^2)} \otimes \tau_j)^{\varphi_3}$ and $\kappa_j \in \mathcal{M}^1(Q)$ with ${\kappa_j}^{\mathrm{Eu}} = \nu_j$, resp. The distributions $\nu_1, \ldots, \nu_4$ have conjugation-invariant $\mu_{SO(3)}$-densities and hence $\tau_1, \ldots, \tau_4$ and $\kappa_1, \ldots, \kappa_4$ have Lebesgue densities $f_{j,[0,2\pi)}$ and $f_{j,Q}$, resp. These densities follow immediately from 4.55(ii),(iii) and from (4.85), resp. The normalizing constants have already been computed in Example 4.58. As the computations in (i) to (iv) require no more than elementary algebraic transformations and trigonometric identities (see, e.g. [14], 2.5.2.1.3) we abstain from further comments (cf. also [67], p. 529 ff.). We recall that $Q := [0,2\pi) \times [0,\pi] \times [0,2\pi)$.
(i) Let $\nu_1 := \mu_{SO(3)}$. From 4.55(ii) and [13], p. 53 (Exercise 6), we obtain

$$f_{1,[0,2\pi)}(\theta) = \frac{1}{2\pi}(1-\cos\theta) \qquad \text{and} \qquad f_{1,Q}(\alpha,\beta,\gamma) = \frac{\sin\beta}{8\pi^2}. \tag{4.86}$$

(ii) Let $\nu_2 := f_2 \cdot \mu_{SO(3)}$ with $f_2(\boldsymbol{S}) = 2\pi|\mathrm{tr}\boldsymbol{S}|/\sqrt{27}$. Then

$$f_{2,[0,2\pi)}(\theta) = \frac{1}{\sqrt{27}}|1+2\cos\theta|(1-\cos\theta) \qquad \text{and} \tag{4.87}$$

$$f_{2,Q}(\alpha,\beta,\gamma) \tag{4.88}$$

$$\begin{aligned}
&= \frac{\sin\beta}{4\pi\sqrt{27}}|\cos\alpha\cos\gamma - \sin\alpha\cos\beta\sin\gamma - \sin\alpha\sin\gamma + \cos\alpha\cos\beta\cos\gamma + \cos\beta| \\
&= \frac{\sin\beta}{4\pi\sqrt{27}}\,|(\cos\alpha\cos\gamma - \sin\alpha\sin\gamma)(\cos\beta+1) + \cos\beta| \\
&= \frac{\sin\beta}{4\pi\sqrt{27}}\,|\cos(\alpha+\gamma)(\cos\beta+1) + \cos\beta|\,.
\end{aligned}$$

(iii) Let $\nu_3 := f_3 \cdot \mu_{SO(3)}$ with $f_3(\boldsymbol{S}) = (\mathrm{tr}\boldsymbol{S})^2$. Then

$$f_{3,[0,2\pi)}(\theta) = \frac{1}{2\pi}(1 + 3\cos\theta - 4\cos^3\theta) = \frac{1}{2\pi}(1 - \cos(3\theta)) \text{ and} \tag{4.89}$$

$$f_{3,Q}(\alpha,\beta,\theta) = \frac{\sin^2\beta}{8\pi^2}\left(\cos(\alpha+\gamma)(\cos\beta+1) + \cos\beta\right)^2. \tag{4.90}$$

(iv) Let $\nu_4 := f_4 \cdot \mu_{SO(3)}$ with $f_4(\boldsymbol{S}) = 2/(3 - \mathrm{tr}\boldsymbol{S})$. Then

$$f_{4,[0,2\pi)}(\theta) = \frac{2(1-\cos\theta)}{2\pi(3-(2\cos\theta+1))} = \frac{1}{2\pi} \qquad \text{and} \tag{4.91}$$

$$\begin{aligned} f_{4,Q}(\alpha,\beta,\theta) &= \frac{2\sin\beta}{4\pi^2\left(4-(\cos(\alpha+\gamma)+1)(\cos\beta+1)\right)} \\ &= \frac{\sin\beta}{8\pi^2\left(1-\cos^2((\alpha+\gamma)/2)\cos^2(\beta/2)\right)}. \end{aligned} \tag{4.92}$$

(v) Let $B_s := \{\boldsymbol{S} \in \mathrm{SO}(3) \mid \mathrm{tr}\boldsymbol{S} = s\}$ with $s \in [-1,3]$. As the trace function is conjugation-invariant $B_s = E_{SO(3)}(B_s)$. From 4.56(ii) we conclude that B_s is an SO(3)-orbit. As the trace function is continuous the set B_s is closed, hence compact and contained in $\mathcal{B}_{SO(3)}(\mathrm{SO}(3))$. In particular there exists a unique conjugation-invariant probability measure ν_5 on SO(3) with total mass on B_s, namely $\nu_5 = (\mu_{(\mathrm{S}^2)} \otimes \varepsilon_{\arccos((s-1)/2)})^{\varphi_3}$. For $s \in (-1,3)$ the subset B_s is a two-dimensional submanifold of SO(3). Clearly, the total mass of a probability measure $\kappa_5 \in \mathcal{M}^1(Q)$ with $\kappa_5^{Eu} = \nu_5$ is concentrated on a Lebesgue zero set. Consequently, for the simulation of κ_5 we needed a two-dimensional parametrization of B_s.

The computation of $\mathrm{Eu}(\alpha,\beta,\gamma)$ costs the time-consuming evaluation of six trigonometric functions whereas $\varphi_3(\mathbf{p},\theta)$ requires only two. Empirical studies have shown that the computation of $\mathrm{Eu}(\alpha,\beta,\gamma)$ needs about 160 per cent relative to the time needed for $\varphi_3(\mathbf{p},\theta)$. Table 4.1 supplies a comparison of the average computation times needed for one ν_j-distributed pseudorandom matrix on SO(3). The programs were written in Turbo Pascal and ran on a 486 MHz processor using a numerical coprocessor ([71], p. 77). As the absolute running times depend essentially on the used machine we divided them by the average time needed by the Torus-algorithm for the generation of a $\mu_{SO(3)}$-distributed pseudorandom matrix.

For the simulation of μ_{S^2} a number of suitable algorithms are well-known. We used an algorithm proposed by Marsaglia ([53], 'the new method') which exploits the following observation: If the random vector (V_1, V_2) is equidistributed on the unit circle and $S := V_1^2 + V_2^2$ then the random vector $(2V_1\sqrt{1-S}, 2V_2\sqrt{1-S}, 1-2S)$ is μ_{S^2}-distributed. For the simulation of the densities $f_{j,[0,2\pi)}$ and $f_{j,Q}$ we used simple acceptance-rejection algorithms (see above or [18], pp. 41f.) with constant dominating densities g_j or $g_{j,Q}$, resp.

For the Torus algorithms the average running times for ν_1, ν_2 and ν_3 are almost equal whereas for the Euler algorithms these times increase notably. This phenomenon is caused by the fact that the average rejection

Table 4.1. Normed average running times

	ν_1	ν_2	ν_3	ν_4	ν_5
Torus	1.00	1.22	1.01	0.72	0.48
Euler	1.13	2.83	4.14	—	—

rates for $f_{2,Q}$ and $f_{3,Q}$ are much higher than those for the simulation of the one-dimensional densities $f_{2,[0,2\pi)}$ and $f_{3,[0,2\pi)}$. The Euler-algorithms may possibly be speeded up by using more sophisticated dominating densities $g_{j,Q}$ (e.g. piecewise constant densities). However, even if the rejection rate decreases this usually makes the setup steps more costly. The generation of $g_{j,Q} \cdot \lambda_3$-distributed pseudorandom matrices becomes more complicated and hence more time-consuming. We should not expect a considerable gain of efficiency. Anyway, as the necessary calculations and possible distinctions of cases are less complicated in $\mathbb{R}$ than in $\mathbb{R}^3$ it seems to be more likely that the Torus algorithms can be accelerated in this fashion. Clearly, the simulation of ν_5 is the fastest of all since we merely have to simulate μ_{S^2} rather than μ_{S^2} and $f_{j,[0,2\pi)}$. Further, the terms $\cos(\arccos((s-1)/2)) = (s-1)/2$ and $\sin(\arccos((s-1)/2))$ have to be computed only once. For ν_4 the constant density $f_{4,[0,2\pi)} \equiv 1/2\pi$ can also be simulated very efficiently. We have already pointed out that Euler angles are no appropriate tool for the simulation of ν_5. The same is true for ν_4 as $f_{4,Q}(\alpha,\beta,\gamma) \le f_{4,Q}(0,\beta,0) = 2\cos(\beta/2)\sin(\beta/2)/8\pi^2\sin^2(\beta/2) = \cot(\beta/2)/4\pi^2$, i.e. the density $f_{4,Q}$ has a pole in $(0,0,0)$. Moreover, although $\int_Q f_{4,Q}(\alpha,\beta,\gamma)\,\mathrm{d}\alpha\mathrm{d}\beta\mathrm{d}\gamma = 1$ for the constant dominating density $\cot(\beta/2)/4\pi^2$ we obtain $\int_Q \cot(\beta/2)/4\pi^2\,\mathrm{d}\alpha\mathrm{d}\beta\mathrm{d}\gamma = \lim_{\epsilon\to 0}\int_{2\epsilon}^{\pi}\cot(\beta/2)\,\mathrm{d}\beta = \lim_{\epsilon\to 0} -2\ln(\sin\epsilon) = \infty$. There does not seem to exist an efficient algorithm to simulate the density $f_{4,Q}$.

Marsaglia's method needs about 2.55 standard random numbers per pseudorandom vector on S. The simulation of the densities $f_{j,[0,2\pi)}$ and $f_{j,Q}$ on $[0,2\pi)$ and $Q \subseteq \mathbb{R}^3$, resp., with acceptance-rejection algorithms (using constant densities $g_{j,[0,2\pi)}$ and $g_{j,Q}$) need $2 \cdot 2\pi \cdot \sup_{\theta\in[0,2\pi)} f^*_{j,[0,2\pi)}(\theta)$ or $4 \cdot 4\pi^3 \cdot \sup_{(\alpha,\beta,\gamma)\in Q} f_{j,Q}(\alpha,\beta,\gamma)$ standard random numbers per pseudorandom matrix. Applying a Torus algorithm the generation of a ν_j-distributed pseudorandom matrix requires 6.55, 7.38, 6.55, 3.55 or 2.55 standard random numbers, resp., in average. The Euler algorithms need 6.28, 13.37 and 25.90, standard random numbers, resp.

The mapping $\mathrm{Eu}\colon Q = [0,2\pi)\times[0,\pi]\times[0,2\pi) \to \mathrm{SO}(3)$ is surjective but not injective. Clearly, the restriction $\mathrm{Eu}_{|(0,2\pi)\times(0,\pi)\times(0,2\pi)}$ is a homeomorphism onto its image, and $\mu_{SO(3)}(\mathrm{SO}(3) \setminus \mathrm{Eu}((0,2\pi)\times(0,\pi)\times(0,2\pi))) = 0$. Concerning the simulation of conjugation-invariant distributions on SO(3) the Torus algorithms are clearly superior to Euler algorithms. Torus algorithms

are suitable for the simulation of conjugation-invariant distributions whose total mass is concentrated on conjugation-invariant subsets of SO(3). However, unlike the Torus algorithms the Euler algorithms can also be applied for the simulation of distributions on SO(3) which are not conjugation-invariant.

Remark 4.60. (i) The rotation matrix $\varphi_3(\mathbf{p}, \theta)$ induces a rotation around the axis $\mathbf{p}$ by the angle θ or equivalently, a rotation around the axis $-\mathbf{p}$ by the angle $2\pi - \theta$ (4.54(iii)). We point out that this insight is not new and already mentioned in [56], p. 466. (Note that in [56] elements of $\mathbb{R}^3$ are interpreted as row vectors and hence the transposed matrix $\varphi_3(\mathbf{p}, \theta)^t$ is considered.) As the mapping φ_3 is not injective a representation $\boldsymbol{S} = \varphi_3(\mathbf{p}, \theta)$ is not unique. In fact $\varphi_3(\mathbf{p}, \theta) = \varphi_3(-\mathbf{p}, 2\pi - \theta)$ and $1_3 = \varphi_3(\mathbf{p}, 0)$ for all $\mathbf{p} \in \mathrm{S}^2$. Apart from [66, 67, 70] the transformation of measures by φ_3 has apparently not been studied in literature.
(ii) The algorithms proposed by Stewart and Heiberger ([77, 34], cf. also Section 4.4) can also be applied to simulate the Haar measure $\mu_{SO(3)}$. However, as these algorithms require nine normally distributed pseudorandom numbers, resp. six normally distributed pseudorandom numbers, per pseudorandom matrix they can not beat the Torus algorithm. Moreover, the algorithms of Stewart and Heiberger can only be used for the simulation of the Haar measure $\mu_{SO(3)}$. Therefore we did not consider them in Example 4.59.
(iii) Moreover, algorithms for the simulation of $\mu_{SO(3)}$ are known which are based on geometric considerations ([4, 76]). WE mention that also false approaches have been published which ground on the erroneous assumption that $\mu_{SO(3)}$ induces equidistributed angles (cf. [76] and 4.57(i)). As the algorithms of Stewart and Heiberger they can only be used for the simulation of $\mu_{SO(3)}$.

In the introductory paragraph of this section we have already pointed at applications from computer aided graphical processing. In the sequel we summarize the central results from [70]. Suppose for the moment that U and V are independent random variables where $U \sim \mu_{\mathrm{S}^2}$ while V assumes values on the interval $[0, 2\pi)$. By 4.54(iii) the random variable $\varphi(U, V)$ is conjugation-invariantly distributed, and U and V can be interpreted as the normed oriented random axis and the random angle, resp. Vice versa, by 4.55(i) each conjugation-invariantly distributed random variable Z can be represented in this way.

In many video games, for example, target objects are rotated randomly to test the reaction and the deftness of the player. Therefore, the player should have minimal information on the future actions of the program. It seems to be reasonable to choose equidistributed random rotation axes being independent of the angle. (Note that these natural requirements characterize the conjugation-invariant probability measures.) In the simplest case the random rotations are iid $(\mu_{\mathrm{S}^2} \otimes \tau)^{\varphi_3}$-distributed and carried out with constant velocity. The player has maximal knowledge of the future actions, of course, if τ is a Dirac measure since he then at least knows all future times when a new

rotation begins. In contrast, as any angle is equally likely, the equidistribution on $[0, 2\pi)$ gives minimal information on the end of the actual rotation, that is, on the time when the axis of rotation changes next. If the random rotations are independently $\mu_{SO(3)}$-distributed the knowledge on the outcome of the subsequent rotations is minimal. In fact, for each $\mathbf{v} \in \mathrm{S}^2$ any point on S^2 is equally likely to be its image under any of the subsequent rotations. Consequently, the future beyond the next change of axis is absolutely unpredictable in this case and the player cannot work out any reasonable long-time strategy with which he could deceive the program about his reaction time and deftness.

We have already noted that a representation $\boldsymbol{S} = \varphi_3(\mathbf{p}, \theta)$ is not unique. If not only the outcome of a rotation $\mathbf{v} \mapsto \boldsymbol{S}\mathbf{v}$ is relevant but also in its course (as in video game applications, for example) one clearly can distinguish between $\varphi_3(\mathbf{p}, \theta)$ and $\varphi_3(-\mathbf{p}, 2\pi - \theta)$ when observing the intermediate values $\varphi_3(\mathbf{p}, \theta/n)\mathbf{v}, \ldots, \varphi_3(\mathbf{p}, (n-1)\theta/n)\mathbf{v}$. In particular, the choice of a particular distribution $\tau_\nu \in \Phi^{-1}(\nu)$ causes visual differences.

Remark 4.61. The representation $\boldsymbol{S} = \varphi_3(\mathbf{p}, \theta)$ supports a division of the angle. Compared with Euler angles this represents a further advantage. As quaternions have the same pleasant property they are often used to carry out rotations (see, e.g. [57]). The rotation which is induced by the matrix $\varphi_3(\mathbf{p}, \theta)$ corresponds with a conjugation with the unit quaternion $qu(\mathbf{p}, \theta) := (\cos(\theta/2), \sin(\theta/2) \cdot \mathbf{p})$ where we identify $\mathbf{v} \in \mathbb{R}^3$ with the quaternion $(0, \mathbf{v}) \in \mathbb{R}^4$. The composition of two rotation matrices corresponds with the multiplication of the respective unit quaternions. The multiplication of two unit quaternions requires fewer arithmetical operations than the multiplication of two (3×3)-matrices. On the other hand the conjugation with a unit quaternion costs more operations than a matrix-vector multiplication. If one has to compute at least five images of the same rotation the use of φ_3-matrices is more efficient than the use of quaternions. Note that both, a φ_3-matrix representation and the representation by a unit quaternion use the same input parameters (axis and angle). As a consequence both can be used at the same time. For example, one could store the composition of all previous rotations as a quaternion. This facilitates a regular normalization (by dividing this quaternion by its norm). Moreover, the quaternion representation enables a smooth composition of rotations without abrupt changes of the axes ([75, 57]). The intermediate states $\varphi_3(\mathbf{p}, k\theta/n)\mathbf{v}$, $\varphi_3(\mathbf{p}, \theta/n)\varphi_3(\mathbf{p}, k\theta/n)\mathbf{v}, \ldots, \varphi_3(\mathbf{p}, \theta/n)^{n-2k}\varphi_3(\mathbf{p}, k\theta/n)\mathbf{v}$, however, can be computed using the matrix $\varphi_3(\mathbf{p}, \theta/n)$. We will not pursue this possible interaction of matrices and quaternions as it lies beyond the scope of this book.

Among the pre-images $\Phi^{-1}(\nu)$ there are two outstanding distributions, namely the 'asymmetric' distribution $\tau_{\nu;\,a} := \nu^\psi$ which is characterized by the condition $\tau_{\nu;\,a}([0, \pi]) = 1$ and the 'symmetric' distribution $\tau_{\nu;s} := \tau_\nu^*$ which satisfies $\tau_{\nu;\,s}(I) = \tau_{\nu;\,s}(2\pi - I)$ for each interval $I \subseteq (0, \pi)$, resp. (cf.

4.56). Suppose that random number $\widetilde{Y}$ is τ'-distributed with $\tau' \in \Phi^{-1}(\nu)$. Then the pseudorandom number $\widetilde{Y}' := \min\{\widetilde{Y}, 2\pi - \widetilde{Y}\}$ is $\tau_{\nu;\,a}$-distributed. If we set $\widetilde{Y}'' := \widetilde{Y}$ or $\widetilde{Y}'' := 2\pi - \widetilde{Y}$, resp., both with probability 1/2 (and identify 2π with 0) then $\widetilde{Y}''$ can be viewed as τ_ν^*-distributed.

Next, we consider the convolution product of conjugation-invariant probability measures. In specific applications one might be interested in the product $Z := Z_2 Z_1$, or more generally, in $Z := Z_k Z_{k-1} \cdots Z_1$ where $Z_1, Z_2, \ldots, Z_k$ denote independent random rotations. If only their composition Z is relevant (which might be the case for simulations of specific stochastic processes or when deriving empirical estimators for the convergence properties of finite sequences of convolution products) it appears to be desirable to find a more efficient method than generating pseudorandom rotation after pseudorandom rotation. For arbitrary distributions η_j it is rather tough even to determine the distribution of $Z_2 Z_1$, i.e. to compute the convolution product $\eta_2 \odot \eta_1$. If $\eta_j \in \mathcal{M}^1_{\mathrm{SO}(3)}(\mathrm{SO}(3))$ for all $j \le k$, however, this problem becomes much easier: Theorem 4.55(iv) says that $\nu_2 \odot \nu_1$ and, by a simple induction argument, that $\nu_k \odot \cdots \odot \nu_1$ are conjugation-invariant. That is, there exist probability measures τ_{21} and $\tau_{k\ldots1} \in \mathcal{M}^1([0, 2\pi))$ with $\nu_2 \odot \nu_1 = (\mu_{(\mathrm{S}^2)} \otimes \tau_{21})^{\varphi_3}$ and $\nu_k \odot \cdots \odot \nu_1 = (\mu_{(\mathrm{S}^2)} \otimes \tau_{k\ldots1})^{\varphi_3}$, resp. We point out that $\mu_{SO(3)} \odot \cdots, \odot \mu_{SO(3)} = \mu_{SO(3)}$.

Clearly, it suffices to determine any distribution $\tau' \in \Phi^{-1}(\eta_2 \odot \eta_1)$ or at least to apply an efficient algorithm to simulate τ'. In this connection we point out that the mapping $\chi\colon [0, 2\pi) \to (-1, 1]$, $\chi(\alpha) := \cos(\alpha/2)$ is bi-measurable and hence $\tau \mapsto \tau^\chi$ induces a $1-1$-correspondence between $\mathcal{M}^1([0, 2\pi))$ und $\mathcal{M}^1((-1, 1])$. Clearly, ${\varepsilon_\alpha}^\chi = \varepsilon_{\chi(\alpha)}$, and the transformation theorem ([18], pp. 11 ff.) implies

$$(g \cdot \lambda_{[0,2\pi)})^\chi = \bar{g} \cdot \lambda_{(-1,1]} \qquad \text{with } \bar{g}(u) := 2g(2\arccos u)/\sqrt{1-u^2}. \quad (4.93)$$

These transformation rules cover the distributions on $[0, 2\pi)$ of practical relevance, namely distributions having a Lebesgue density, discrete distributions and convex combinations of both types. In the sequel we assume $0 \le \alpha, \beta, \gamma < 2\pi$.

Theorem 4.62. (i)*Let* $\varphi_3(\mathbf{r}, \gamma) := \varphi_3(\mathbf{q}, \beta)\varphi_3(\mathbf{p}, \alpha)$. *Then*

$$\mathrm{tr}(\varphi_3(\mathbf{r}, \gamma)) = 4\left(\mathbf{q} \cdot \mathbf{p} \sin(\beta/2) \sin(\alpha/2) - \cos(\beta/2)\cos(\alpha/2)\right)^2 - 1. \quad (4.94)$$

In particular,

$$\cos(\gamma/2) = \pm\left(\mathbf{q} \cdot \mathbf{p} \sin(\beta/2) \sin(\alpha/2) - \cos(\beta/2)\cos(\alpha/2)\right). \quad (4.95)$$

(ii) *Suppose that* U, Y_1 *and* Y_2 *denote independent random variables where* U *is equidistributed on* $(-1, 1]$ *while* Y_1 *and* Y_2 *are* τ_1- *and* τ_2-*distributed, resp., with* $\tau_j \in \Phi^{-1}(\nu_j)$. *Let* ${\tau_{21}}^\chi$ *denote the distribution of the random variable*

$$U \sin(Y_2/2) \sin(Y_1/2) - \cos(Y_2/2) \cos(Y_1/2). \tag{4.96}$$

Then ${\tau_{21}}^{\chi} \in (\Phi^{-1}(\nu_2 \odot \nu_1))^{\chi}$.
(iii) *Assume that the random variables* U, W_1, W_2 *are independent where* W_j *is* τ_j^{χ}*-distributed. Then the random variable*

$$U\sqrt{1-W_2^2}\sqrt{1-W_1^2} - W_2W_1 \tag{4.97}$$

is ${\tau_{21}}^{\chi}$*-distributed.*

Proof. A careful computation yields

$$\begin{aligned}
\operatorname{tr}(\varphi_3(\mathbf{q},\beta)\varphi_3(\mathbf{p},\alpha)) &= \sum_{i,j=1}^{3} \varphi_3(\mathbf{q},\beta)_{ij}\varphi_3(\mathbf{p},\alpha)_{ji} \\
&= (1-\cos\beta)(1-\cos\alpha)(\mathbf{q}\cdot\mathbf{p})^2 - 2\mathbf{q}\cdot\mathbf{p}\sin\beta\sin\alpha + \cos\beta + \cos\alpha \;\times \\
&\quad\times \quad + \cos\beta\cos\alpha \\
&= \big(2\sin^2(\beta/2)\big)\big(2\sin^2(\alpha/2)\big)(\mathbf{q}\cdot\mathbf{p})^2 - 2\,(2\sin(\beta/2)\cos(\beta/2)) \quad\times \\
&\quad\times \quad (2\sin(\alpha/2)\cos(\alpha/2)) + \cos\beta + \cos\alpha + \cos\beta\cos\alpha \\
&= 4\,(\mathbf{q}\cdot\mathbf{p}\sin(\beta/2)\sin(\alpha/2) - \cos(\beta/2)\cos(\alpha/2))^2 - 1
\end{aligned}$$

where we repeatedly used the identities $1-\cos 2x = 2\sin^2 x$ and $\sin 2x = 2\sin x\cos x$. On the other hand, $\operatorname{tr}(\varphi_3(\mathbf{r},\gamma)) = 2\cos\gamma + 1 = 4\cos^2(\gamma/2) - 1$ which proves (i). Let X_1 and X_2 be independent μ_{S^2}-distributed random vectors. The symmetry of the sphere implies $\operatorname{Prob}(X_2 \cdot X_1 \le z) = \operatorname{Prob}(X_2 \cdot (0,0,1)^t \le z)$. The z-component of X_2 and thus the scalar product $X_2 \cdot X_1$ are equidistributed on $[-1,1]$ ([18], p. 230). As $\cos(\gamma/2) \ge 0$ for $0 \le \gamma \le \pi$ and $\cos(\gamma/2) \le 0$ for $\pi \le \gamma < 2\pi$ it follows from (i) that $|U\sin(Y_2/2)\sin(Y_1/2) - \cos(Y_2/2)\cos(Y_1/2)|$ is ${\tau_{\eta_2\odot\eta_1;\,a}}^{\chi}$-distributed. Further, as $\cos(\gamma/2) = -\cos(\gamma'/2)$ iff $\gamma = 2\pi - \gamma'$ we obtain $\tau_{21}(I \cup 2\pi - I) = \tau_{\eta_2\odot\eta_1;\,a}(I \cup 2\pi - I)$ for each interval $I \subseteq (0,\pi)$ which proves (ii). Finally, as $Y_j/2 \in [0,\pi]$ we have $\sin(Y_j/2) \ge 0$, and (iii) is an immediate corollary of (ii). □

Theorem 4.62 can be used to simulate the distributions $\nu_2 \odot \nu_1$ and, more generally, $\nu_k \odot \cdots \odot \nu_1$, without computing the distribution explicitly. The algorithm below makes this idea precise. The pseudorandom numbers $\widetilde{W_1}, \widetilde{W_2} \ldots, \widetilde{W_k}, \widetilde{U_2}, \widetilde{U_3}, \ldots, \widetilde{U_k}$ are assumed to be independent. Further, $\widetilde{W_j}$ is assumed to be τ_j^{χ}-distributed (with $\tau_j \in \Phi^{-1}(\nu_j)$) and $\widetilde{U_2}, \widetilde{U_3}, \ldots, \widetilde{U_k}$ to be equidistributed on $(-1,1]$. The pseudorandom number S_k may be viewed as ${\tau_{k\cdots 1}}^{\chi}$-distributed with ${\tau_{k\cdots 1}}^{\chi} \in (\Phi^{-1}(\nu_k \odot \cdots \odot \nu_1))^{\chi}$. Step 5.) of the following algorithm uses the trigonometric identities $\cos\gamma = 2\cos^2(\gamma/2) - 1$ and $\sin\gamma = 2\sin(\gamma/2)\cos(\gamma/2)$.

Algorithm (Composition of random rotations)

1.) Generate $\widetilde{W_1}$; $\widetilde{S_1} := \widetilde{W_1}$.
FOR j:=2 TO k DO {
2.) Generate $\widetilde{W_j}$ and $\widetilde{U_j}$.
3.) $\widetilde{S_j} := \widetilde{U_j}\sqrt{1-\widetilde{W_j}^2}\sqrt{1-\widetilde{S_{j-1}}^2} - \widetilde{W_j}\widetilde{S_{j-1}}$.
}
4.) Generate a $\mu_{(S^2)}$-distributed pseudorandom vector $\widetilde{X}$ (pseudorandom axis).
5.) $\widetilde{cos} := 2\widetilde{S_k}^2 - 1$ and $\widetilde{sin} := 2\widetilde{S_k}\sqrt{1-\widetilde{S_k}^2}$. Insert $\widetilde{X}$, $\widetilde{sin}$ and $\widetilde{cos}$ (for (p_1, p_2, p_3), $\cos\theta$ and $\sin\theta$) into the matrix representation (4.72).
6.) END

Remark 4.63. (i) This algorithm is obviously much faster than the naive variant, i.e. simulating rotation by rotation, which requires at least $2k$ evaluations of trigonometric functions, k evaluations of $\varphi_3(\cdot,\cdot)$ and $(k-1)$ matrix-matrix multiplications.
(ii) Replacing Step 5.) by
5′.) $\widetilde{qu} := (\widetilde{S_k}, \sqrt{1-\widetilde{S_k}^2}\cdot\widetilde{X})$,
yields the quaternion representation.
(iii) Replacing $\widetilde{S}_k$ by $|\widetilde{S}_k|$ yields the simulation of the 'asymmetric' distribution $\tau_{\nu_k\odot\cdots\odot\nu_1;a}{}^{\chi}$. Replacing S_k by $-S_k$ with probability 1/2 provides the simulation of the 'symmetric' distribution $\tau_{\nu_k\odot\cdots\odot\nu_1;s}{}^{\chi}$.

If one requires a large number of pseudorandom numbers for particular distributions $\nu_1, \nu_2, \ldots, \nu_k$ it may be economic to compute $\tau_{k\cdots1}{}^{\chi}$ first. Instead of Step 1.) and the loop 2.) – 3.) merely a single $\tau_{k\cdots1}{}^{\chi}$-distributed pseudorandom number has to be generated then. Theorem 4.64 shows that $\tau_{21}{}^{\chi}$ and, by induction, $\tau_{k\cdots1}{}^{\chi}$ have Lebesgue densities in all cases of practical interest.

Theorem 4.64. *For $x, y \in [-1,1]$ let $ug(x,y) := (xy/2) - \sqrt{1-x^2}\sqrt{1-y^2}$ and $og(x,y) := (xy/2) + \sqrt{1-x^2}\sqrt{1-y^2}$. Assume further that ν_j is conjugation-invariant for $j \in \{1,2\}$ while $\tau_j \in \Phi^{-1}(\nu_j)$ either has a Lebesgue density g_j or is a Dirac measure $\neq \varepsilon_0$. Then*

$$\tau_{21}{}^{\chi} = \bar{r}_{(\tau_2,\tau_1)} \cdot \lambda_{(-1,1]} \tag{4.98}$$

where we have to distinguish three cases:
(i) $\bar{r}_{(g_1\cdot\lambda_{[0,2\pi)}, g_2\cdot\lambda_{[0,2\pi)})}(a) = \int_{-1}^{1}\int_{ug(a,b)}^{og(a,b)} \bar{g}_1(b)\bar{g}_2(c)/2\sqrt{1-b^2}\sqrt{1-c^2}\,\mathrm{d}c\,\mathrm{d}b.$
(ii) $\bar{r}_{(g\cdot\lambda_{[0,2\pi)},\varepsilon_\alpha)}(a) = \int_{ug(a,\cos(\alpha/2))}^{og(a,\cos(\alpha/2))} \bar{g}(c)/2\sqrt{1-c^2}\sin(\alpha/2)\,\mathrm{d}c.$
(iii) $\bar{r}_{(\varepsilon_\alpha,\varepsilon_\beta)}(a) = (2\sin(\alpha/2)\sin(\beta/2))^{-1}$ *for*
$a \in (-\cos((\beta-\alpha)/2), -\cos((\beta+\alpha)/2))$ *and* $\bar{r}_{(\varepsilon_\alpha,\varepsilon_\beta)}(a) = 0$ *else.*

Proof. The proof of this theorem requires careful but elementary computations. We refer the interested reader to the proof of Theorem 3.2 in [70]. □

To avoid non-necessary distinctions of cases we excluded the trivial case $\tau_j = \varepsilon_0$. Then $Z_j = 1_3$ with probability one which does not affect the convolution product. Theorem 4.64 supports the computation of $\bar{r}_{(\tau_2,\tau_1)}$ if the $\tau_j \in \mathcal{M}^1([0,2\pi))$ have a Lebesgue density, are discrete or a convex combination of both types. To compute ${\tau_{k\cdots 1}}^{\chi}$ one subsequently computes $\bar{r}_{(\tau_2,\tau_1)}, \bar{r}_{(\tau_3,\tau_{21})}, \cdots, \bar{r}_{(\tau_k,\tau_{k-1\cdots 1})}$. (Note that $\bar{r}_{(\tau_{j+1},\tau_{j\cdots 1})}$ depends on ${\tau_{j+1}}^{\chi}$ and ${\tau_{j\cdots 1}}^{\chi}$ which is known at the time of computation.) If this is needed we may obtain $\tau_{k\cdots 1}$ from ${\tau_{k\cdots 1}}^{\chi}$ by applying the transformation theorem to the inverse transformation $\chi^{-1}\colon (-1,1] \to [0,2\pi)$. In particular,

$$\big(\bar{g}\cdot\lambda_{(-1,1]}\big)^{\chi^{-1}} = g\cdot\lambda_{[0,2\pi)} \qquad \text{with } g(\alpha) = \bar{g}(\cos(\alpha/2))\sin(\alpha/2)/2. \tag{4.99}$$

Finally, we consider statistical applications. The sufficient statistics ρ_1^m and tr^m, resp., transfer test problems from $\mathrm{SO}(3)^m$ to compact intervals. This facilitates the necessary computations considerably. Moreover, for test problems on subsets of $\mathbb{R}^m$ there exist many well-known tests. We will not pursue this aspect as this would lead too far beyond from the scope of this book. Instead, for readers who are not familiar with statistics we just discuss two examples to illustrate the procedure.

Example 4.65. (i) Let $\Gamma = \{0,1\}$, $\Gamma_0 = \{0\}$, $p_0 = (\mu_{SO(3)})^m$ and $p_1 = (2(3-\mathrm{tr})^{-1}\cdot\mu_{\mathrm{SO}(3)})^m$ (cf. 4.59(i), (iv)). We want to determine a most powerful test $\Psi\colon \mathrm{SO}(3)^m \to [0,1]$ for the significance level $\alpha \in (0,1)$. That is, we have to determine a test Ψ with

$$\mathrm{p}ow_\Psi(p_0) = \int_{\mathrm{SO}(3)^m} \Psi(\mathbf{m})\, p_0(\mathrm{d}\mathbf{m}) \le \alpha \qquad \text{and} \tag{4.100}$$

$$\mathrm{p}ow_\Psi(p_1) = \sup\{\mathrm{p}ow_{\Psi'}(p_1) \mid \Psi'\colon \mathrm{SO}(3)^m \to [0,1],\ \text{measurable},\ \mathrm{p}ow_{\Psi'}(p_0) \le \alpha\}. \tag{4.101}$$

For simple null and alternative hypotheses the Neyman-Pearson Lemma ([51], pp. 71) guarantees the existence of a most powerful test. The sufficient statistic $\mathrm{tr}^m\colon \mathrm{SO}(3)^m \to [-1,3]^m$ transfers the computation of a most powerful test to the transformed test problem $\{p_0^{\mathrm{tr}^m}, p_1^{\mathrm{tr}^m}\}$. As repeatedly mentioned before this simplifies the necessary computations considerably. From 4.56(v) we obtain

$$p_0^{\mathrm{tr}^m} = \big(h_{0;3}\cdot\lambda_{[-1,3]}\big)^m \qquad \text{and} \tag{4.102}$$

$$p_1^{\mathrm{tr}^m} = \big(f_1\cdot\lambda_{[-1,3]}\big)^m \qquad \text{with } f_1(x) = \frac{\sqrt{3-x}}{\pi(3-x)\sqrt{1+x}}. \tag{4.103}$$

For each $c \in \mathbb{R}$ the set

$$\left\{ \mathbf{x} \in [-1,3]^m \mid \prod_{j=1}^{m} \frac{f_1(x_j)}{h_{0;3}(x_j)} = c \right\} \tag{4.104}$$

is a ${\lambda_{[-1,3]}}^m$-zero set and hence also a ${p_0^{\mathrm{tr}}}^m$-zero set. Consequently, there exists a deterministic most powerful test Ψ_1 ([51], p. 74). That is, there exists a measurable subset $A \subseteq [-1,3]^m$ with $\Psi_{1|A} = 1$ and $\Psi_{1|A^c} = 0$. More precisely: The most powerful test Ψ_1 for the transformed test problem $\{{p_0^{\mathrm{tr}}}^m, {p_1^{\mathrm{tr}}}^m\}$ to significance level α has the form

$$\Psi_1(x_1,\ldots,x_m) := \begin{cases} 1 & \text{if } \prod_{j=1}^m \frac{f_1(x_j)}{h_{0;3}(x_j)} \geq r_\alpha \\ 0 & \text{else.} \end{cases} \tag{4.105}$$

The constant $r_\alpha \in \mathbb{R}$ is implicitly given by $\int_{[-1,3]^m} 1_{\{f_1^m/h_{0;3}^m \geq r_\alpha\}} h_{0;3}^m \,\mathrm{d}\lambda^m = \alpha$. Hence

$$\Psi(\boldsymbol{S}_1,\ldots,\boldsymbol{S}_m) := \begin{cases} 1 & \text{if } \prod_{j=1}^m \frac{2}{3-\mathrm{tr}(\boldsymbol{S}_j)} \geq r_\alpha \\ 0 & \text{else.} \end{cases} \tag{4.106}$$

defines a most powerful test for the test problem $\{p_0, p_1\}$.

An exact computation of the threshold value r_α seems hardly be possible unless m is very small. Instead, numerical methods or simulation techniques can be applied. We describe the latter. As the simulation of the densities $h_{0;3} \cdot \lambda_{[-1,3]}$ and $f_1 \cdot \lambda_{[-1,3]}$ is costly we apply the transformation $g\colon [0,\pi] \to [-1,3]$ from 4.56(v). By this, we transfer the simulation part from $[-1,3]^m$ to $[0,\pi]^m$. The transformation theorem yields

$$\begin{aligned} \alpha &= \int_{[-1,3]^m} 1_{\{f_1^m(\mathbf{x})/h_{0;3}^m(\mathbf{x}) \geq r_\alpha\}}(\mathbf{x}) h_{0;3}^m(\mathbf{x}) \,\mathrm{d}\mathbf{x} \qquad (4.107) \\ &= \int_{[0,\pi]^m} 1_{\{(f_1\circ g)^m(\mathbf{y})/(h_{0;3}\circ g)^m(\mathbf{y}) \geq r_\alpha\}}(\mathbf{y}) (h_{0;3}\circ g)^m(\mathbf{y}) |\det D(g^m)|(\mathbf{y}) \,\mathrm{d}\mathbf{y} \\ &= \pi^{-m} \int_{[0,\pi]^m} 1_{\left\{\prod_{j=1}^m (1-\cos y_j) \leq r_\alpha^{-1}\right\}}(y_1,\ldots,y_m) \prod_{j=1}^m (1-\cos y_j)\, dy_1 \cdots \mathrm{d}y_m. \end{aligned}$$

Stochastic simulations provide an empirical cumulative distribution function of the random variable $\prod_{j=1}^m (1-\cos Y_j)$ where $Y_1,\ldots,Y_m$ are iid $(2 \cdot 1_{[0,\pi]} h_3 \cdot \lambda_{[0,\pi]})$-distributed. This gives an estimator for the threshold value r_α. Analogously, one estimates $pow_{\Psi_1}({p_1}^{\mathrm{tr}^m}) = pow_\Psi(p_1)$. For sample size $m = 5$ practical experiments yielded Table 4.2.

Applying this test to the samples $(\boldsymbol{S}_1,\ldots,\boldsymbol{S}_5),\ldots,(\boldsymbol{S}_{96},\ldots,\boldsymbol{S}_{100})$ gives twenty test values $\Psi(\boldsymbol{S}_1,\ldots,\boldsymbol{S}_5),\ldots,\Psi(\boldsymbol{S}_{96},\ldots,\boldsymbol{S}_{100})$. If the alternative hypothesis is true one may expect about 10 or 11 rejections of the null hypothesis (i.e. the respective test value $\Psi(\cdots)$ equals 1) at the significance level 0.01. On the other hand if the null hypothesis is true the probability for at least four rejections is about 6×10^{-5}. Note that we used the sufficient statistic tr^m to determine the most powerful test $\Psi(\cdot,\ldots,\cdot)$ besides the threshold value

Table 4.2. Threshold values and the power function for $m = 5$

α	0.05	0.02	0.01
r_α	2.17	5.26	69.7
$pow_\Psi(p_1)$	0.72	0.62	0.54

r_α. To determine the threshold value r_α, however, we applied the mapping ρ_1^m.

If the hypotheses are composite tests which are simultaneously most powerful for each pair $(\gamma_0, \gamma_1) \in \Gamma_0 \times (\Gamma \setminus \Gamma_0)$ do only exist under specific conditions ([51], pp. 78 ff.). A generalization of this concept are statistical tests which are based on minimax principles ([51], pp. 505 ff.). However, for many examples which are relevant for applications the necessary computations are very costly. It may happen that they are not practically feasible at all.

In the final example of this section we treat composite hypotheses. We derive a test which is powerful although it does not meet minimax principles. In fact the necessary computations are simple and the pre-considerations are not complicated. This test exploits the fact that different admissible hypotheses have different mean values.

Example 4.66. Let $\Gamma = [0, 1]$, $\Gamma_0 = [0, 0.2]$ and $p_\gamma = (f_\gamma \cdot \mu_{SO(3)})^m$ with $f_\gamma := (1-\gamma)\cdot 1 + \gamma \cdot \mathrm{tr}^2$. In other words we interpret the sample $\boldsymbol{S}_1, \ldots, \boldsymbol{S}_m \in \mathrm{SO}(3)$ as a realization of iid $(f_\gamma \cdot \mu_{SO(3)})$-distributed random variables $Z_1, \ldots, Z_m$ where $\gamma \in \Gamma$ is unknown. The null hypothesis claims that the unknown parameter γ is contained in Γ_0. At first we determine the expectation $\mathrm{E}(\mathrm{tr}(Z_j))$ and the variance $\mathrm{Var}(\mathrm{tr}(Z_j))$ as a function of γ. This requires some preparatory work.

Suppose that the random variables Z and Z' are $\mu_{SO(3)}$-distributed and $\mathrm{tr}^2 \cdot \mu_{SO(3)}$-distributed, resp. From 4.55(iii) we conclude

$$\mathrm{E}(\mathrm{tr}(Z')^k) = \int_{-1}^{3} \frac{x^k x^2 \sqrt{3-x}}{2\pi\sqrt{1+x}}\,\mathrm{d}x = \mathrm{E}\left(\mathrm{tr}(Z)^{k+2}\right). \tag{4.108}$$

The substitution $x = 2\cos z + 1$ yields

$$\begin{aligned}
\mathrm{E}\left(\mathrm{tr}(Z)^k\right) &= \int_{-1}^{3} x^k \frac{\sqrt{3-x}}{2\pi\sqrt{1+x}}\,\mathrm{d}x && (4.109)\\
&= \int_0^\pi (2\cos z + 1)^k \frac{\sqrt{2-2\cos z}}{\sqrt{2+2\cos z}} \frac{\sqrt{1-\cos z}}{\sqrt{1-\cos z}} \frac{2\sin z}{2\pi}\,\mathrm{d}z \\
&= \frac{1}{\pi}\int_0^\pi (2\cos z + 1)^k (1-\cos z)\,\mathrm{d}z \\
&= \sum_{j=0}^{k} \binom{k}{j} 2^j \frac{1}{\pi} \int_0^\pi \left((\cos z)^j - (\cos z)^{j+1}\right)\,\mathrm{d}z. && (4.110)
\end{aligned}$$

Of course, only the terms with even exponents give non-zero contributions. Elementary computations yield $\mathrm{E}(\mathrm{tr}(Z)) = 0$, $\mathrm{E}(\mathrm{tr}(Z)^2) = \mathrm{E}(\mathrm{tr}(Z)^3) = 1$ and $\mathrm{E}(\mathrm{tr}(Z)^4) = 3$. From this, we obtain $\mathrm{E}(\mathrm{tr}(Z_j)) = \gamma$ and $\mathrm{E}(\mathrm{tr}(Z_1)^2) = 1 + 2\gamma$, i.e. $\mathrm{Var}(\mathrm{tr}(Z_j)) = 1 + 2\gamma - \gamma^2$. In particular, mean and variance increase monotonously in γ. For $\gamma = 0.2$ the term $(\sum_{j=1}^m \mathrm{tr}(Z_j) - 0.2m)/\sqrt{m}\sqrt{1.36}$ is asymptotically normal distributed (Central Limit Theorem). Moreover,

$$\begin{aligned} & P\left(\frac{\sum_{j=1}^m \mathrm{tr}\,(Z_j) - 0.2m}{\sqrt{m}\sqrt{1.36}} \geq x\right) \qquad (4.111) \\ & = P\left(\frac{\sum_{j=1}^m \mathrm{tr}\,(Z_j) - m\gamma}{\sqrt{m}\sqrt{1+2\gamma-\gamma^2}} \geq \sqrt{\frac{1.36}{1+2\gamma-\gamma^2}}x - \sqrt{m}\frac{\gamma - 0.2}{\sqrt{1+2\gamma-\gamma^2}}\right) \\ & = 1 - F_{N(0,1)}\left(\sqrt{\frac{1.36}{1+2\gamma-\gamma^2}}x + \sqrt{m}\frac{0.2-\gamma}{\sqrt{1+2\gamma-\gamma^2}}\right) \end{aligned}$$

where $F_{N(0,1)}$ denotes the cumulative distribution function of the standard normal distribution. The last term is strictly increasing in γ. For significance level α this suggests the following test $\Psi\colon \mathrm{SO}(3)^m \to [0,1]$,

$$\Psi\left(\boldsymbol{S}_1, \ldots, \boldsymbol{S}_m\right) = \begin{cases} 1 & \text{if } \left(\sum_{j=1}^m \mathrm{tr}(\boldsymbol{S}_j) - 0.2m\right)/\sqrt{1.36m} > x_\alpha \\ 0 & \text{else.} \end{cases} \qquad (4.112)$$

Clearly, for $\gamma < 0.2$ (resp. $\gamma = 0.2$, resp. $\gamma > 0.2$) the power function $pow_\Psi(p_\gamma)$ is $< \alpha$ (resp. $= \alpha$. resp. $> \alpha$). For $m = 15$ and $\alpha = 0.05$ it is $x_\alpha = 1.65$, and we obtain

Table 4.3. Power function for $m = 15$, $\alpha = 0.05$

γ	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
$pow_\Psi(p_\gamma)$	.004	.014	.050	.107	.186	.284	.393	.505	.614	.713	.798.

4.6 The Theorem of Iwasawa and Invariant Measures on Lie Groups

Similarly as in Section 4.3 we consider a general situation. In Sections 4.7 and 4.8 below we will investigate two well-known special cases in detail, namely the QR- and the polar decomposition on $\mathrm{GL}(n)$. Within this section we assume that H is a finite dimensional Lie group with finitely many components and that $G \leq H$ is a maximal compact subgroup of H. We point out that H in particular is a manifold and hence a second countable locally compact

space. Consequently, there exists a left invariant Haar measure $\mu_{H;l}$ on H (cf. Remark 4.15). The compact group G acts on H and G by left multiplication. The G-action onto itself is transitive, and the Haar measure μ_G is the unique G-invariant probability measure on G. We point out that $\mu_{H;l} \in \mathcal{M}(H)$.

Lemma 4.67. (i) *(Theorem of Iwasawa) Let H be a Lie group with finitely many components. Then there exist maximal compact subgroups, and any two maximal subgroups are conjugate. For any maximal compact subgroup $G \leq H$ there is an integer $k \geq 0$ and a submanifold $V \subseteq H$ which is diffeomorphic to $\mathbb{R}^k$ so that*

$$\varphi\colon G \times V \to H, \qquad \varphi(g,v) := gv \tag{4.113}$$

is a diffeomorphism.
(ii) *Let $\varphi\colon G \times V \to H$ be given by $\varphi(g,v) := gv$. With the G-actions defined above the 5-tuple (G, G, V, H, φ) has Property $(*)$.*

Proof. Statement (i) combines Theorems A, B and C from [39], p. 623. (For a stronger version of 4.67(i) the interested reader is referred to [38], pp. 180ff, while the pioneering work [41] provides a weaker version.) As φ is a diffeomorphism it remains to verify the equivariance of φ. For $g, g', v \in G \times G \times V$ we have $\varphi(g'g, v) = g'gv = g'\varphi(g,v)$ which completes the proof of (ii). □

Theorem 4.68. *Let $\Phi\colon \mathcal{M}^\sigma(V) \to \mathcal{M}_G^+(H)$, $\Phi(\tau) := (\mu_G \otimes \tau)^\varphi$. Then the following statements are valid.*
(i) *$\Phi(\mathcal{M}(V)) = \mathcal{M}_G(H)$ and $\Phi^{-1}(\mathcal{M}_G(H)) = \mathcal{M}(V)$. In particular, Φ is bijective onto its image.*
(ii) *Let $\nu \in \mathcal{M}_G(H)$ and $\eta = f \cdot \nu \in \mathcal{M}(H)$ with G-invariant density f. Then $\eta \in \mathcal{M}_G(H)$, and $\Phi^{-1}(\nu) = \tau_\nu$ implies $\Phi^{-1}(\eta) = f_{|V} \cdot \tau_\nu$.*
(iii) *Let $\mathrm{pr}_2\colon G \times V \to V \subseteq H$ denote the projection onto the second component. The mapping $(\mathrm{pr}_2 \circ \varphi^{-1})^m$ is a sufficient statistic for all test problems $(\mathrm{P}_{\Gamma_0}, \mathrm{P}_{\Gamma \setminus \Gamma_0})$ with $\mathrm{P}_\Gamma \subseteq \mathcal{M}_G^1(H)^m$.*

Proof. As φ is a diffeomorphism the inverse mapping φ^{-1} is continuous and $(\varphi \circ \iota)^{-1}(K) = \mathrm{pr}_2(\varphi(K)) \subseteq V$ is compact for each G-invariant compact subset $K \subseteq H$. Hence the first assertion of (i) follows from 4.9(i). As φ is a homeomorphism $\Phi(\tau_1) = \Phi(\tau_2)$ i.e. $(\mu_G \otimes \tau_1) = (\mu_G \otimes \tau_2)$, implies $\tau_1 = \tau_2$ which completes the proof of (i). Assertion (ii) is an immediate consequence of Theorem 4.7(vii). As φ is bijective the equivalence classes on V are singleton, that is, V itself is a section. □

Remark 4.69. (i) As will become clear in the Sections 4.7 and 4.8 the bijectivity of φ simplifies computations considerably. We mention that Theorem 4.68 can also be proved with elementary methods (cf. the proof of Theorem 2.4 (ii) in [65] which equals statement (ii)). Also techniques from [81] could be applied. The submanifold V is a global cross section (cf. Remark 2.46). Note, however, that the proof of 4.68 neither exploits that φ and φ^{-1} are differentiable nor the maximality of G nor that the submanifold $V \subseteq H$ is

diffeomorphic to $\mathbb{R}^k$.
(ii) If H is compact then $G = H$ and $k = 0$, i.e. $H \cong H \times \{e_H\}$.
(iii) Let $\nu \in \mathcal{M}_G(H)$ and $E \in \mathcal{B}(V)$ with $0 < \nu(GE) < \infty$. Then

$$\mu_G(B) = \frac{\nu(BE)}{\nu(GE)} \qquad \text{for each } B \in \mathcal{B}(G). \tag{4.114}$$

In fact, $\eta(B) := \nu(BE)/\nu(GE)$ defines a probability measure on G. As $\nu \in \mathcal{M}_G(H)$ the nominator is invariant under the G-action and hence $\eta = \mu_G$.

4.7 QR-Decomposition on GL(n)

In the Sections 4.7, 4.8 and 4.10 we investigate symmetries on GL(n). This and the following section consider special cases of the more general situation treated in Section 4.6. We begin with some remarks on GL(n).

Clearly, $\boldsymbol{M} := (m_{ij})_{1 \le i,j \le n} \mapsto (m_{11}, \ldots, m_{nn})$ defines an vector space isomorphism between $\mathrm{Mat}(n,n)$ and $\mathbb{R}^{n^2}$. The pre-image of the Lebesgue measure λ_{n^2} on $\mathbb{R}^{n^2}$ is suggestively denoted with $\lambda_{\mathrm{Mat}(n,n)}$. Clearly,

$$\lambda_{\mathrm{Mat}(n,n)}\left(\{\boldsymbol{M} \in \mathrm{Mat}(n,n) \mid a_{ij} \le m_{ij} \le b_{ij}\}\right) = \prod_{i,j} (b_{ij} - a_{ij}) \tag{4.115}$$

for all $(a_{11}, \ldots, a_{nn}), (b_{11}, \ldots, b_{nn})$ with $a_{ij} \le b_{ij}$. In particular, $\lambda_{\mathrm{Mat}(n,n)}$ is invariant under the addition of any fixed matrix. We call $\lambda_{\mathrm{Mat}(n,n)}$ the *Lebesgue measure* on $\mathrm{Mat}(n,n)$. The general linear group GL(n) is an open subset of $\mathrm{Mat}(n,n)$. Its complement is the zero set of the determinant function on $\mathrm{Mat}(n,n)$ and hence is a $\lambda_{\mathrm{Mat}(n,n)}$-zero set (4.21(vi)). We denote the restriction of $\lambda_{\mathrm{Mat}(n,n)}$ to GL(n) with $\lambda_{\mathrm{GL}(n)}$. The multiplicative group GL(n) is locally compact and hence there exists a left invariant Haar measure $\mu_{\mathrm{GL}(n)}$. Moreover, the GL(n) is unimodular, i.e. $\mu_{\mathrm{GL}(n)}$ is also right invariant. Consequently, we call $\mu_{\mathrm{GL}(n)}$ briefly a Haar measure (cf. Remark 4.15). In particular,

$$\mu_{\mathrm{GL}(n)} = h_{\mu;n} \cdot \lambda_{\mathrm{GL}(n)} \qquad \text{with} \qquad h_{\mu;n}(\boldsymbol{M}) = |\det \boldsymbol{M}|^{-n} \tag{4.116}$$

([54], p. 70). Obviously,

$$\mu_{\mathrm{GL}(n)} \ll \lambda_{\mathrm{GL}(n)} \qquad \text{and} \qquad \lambda_{\mathrm{GL}(n)} \ll \mu_{\mathrm{GL}(n)}. \tag{4.117}$$

Finally, we point out that GL(n) is a Lie group with Lie algebra $\mathrm{Mat}(n,n)$. The GL(n) falls into two components which are given by the matrices with positive or negative determinant, resp. The following definitions will be used in the present and the following sections.

Definition 4.70. *Let* $\mathrm{R}(n)$ *denote the group of all upper triangular matrices with positive diagonal elements and* $\mathrm{Pos}(n)$ *the set of all real symmetric positive definite matrices. For* $\boldsymbol{M} \in \mathrm{Mat}(n,n)$ *the* Frobenius norm *of* $\boldsymbol{M}$ *is given by* $\|\boldsymbol{M}\|_{\mathrm{F}} := \sqrt{\sum_{i,j=1}^{n} m_{ij}^2}$. *For any subset* $W \subseteq \mathrm{Mat}(n,n)$ *we define* $K_W(r) := \{\boldsymbol{M} \in W \mid 0 \leq \|\boldsymbol{M}\|_{\mathrm{F}} \leq r\}$. *Different from Sections 4.4 and 4.5 the terms* $\rho_1(\mathbf{M}), \ldots, \rho_n(\mathbf{M})$ *do not denote the phases of the eigenvalues of* $\boldsymbol{M}$ *but the eigenvalues themselves. If* $\boldsymbol{M}$ *is symmetric we always assume* $\rho_1(\mathbf{M}) \geq \rho_2(\mathbf{M}) \geq \cdots \geq \rho_n(\mathbf{M})$. *For* $\boldsymbol{P} \in \mathrm{Pos}(n)$ *the unique* $\boldsymbol{P}_1 \in \mathrm{Pos}(n)$ *with* $\boldsymbol{P}_1^2 = \boldsymbol{P}$ *is denoted with* $\sqrt{\boldsymbol{P}}$.

Any $H' \in \{\mathrm{Mat}(n,n), \mathrm{GL}(n), \mathrm{R}(n), \mathrm{Pos}(n)\}$ can be identified with $\mathbb{R}^{n^2}$ or an open subset of $\mathbb{R}^{n^2}$ or $\mathbb{R}^{n(n+1)/2}$, resp., via $\boldsymbol{M} \mapsto (m_{11}, m_{12}, \ldots, m_{1n}, m_{21}, \ldots, m_{nn})$ or $\boldsymbol{M} \mapsto (m_{11}, m_{12}, \ldots, m_{1n}, m_{22}, m_{23}, \ldots, m_{nn})$, resp. In particular, we denote the pre-image of λ_{n^2} or $\lambda_{n(n+1)/2}$, resp., as Lebesgue measure on H' and use the term $\lambda_{H'}$. If it unambiguous we write briefly dh' instead of $\lambda_{H'}(dh')$.

Each regular $(n \times n)$-matrix $\boldsymbol{M}$ can be represented as a product $\boldsymbol{M} = \boldsymbol{T}_{\mathbf{M}} \boldsymbol{R}_{\mathbf{M}}$ with $\boldsymbol{T}_{\mathbf{M}} \in \mathrm{O}(n)$ and $\boldsymbol{R}_{\mathbf{M}} \in \mathrm{R}(n)$. Both factors are unique (*QR-decomposition*). We point out that the orthogonal group $\mathrm{O}(n)$ acts on $\mathrm{GL}(n)$ and $\mathrm{O}(n)$ by left multiplication. By definition, $\mathcal{M}_{\mathrm{O}(n)}(\mathrm{GL}(n))$ denotes the set of all Borel measures on $\mathrm{GL}(n)$ which are invariant under left multiplication with orthogonal matrices.

Lemma 4.71. *(i) The orthogonal group* $\mathrm{O}(n)$ *is a maximal compact subgroup of* $\mathrm{GL}(n)$.
(ii) $\mathrm{R}(n)$ *is a locally compact group. A right-invariant Haar measure on* $\mathrm{R}(n)$ *is given by*

$$\mu_{\mathrm{R}(n)} = h_{\mu;\mathrm{R}(n)} \cdot \lambda_{\mathrm{R}(n)} \qquad \textit{with} \qquad h_{\mu;\mathrm{R}(n)}(\mathbf{R}) := c_{\mathrm{Q};n} \prod_{j=1}^{n} r_{jj}^{-j}. \tag{4.118}$$

(iii) The mapping

$$\varphi_Q \colon \mathrm{O}(n) \times \mathrm{R}(n) \to \mathrm{GL}(n), \qquad \varphi_Q(\boldsymbol{T}, \boldsymbol{R}) := \boldsymbol{T}\boldsymbol{R} \tag{4.119}$$

is a diffeomorphism.
(iv) The orthogonal group $\mathrm{O}(n)$ *acts on* $\mathrm{GL}(n)$ *and* $\mathrm{O}(n)$ *by left multiplication. The* 5*-tuple* $(\mathrm{O}(n), \mathrm{O}(n), \mathrm{R}(n), \mathrm{GL}(n), \varphi_Q)$ *has Property* $(*)$.

Proof. Assume $\mathrm{O}(n) \leq G' \leq \mathrm{GL}(n)$ with $\mathrm{O}(n) \neq G'$. Let $\boldsymbol{F} \in G' \setminus \mathrm{O}(n)$. There exist $\boldsymbol{T}_1, \boldsymbol{T}_2 \in \mathrm{O}(n)$ such that $\boldsymbol{T}_1\boldsymbol{F}$ is a positive definite symmetric matrix (polar decomposition) and $\boldsymbol{D} := \boldsymbol{T}_2^t \boldsymbol{T}_1 \boldsymbol{F} \boldsymbol{T}_2 \in G'$ is diagonal with positive entries. If $\boldsymbol{D} \neq 1_n$ then $\{\boldsymbol{D}^z \mid z \in \mathbb{Z}\}$ is not bounded and hence G' is not compact. If $\boldsymbol{D} = 1_n$ then $\boldsymbol{F} \in \mathrm{O}(n)$ which leads to a contradiction. The subgroup $\mathrm{R}(n) \leq \mathrm{GL}(n)$ is the intersection of an

open and closed subset of GL(n) and by 2.4(ii) locally compact in the induced topology. In particular, there is a unique right-invariant Haar measure $\mu_{\mathrm{R}(n)}$ on R(n) (cf. 4.15). Formula (4.118) is proved in [11], Chap. 7, par. 3, no. 3, Exemple 7. The uniqueness of the QR-decomposition implies the bijectivity of φ_Q. Clearly, φ_Q is smooth. By 4.16 it remains to show that $\det D\varphi_Q(\boldsymbol{T}_0, \boldsymbol{P}_0) \neq 0$ for each pair $(\boldsymbol{T}_0, \boldsymbol{R}_0) \in \mathrm{O}(n) \times \mathrm{R}(n)$. (The existence of a submanifold $V \subseteq \mathrm{GL}(n)$ with this property follows from 4.67(i).) Let temporarily $L_1 \colon \mathrm{O}(n) \times \mathrm{R}(n) \to \mathrm{O}(n) \times \mathrm{R}(n)$ and $L_2 \colon \mathrm{GL}(n) \to \mathrm{GL}(n)$ be defined by $L_1(\boldsymbol{T}, \boldsymbol{R}) := (\boldsymbol{T}_0^{-1}\boldsymbol{T}, \boldsymbol{R})$ and $L_2(\boldsymbol{M}) := \boldsymbol{T}_0\boldsymbol{M}$. If $\|\boldsymbol{T} - \boldsymbol{T}_0\|_F$ is sufficiently small then

$$\varphi_Q(\boldsymbol{T}, \boldsymbol{R}) = L_2 \circ \varphi_Q \circ \Big(\exp_{so(n)} \times \mathrm{id}\Big) \circ \Big(\exp_{so(n)}^{-1} \times \mathrm{id}\Big) \circ L_1(\boldsymbol{T}, \boldsymbol{R}). \quad (4.120)$$

As L_1 and L_2 are bijective linear mappings and since the exponential map $\exp_{so(n)}$ is diffeomorphic in a neighbourhood of $0 \in \mathrm{so}(n)$ (cf. Remark 4.17) the chain rule implies that $\det D\varphi_Q(\boldsymbol{T}_0, \boldsymbol{R}_0) \neq 0$ iff $\det Df(0, \boldsymbol{R}_0) \neq 0$ where $f := \varphi_Q \circ \Big(\exp_{so(n)} \times \mathrm{id}\Big)$ for the moment. A careful computation yields

$$\begin{aligned}
&\varphi_Q(\exp_{so(n)}(\Delta\boldsymbol{F}), \boldsymbol{R}_0 + \Delta\boldsymbol{R}) - \varphi_Q(\exp_{so(n)}(0), \boldsymbol{R}_0) \\
&\qquad = (1_n + \Delta\boldsymbol{F} + O(\|\Delta\boldsymbol{F}\|^2))(\boldsymbol{R}_0 + \Delta\boldsymbol{R}) - 1_n\boldsymbol{R}_0 \\
&\qquad = \Delta\boldsymbol{F}\boldsymbol{R}_0 + \Delta\boldsymbol{R} + O(\|\Delta\boldsymbol{F}\|^2) + O(\|\Delta\boldsymbol{F}\|\,\|\Delta\boldsymbol{R}\|).
\end{aligned}$$

In other words, $Df(0, \boldsymbol{R})(\Delta\boldsymbol{F}, \Delta\boldsymbol{R}) = \Delta\boldsymbol{F}\boldsymbol{R}_0 + \Delta\boldsymbol{R}$. Clearly, $\Delta\boldsymbol{F}\boldsymbol{R}_0 + \Delta\boldsymbol{R} = 0$ iff $\Delta\boldsymbol{F} = -\Delta\boldsymbol{R}\boldsymbol{R}_0^{-1}$ which implies $\Delta\boldsymbol{F} = 0$ and hence also $\Delta\boldsymbol{R} = 0$ as $\boldsymbol{R}_0$ and $\Delta\boldsymbol{R}$ are upper triangular and $\Delta\boldsymbol{F}$ is skew-symmetric. This means that the differential $Df(0, \boldsymbol{R})$ is bijective which finishes the proof of (iii). Assertion (iv) follows from (iii) and Lemma 4.67(ii). □

The QR-decomposition represents a special case of Iwasawa's Theorem. (Note that $\boldsymbol{R} \mapsto (\log r_{11}, \ldots, \log r_{nn}, r_{12}, \ldots, r_{n-1,n})$ induces a diffeomorphism between R(n) and $\mathbb{R}^{n(n+1)/2}$.) Keep in mind, however, that Theorem 4.72 below exploits only the fact that φ_Q is a homeomorphism (cf. 4.69(i)).

Theorem 4.72. *Let* $\Phi \colon \mathcal{M}^\sigma(\mathrm{R}(n)) \to \mathcal{M}^+_{\mathrm{O}(n)}(\mathrm{GL}(n))$, $\Phi(\tau) := \big(\mu_{\mathrm{O}(n)} \otimes \tau\big)^{\varphi_Q}$. *Then the following statements are valid.*
(i) $\Phi(\mathcal{M}(\mathrm{R}(n))) = \mathcal{M}_{\mathrm{O}(n)}(\mathrm{GL}(n))$ *and* $\Phi^{-1}(\mathcal{M}_{\mathrm{O}(n)}(\mathrm{GL}(n))) = \mathcal{M}(\mathrm{R}(n))$. *In particular,* Φ *is bijective onto its image.*
(ii) *Let* $\nu \in \mathcal{M}_{\mathrm{O}(n)}(\mathrm{GL}(n))$ *and* $\eta = f \cdot \nu \in \mathcal{M}(\mathrm{R}(n))$ *where* f *is invariant under the left multiplication with orthogonal matrices. Then* $\eta \in \mathcal{M}_{\mathrm{O}(n)}(\mathrm{GL}(n))$, *and* $\Phi^{-1}(\nu) = \tau_\nu$ *implies* $\Phi^{-1}(\eta) = f_{|\mathrm{R}(n)} \cdot \tau_\nu$.
(iii) *In particular*

$$\lambda_{\mathrm{GL}(n)} = \big(\mu_{\mathrm{O}(n)} \otimes \mathrm{h}_{\mathrm{R}(n)} \cdot \lambda_{\mathrm{R}(n)}\big)^{\varphi_Q} \quad \textit{with } \mathrm{h}_{\mathrm{R}(n)}(\mathbf{R}) := c_{\mathrm{Q};n} \prod_{j=1}^{n} r_{jj}^{n-j} \quad (4.121)$$

for a suitable constant $c_{\mathrm{Q};n}$. *More precisely,*

$$c_{\mathrm{Q};n} = \frac{2^n \sqrt{\pi}^{\,n(n+1)/2}}{\prod_{j=1}^{n} \Gamma(\frac{j}{2})}. \tag{4.122}$$

That is,

$$c_{\mathrm{Q};2k} = \frac{2^{k^2+k}\pi^{k^2}}{\prod_{j=1}^{k-1}(2j)!} \qquad \textit{and} \qquad c_{\mathrm{Q};2k+1} = \frac{2^{k^2+2k+1}\pi^{k^2+k}}{\prod_{j=1}^{k-1}(2j+1)!}. \tag{4.123}$$

(iv) *Let* $\mathrm{pr}_2 \colon \mathrm{O}(n) \times \mathrm{R}(n) \to \mathrm{R}(n)$ *denote a projection onto the second component. The mapping* $(\mathrm{pr}_2 \circ \varphi_Q^{-1})^m$ *is a sufficient statistic for all test problems* $(\mathrm{P}_{\Gamma_0}, \mathrm{P}_{\Gamma \setminus \Gamma_0})$ *with* $\mathrm{P}_\Gamma \subseteq \mathcal{M}^1_{\mathrm{O}(n)}(\mathrm{GL}(n))^m$.

Proof. The statements (i), (ii) and (iv) follow from Lemma 4.71 and Theorem 4.68. It remains to prove (iii). We first note that $\Phi^{-1}(\mu_{GL(n)}) = \mu_{GL(n)}{}^{\mathrm{pr}_2 \circ \varphi_Q^{-1}}$ is a right invariant Haar measure on the group $\mathrm{R}(n)$. In fact, for $(\boldsymbol{R}, B) \in \mathrm{R}(n) \times \mathcal{B}(\mathrm{R}(n))$ we conclude

$$\begin{aligned}
&\mu_{GL(n)}{}^{\mathrm{pr}_2 \circ \varphi_Q^{-1}}(B\boldsymbol{R}) = \left(\mu_{\mathrm{O}(n)} \otimes \mu_{GL(n)}{}^{\mathrm{pr}_2 \circ \varphi_Q^{-1}}\right)(\mathrm{O}(n) \times B\boldsymbol{R}) \\
&= \mu_{GL(n)}\left(\mathrm{O}(n)B\boldsymbol{R}\right) = \mu_{GL(n)}\left(\mathrm{O}(n)B\right) = \mu_{GL(n)}{}^{\mathrm{pr}_2 \circ \varphi_Q^{-1}}(B).
\end{aligned}$$

Equation (4.121) finally follows from (4.118), (4.116) and assertion (ii). The evaluation of a reference integral yields the constant $c_{\mathrm{Q};n}$. We first note that $\|\boldsymbol{M}\|_F = \|\boldsymbol{T}\boldsymbol{M}\|_F$ for all $(\boldsymbol{T}, \boldsymbol{M}) \in \mathrm{O}(n) \times \mathrm{GL}(n)$ as the left multiplication with an orthogonal matrix does not change the lengths of the column vectors. In particular, $\varphi_Q\left(\mathrm{O}(n) \times K_{\mathrm{R}(n)}(1)\right) = K_{\mathrm{GL}(n)}(1)$ and consequently

$$\int_{\mathrm{GL}(n)} 1_{K_{\mathrm{GL}(n)}}(\boldsymbol{M})\, \lambda_{\mathrm{GL}(n)}(\mathrm{d}\boldsymbol{M}) = \int_{\mathrm{R}(n)} 1_{K_{\mathrm{R}(n)}}(\boldsymbol{R})\, \Phi^{-1}(\lambda_{\mathrm{GL}(n)})(\mathrm{d}\boldsymbol{R}) \tag{4.124}$$

As $\mathrm{Mat}(n,n) \setminus \mathrm{GL}(n)$ is a $\lambda_{\mathrm{Mat}(n,n)}$-zero set the left-hand integral equals the volume of the unit ball in $\mathbb{R}^{n^2}$ whereas the right-hand integral can be evaluated by applying (4.121). This yields (4.122), and (4.123) follows by induction (cf. the proof of Theorem 2.4(iii) in [69]). □

As $\mu_{\mathrm{GL}(n)} \ll \lambda_{\mathrm{GL}(n)}$ and $\lambda_{\mathrm{GL}(n)} \ll \mu_{\mathrm{GL}(n)}$ and since the respective densities are invariant under the left multiplication with orthogonal matrices applying 4.72(ii) or 4.78(ii) (Section 4.8) one immediately obtains $\Phi^{-1}(\mu_{\mathrm{GL}(n)})$ from $\Phi^{-1}(\lambda_{\mathrm{GL}(n)})$ and, vice versa, $\Phi^{-1}(\lambda_{\mathrm{GL}(n)})$ from $\Phi^{-1}(\mu_{\mathrm{GL}(n)})$. This observation is significant for various problems (see, e.g., the proof of Theorem 4.72(iii)) as we can choose that measure which is more suitable in the concrete situation.

Next, we formulate a lemma which summarizes some interesting properties which will be needed in the sequel. We note that (iii) is from [24], and

(iv) says that for the evaluation of Lebesgue integrals we may change the domain of integration from GL(n) to Mat(n,n) and vice versa. This means that we can always use that domain of integration which is more suitable in the concrete situation.

Lemma 4.73. *(i) For each $x>0$ we have*

(a) $\Gamma(x+1)=x\Gamma(x)$.
(b) $\Gamma(x)\Gamma\left(x+\frac{1}{2}\right)=\sqrt{\pi}\,\Gamma(2x)/2^{2x-1}$.

(ii) The volume of the unit ball in $\mathbb{R}^m$ equals $\pi^{\frac{m}{2}}/\Gamma\left(\frac{m}{2}+1\right)$.
(iii) Assume that $g\colon\mathbb{R}\to\mathbb{R}$ is a measurable mapping, and for $j\le m$ let β_j, a_j and p_j denote positive real numbers. Then

$$\int\limits_{\substack{\left(\frac{x_1}{a_1}\right)^{\beta_1}+\cdots+\left(\frac{x_m}{a_m}\right)^{\beta_m}\le 1\\ x_1,\ldots,x_m\ge 0}}\cdots\int g\left(\left(\frac{x_1}{a_1}\right)^{\beta_1}+\cdots+\left(\frac{x_m}{a_m}\right)^{\beta_m}\right)x_1^{p_1-1}\cdots x_m^{p_m-1}\,\mathrm{d}x_1\cdots\mathrm{d}x_m$$
$$=\frac{a_1^{p_1}\cdots a_m^{p_m}\,\Gamma\left(\frac{p_1}{\beta_1}\right)\cdots\Gamma\left(\frac{p_m}{\beta_m}\right)}{\beta_1\cdots\beta_m\;\Gamma\left(\frac{p_1}{\beta_1}+\cdots+\frac{p_m}{\beta_m}\right)}\int_0^1 g(t)\,t^{\frac{p_1}{\beta_1}+\cdots+\frac{p_m}{\beta_m}-1}\,\mathrm{d}t. \tag{4.125}$$

(iv)

$$\lambda_{Mat(n,n)}\left(\mathrm{Mat}(n,n)\setminus\mathrm{GL}(n)\right)=0 \tag{4.126}$$

(v) (Transformation Theorem in $\mathbb{R}^n$) Assume that U and V are open subsets of $\mathbb{R}^n$, and $\chi\colon U\to V$ is a continuously differentiable homeomorphism. Then the function $f\colon V\to\mathbb{R}$ is λ_n-integrable iff $(f\circ\chi)|\det D\chi|\colon U\to\mathbb{R}$ is λ_n-integrable. In this case

$$\int_U f\left(\chi(\mathbf{x})\right)\left|\det D\chi(\mathbf{x})\right|\,\mathrm{d}\mathbf{x}=\int_V f(\mathbf{y})\,\mathrm{d}\mathbf{y}. \tag{4.127}$$

If X is a $g\cdot\lambda_n$-distributed random variable on V then $\chi^{-1}(X)$ assumes values in U. The random variable $\chi^{-1}(X)$ has the λ_n-density $\left(g\circ\chi^{-1}\right)\left|\det D\chi^{-1}\right|$.

Proof. The statements (i)–(iv) correspond with Lemma 2.5 in [65]. The first assertion of (v) is shown in [27], pp. 120f., whereas the second is an immediate consequence from the first. □

Example 4.74. We compute the integral $\int_{K_{GL(n)}(b)}|\det\boldsymbol{M}|^u\,\mathrm{d}\boldsymbol{M}$ with $b>0$ and $u>-1$.
Theorem 4.72(iii) and the transformation theorem in $\mathbb{R}^{n(n+1)/2}$ imply

$$\int\limits_{K_{GL(n)}(b)}|\det\boldsymbol{M}|^u\,\mathrm{d}\boldsymbol{M}=c_{\mathrm{Q};n}\int\limits_{K_{\mathrm{R}(n)}(b)}\prod_{j=1}^n r_{jj}^{n-j+u}\,\mathrm{d}r_{11}\cdots\mathrm{d}r_{nn} \tag{4.128}$$
$$=2^{n(n-1)/2}\,c_{\mathrm{Q};n}\,b^{n^2+nu}\int\limits_{K_{\mathrm{R}(n)}(1)\cap\{\mathbf{R}|r_{ij}\ge 0\}}\prod_{j=1}^n r_{jj}^{n-j+u}\,\mathrm{d}r_{11}\cdots\mathrm{d}r_{nn}.$$

Applying 4.73(iii) with $p_{jj} = n - j + u + 1$ and $p_{ij} = 1$ we obtain for $i < j$

$$= c_{\mathrm{Q};n} \frac{2^{n(n-1)/2} b^{n^2+nu} \prod_{j=1}^{n} \Gamma\left(\frac{n-j+u+1}{2}\right) \left(\Gamma\left(\frac{1}{2}\right)\right)^{n(n-1)/2}}{2^{n(n+1)/2} \Gamma\left(\frac{n^2+nu}{2}\right)} \int_0^b t^{\frac{n^2+nu}{2}-1} \,\mathrm{d}t$$

$$= b^{n^2+nu} \frac{2^n \sqrt{\pi}^{\,n(n+1)/2}}{\prod_{j=1}^{n} \Gamma\left(\frac{j}{2}\right)} \frac{2^{-n} \prod_{j=1}^{n} \Gamma\left(\frac{j+u}{2}\right) \sqrt{\pi}^{\,n(n-1)/2}}{\Gamma\left(\frac{n^2+nu}{2}\right) \left(\frac{n^2+nu}{2}\right)}$$

as $\sum_{j=1}^{n}(n - j + u + 1) + n(n-1)/2 = n^2 + nu$. Elementary arithmetical operations finally yield

$$\int\limits_{K_{\mathrm{GL}(n)}(b)} |\det \boldsymbol{M}|^u \,\mathrm{d}\boldsymbol{M} = b^{n^2+nu} \frac{\pi^{\frac{n^2}{2}}}{\Gamma\left(\frac{n^2+nu+2}{2}\right)} \prod_{j=1}^{n} \frac{\Gamma\left(\frac{j+u}{2}\right)}{\Gamma\left(\frac{j}{2}\right)}. \quad (4.129)$$

Example 4.75. From 4.72 and 4.73(iii) we obtain

$$\int\limits_{\mathrm{GL}(n)} \mathrm{e}^{-\frac{1}{2}\|\mathbf{M}\|_{\mathrm{F}}^2} |\det \boldsymbol{M}| \,\mathrm{d}\boldsymbol{M} = c_{\mathrm{Q};n} \int\limits_{\mathrm{R}(n)} e^{-\frac{1}{2}\|\mathbf{R}\|_{\mathrm{F}}^2} \prod_{j=1}^{n} r_{jj}^{n-j+1} \,\mathrm{d}r_{11} \cdots r_{nn}$$

$$= c_{\mathrm{Q};n} \prod_{j=1}^{n} \int_0^\infty \mathrm{e}^{-\frac{1}{2}x^2} x^{n-j+1} \,\mathrm{d}x \prod_{1 \le i < j \le n} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} \,\mathrm{d}x$$

$$= c_{\mathrm{Q};n} \prod_{j=1}^{n} \frac{\Gamma\left(\frac{n-j+2}{2}\right) 2^{\frac{n-j+2}{2}}}{2} \sqrt{2\pi}^{\,n(n-1)/2}. \quad (4.130)$$

The last equation follows from formula 1.1.3.4.2 in [14]. As $\prod_{j=1}^{n} \Gamma\left(\frac{n-j+2}{2}\right) = \prod_{j=2}^{n+1} \Gamma\left(\frac{j}{2}\right)$ we finally obtain

$$\int\limits_{\mathrm{GL}(n)} \mathrm{e}^{-\frac{1}{2}\|\mathbf{M}\|_{\mathrm{F}}^2} |\det \boldsymbol{M}| \,\mathrm{d}\boldsymbol{M} = 2^{n(n+1)/2} \pi^{\frac{n^2-1}{2}} \Gamma\left(\tfrac{n+1}{2}\right). \quad (4.131)$$

The preceding examples illustrate the applicability of Theorem 4.72 and the main advantages of the QR-decomposition for integration problems on $\mathrm{Mat}(n,n)$ and $\mathrm{GL}(n)$: If the integrand and the measure are invariant under the left multiplication with orthogonal matrices the dimension of the domain of integration reduces from n^2 to $n(n+1)/2$ which is favourable for both analytical and numerical computations. Moreover, in many cases the integrand simplifies considerably. The determinant function, for example, simplifies to the product of the diagonal elements. We note that these examples correspond with Examples 2.6 and 2.7 in [69]. Equation (4.131) will be used in the proof of Theorem 4.91 for the computation of the constant b_n. Example 4.76 considers a Borel measure $\eta = f \cdot \lambda_{\mathrm{GL}(n)} \notin \mathcal{M}_{\mathrm{O}(n)}(\mathrm{GL}(n))$. We determine the unique $\mathrm{O}(n)$-invariant measure η^* which coincides with η on the sub-σ-algebra $\mathcal{B}_{\mathrm{O}(n)}(\mathrm{GL}(n))$ (cf. Theorem 4.7).

Example 4.76. Let $f\colon \mathrm{GL}(n) \to \mathbb{R}$, $f(\boldsymbol{M}) := m_{11}^2$. Clearly, $\eta := f \cdot \lambda_{GL(n)} \in \mathcal{M}(\mathrm{GL}(n))$ but $\eta \notin \mathcal{M}_{\mathrm{O}(n)}(\mathrm{GL}(n))$. Applying 4.7(iii) yields

$$\begin{aligned} f^*(\boldsymbol{M}) &= \int_{\mathrm{O}(n)} f(\boldsymbol{T}\boldsymbol{M})\, \mu_{\mathrm{O}(n)}(\mathrm{d}\boldsymbol{T}) = \int_{\mathrm{O}(n)} (\mathbf{e}_1^t \boldsymbol{T}\boldsymbol{M}\mathbf{e}_1)^2\, \mu_{\mathrm{O}(n)}(\mathrm{d}\boldsymbol{T}) \\ &= \int_{\mathrm{O}(n)} (\mathbf{e}_1^t \boldsymbol{T}\mathbf{m_1})^2\, \mu_{\mathrm{O}(n)}(\mathrm{d}\boldsymbol{T}) \end{aligned}$$

where $\mathbf{m_1}$ denotes the first column vector of $\boldsymbol{M} \in \mathrm{GL}(n)$. The group $\mathrm{O}(n)$ acts onto itself and on S^{n-1} by left multiplication, and with respect to these actions the mapping $\mathrm{O}(n) \to \mathrm{S}^{n-1}$, $\boldsymbol{T} \mapsto \boldsymbol{T}\mathbf{m_1}/\|\mathbf{m_1}\|$ is $\mathrm{O}(n)$-equivariant. Applying 4.6(i), (iii) yields

$$f^*(\boldsymbol{M}) = \int_{\mathrm{S}^{n-1}} (\mathbf{e}_1^t \mathbf{y}\|\mathbf{m_1}\|)^2\, \mu_{(\mathrm{S}^{n-1})}(\mathrm{d}\mathbf{y}) = \|\mathbf{m_1}\|^2 \int_{\mathrm{S}^{n-1}} y_1^2\, \mu_{(\mathrm{S}^{n-1})}(\mathrm{d}\mathbf{y})$$

where $\mu_{(\mathrm{S}^{n-1})}$ denotes the unique $\mathrm{O}(n)$-invariant probability measure on S^{n-1}. Consequently,

$$f^*(\boldsymbol{M}) = \frac{\|\mathbf{m_1}\|^2}{n} \int_{\mathrm{S}^{n-1}} (y_1^2 + \cdots + y_n^2)\, \mu_{(\mathrm{S}^{n-1})}(\mathrm{d}\mathbf{y}) = \frac{\|\mathbf{m_1}\|^2}{n} \tag{4.132}$$

From 4.7(ii) and (iii) $\eta^* := f^* \cdot \lambda_{GL(n)} \in \mathcal{M}_{\mathrm{O}(n)}(\mathrm{GL}(n))$ and $\eta^*_{|\mathcal{B}_{\mathrm{O}(n)}(GL(n))} = \eta_{|\mathcal{B}_{\mathrm{O}(n)}(GL(n))}$.

4.8 Polar Decomposition on GL(n)

In the present section we treat the polar decomposition on $\mathrm{GL}(n)$. As the QR-decomposition in the preceding section the polar decomposition represents a further special case of Iwasawa's Theorem (cf. Section 4.6). The situation here is very similar to that in Section 4.7. Again, we consider the Borel measures on $\mathrm{GL}(n)$ which are invariant under the left multiplication with orthogonal matrices.

Each regular $(n \times n)$-matrix $\boldsymbol{M}$ can be represented as a product $\boldsymbol{M} = \boldsymbol{T}_{\mathbf{M}}\boldsymbol{P}_{\mathbf{M}}$ with $\boldsymbol{T}_{\mathbf{M}} \in \mathrm{O}(n)$ and $\boldsymbol{P}_{\mathbf{M}} \in \mathrm{Pos}(n)$. Both factors are unique (*polar decomposition*) ([47], p. 198). We recall that the orthogonal group $\mathrm{O}(n)$ acts on $\mathrm{GL}(n)$ and $\mathrm{O}(n)$ by left multiplication and that the set of all $\mathrm{O}(n)$-invariant Borel measures on $\mathrm{GL}(n)$ is denoted with $\mathcal{M}_{\mathrm{O}(n)}(\mathrm{GL}(n))$. The term $\mathrm{Sym}(n)$ denotes the vector space of all real symmetric $(n \times n)$-matrices. Moreover, we refer to the Definition 4.70.

Lemma 4.77. *(i) The orthogonal group* $\mathrm{O}(n)$ *is a maximal compact subgroup of* $\mathrm{GL}(n)$.
(ii) The mapping

$$\Theta_r\colon \mathrm{Pos}(n) \times \mathrm{GL}(n) \to \mathrm{Pos}(n), \qquad \Theta_r(\boldsymbol{P}, \boldsymbol{M}) := \boldsymbol{M}^t \boldsymbol{P}\boldsymbol{M} \tag{4.133}$$

defines a right action of $\mathrm{GL}(n)$ *on* $\mathrm{Pos}(n)$. *More precisely,* $\mathrm{Pos}(n)$ *is a homogeneous* $\mathrm{GL}(n)$*-space with respect to this action. The* $\mathrm{GL}(n)$*-invariant measures on* $\mathrm{Pos}(n)$ *are of the form*

$$\mu_{(\mathrm{P}os(n))} = h_{\mu;\mathrm{P}os(n)} \cdot \lambda_{\mathrm{P}os(n)} \quad \textit{with} \quad h_{\mu;\mathrm{P}os(n)} := c_\mu \,(\det \boldsymbol{P})^{-(n+1)/2} \tag{4.134}$$

where c_μ *is a positive constant.*
(iii) The mapping

$$\varphi_P \colon \mathrm{O}(n) \times \mathrm{Pos}(n) \to \mathrm{GL}(n), \qquad \varphi_P(\boldsymbol{T}, \boldsymbol{P}) := \boldsymbol{T}\boldsymbol{P} \tag{4.135}$$

is a diffeomorphism with inverse mapping

$$\varphi_P^{-1}(\boldsymbol{M}) = \left(\boldsymbol{M} \sqrt{\boldsymbol{M}^t \boldsymbol{M}}^{\,-1}, \sqrt{\boldsymbol{M}^t \boldsymbol{M}} \right). \tag{4.136}$$

(iv) The mapping

$$\chi_S \colon \mathrm{Pos}(n) \to \mathrm{Pos}(n), \qquad \chi_S(\boldsymbol{P}) := \boldsymbol{P}^2 \tag{4.137}$$

is a diffeomorphism. More precisely,

$$\det D\chi_S(\boldsymbol{P}) = 2^n \det \boldsymbol{P} \prod_{1 \le i < j \le n} (\rho_i(\boldsymbol{P}) + \rho_j(\boldsymbol{P})). \tag{4.138}$$

(v) The orthogonal group $\mathrm{O}(n)$ *acts on* $\mathrm{GL}(n)$ *and* $\mathrm{O}(n)$ *by left multiplication. The* 5*-tuple* $(\mathrm{O}(n), \mathrm{O}(n), \mathrm{Pos}(n), \mathrm{GL}(n), \varphi_P)$ *has Property (*).*

Proof. Assertion (i) equals 4.71(i), and (ii) is shown in [11], Chap. 7, par. 3, no. 3, Exemple 8. The uniqueness of the polar decomposition implies the bijectivity of φ_P. Since $\mathbf{v}\boldsymbol{M}^t\boldsymbol{M}\mathbf{v} = (\boldsymbol{M}\mathbf{v}) \cdot (\boldsymbol{M}\mathbf{v}) > 0$ for each $\mathbf{v} \neq 0$ the second component of φ_P^{-1} is positive definite and symmetric which verifies (4.136). Clearly, φ_P is smooth. It remains to verify that $\det D\varphi_P(\boldsymbol{T}_0, \boldsymbol{P}_0) \neq 0$ for each $(\boldsymbol{T}_0, \boldsymbol{P}_0) \in \mathrm{O}(n) \times \mathrm{Pos}(n)$ (cf. 4.16). The proof of Lemma 4.71(iii) can almost literally be transferred. In a final step it remains to be proved that $\Delta\boldsymbol{F}\boldsymbol{P}_0 = -\Delta\boldsymbol{P}$ implies $\Delta\boldsymbol{F} = 0 = \Delta\boldsymbol{P}$ for skew-symmetric $\Delta\boldsymbol{F}$ and symmetric $\Delta\boldsymbol{P}$. In fact, $-\Delta\boldsymbol{P} = \Delta\boldsymbol{F}\boldsymbol{P}_0 = (\Delta\boldsymbol{F}\boldsymbol{P}_0)^t = -\boldsymbol{P}_0\Delta\boldsymbol{F}$ as $\Delta\boldsymbol{F}$ is skew-symmetric. Let $\mathbf{v} \in \mathbb{R}^n$ be an eigenvector of $\boldsymbol{P}_0$, i.e. $\boldsymbol{P}_0\mathbf{v} = \rho_j(\boldsymbol{P}_0)\mathbf{v}$ for any $j \le n$. Then $-\rho_j(\boldsymbol{P}_0)(\Delta\boldsymbol{F}\mathbf{v}) = \boldsymbol{P}_0(\Delta\boldsymbol{F}\mathbf{v})$. As $\boldsymbol{P}_0$ is positive definite $\Delta\boldsymbol{F}\mathbf{v}_j = 0$ for each eigenvector of $\boldsymbol{P}_0$ and hence $\Delta\boldsymbol{F} = 0$ which completes the proof of (iii). To prove (iv) we first note that $\mathrm{Pos}(n) \to \mathrm{Pos}(n)$, $\boldsymbol{P} \mapsto \sqrt{\boldsymbol{P}}$ is unique ([47], p. 198). Hence χ_S is bijective with inverse mapping $\chi_S^{-1}(\boldsymbol{P}) := \sqrt{\boldsymbol{P}}$. It remains to prove (4.138) (cf. Remark 4.16). Clearly, $\chi_S(\boldsymbol{P} + \Delta\boldsymbol{P}) - \chi_S(\boldsymbol{P}) = \boldsymbol{P}\Delta\boldsymbol{P} + \Delta\boldsymbol{P}\boldsymbol{P} + (\Delta\boldsymbol{P})^2$ and hence $D\chi_S(\boldsymbol{P})(\Delta\boldsymbol{P}) = \boldsymbol{P}\Delta\boldsymbol{P} + \Delta\boldsymbol{P}\boldsymbol{P}$. For each $\boldsymbol{T} \in \mathrm{O}(n)$ the mapping $\mathrm{C}_{\mathbf{T}} \colon \mathrm{Sym}(n) \to \mathrm{Sym}(n)$, $\mathrm{C}_{\mathbf{T}}(\boldsymbol{S}) := \boldsymbol{T}\boldsymbol{S}\boldsymbol{T}^t$ is linear with inverse $\mathrm{C}_{\mathbf{T}}^{-1} = \mathrm{C}_{\mathbf{T}^t}$. In particular, $\mathrm{C}_{\mathbf{T}}^{-1} \circ D\chi_S(\boldsymbol{T}\boldsymbol{P}\boldsymbol{T}^t) \circ \mathrm{C}_{\mathbf{T}} = D\chi_S(\boldsymbol{P})$. As all mappings are linear the computation of the determinant

function on both sides yields $\det D\chi_S(\boldsymbol{P}) = \det D\chi_S(\boldsymbol{TPT}^t)$. For suitably chosen $\boldsymbol{T} \in \mathrm{O}(n)$ the matrix $\boldsymbol{D} = \boldsymbol{TPT}^t$ is diagonal. Then $[D\chi_S(\boldsymbol{D})(\boldsymbol{S})]_{ij} = [\boldsymbol{DS} + \boldsymbol{SD}]_{ij} = (d_{ii} + d_{jj})s_{ij}$ for all $\boldsymbol{S} \in \mathrm{Sym}(n)$. Let for the moment $\mathbf{E}^{kl} = (e_{ij}^{kl})_{1\le i,j\le n}$ be given by $e_{kl}^{kl} = e_{lk}^{kl} = 1$ and $e_{ij}^{kl} = 0$ else $(1 \le k \le l \le n)$. Clearly, $\{\mathbf{E}^{kl} \mid 1 \le k \le l \le n\}$ is a basis of the vector space $\mathrm{Sym}(n)$. With respect to this basis the nondiagonal elements of the matrix representation of $D\chi_S(\boldsymbol{D})$ are zero while its diagonal entries equal $\{d_{ii} + d_{jj} \mid 1 \le i \le j \le n\}$. This proves (4.138) since $\prod_{i=1}^n 2\rho_i(\boldsymbol{P}) = 2^n \det \boldsymbol{P}$. Statement (v) is an immediate consequence from (i), (iii) and 4.67(ii). □

The polar decomposition represents a further special case of Iwasawa's Theorem. Note that Theorem 4.78 below does only use the fact that φ_P is a homeomorphism (cf. 4.69(i)).

Theorem 4.78. *Let* $\Phi\colon \mathcal{M}^\sigma(\mathrm{Pos}(n)) \to \mathcal{M}^+_{\mathrm{O}(n)}(\mathrm{GL}(n)), \Phi(\tau) := \big(\mu_{\mathrm{O}(n)} \otimes \tau\big)^{\varphi_P}$. *Then the following statements are valid:*
(i) $\Phi(\mathcal{M}(\mathrm{Pos}(n))) = \mathcal{M}_{\mathrm{O}(n)}(\mathrm{GL}(n))$ *and* $\Phi^{-1}(\mathcal{M}_{\mathrm{O}(n)}(\mathrm{GL}(n))) = \mathcal{M}(\mathrm{Pos}(n))$. *In particular,* Φ *is bijective onto its image.*
(ii) *Let* $\nu \in \mathcal{M}_{\mathrm{O}(n)}(\mathrm{GL}(n))$ *and* $\eta = f \cdot \nu \in \mathcal{M}(\mathrm{Pos}(n))$ *where* f *is invariant under the left multiplication with orthogonal matrices. Then* $\eta \in \mathcal{M}_{\mathrm{O}(n)}(\mathrm{GL}(n))$, *and* $\Phi^{-1}(\nu) = \tau_\nu$ *implies* $\Phi^{-1}(\eta) = f_{|\mathrm{P}os(n)} \cdot \tau_\nu$.
(iii) *In particular*

$$\begin{aligned} &\lambda_{\mathrm{G}L(n)} = \big(\mu_{\mathrm{O}(n)} \otimes \mathrm{h}_{\mathrm{P}os(n)} \cdot \lambda_{\mathrm{P}os(n)}\big)^{\varphi_P} \quad \textit{with} \\ &\mathrm{h}_{\mathrm{P}os(n)}(\mathbf{P}) := c_{\mathrm{P};n} \prod_{1\le i<j\le n} (\rho_i(\boldsymbol{P}) + \rho_j(\boldsymbol{P})) \end{aligned} \tag{4.139}$$

for a suitable constant $c_{\mathrm{P};n} > 0$.
(iv) *Let*

$$q_*\colon \mathrm{GL}(n) \to \mathrm{Pos}(n), \qquad q_*(\boldsymbol{M}) := \boldsymbol{M}^t\boldsymbol{M}. \tag{4.140}$$

Suppose that $\nu = f \cdot \lambda_{\mathrm{GL}(n)}$ *with* $\mathrm{O}(n)$*-invariant density* f. *Then*

$$\nu^{q_*} = f_{q*} \cdot \lambda_{\mathrm{P}os(n)} \qquad \textit{with} \qquad f_{q*}(\boldsymbol{P}) := \frac{c_{\mathrm{P};n}\, f\left(\sqrt{\boldsymbol{P}}\right)}{2^n \sqrt{\det \boldsymbol{P}}}. \tag{4.141}$$

(v) *Let* $\mathrm{pr}_2\colon \mathrm{O}(n) \times \mathrm{Pos}(n) \to \mathrm{Pos}(n)$ *denote a projection onto the second component. The mapping* $(\mathrm{pr}_2 \circ \varphi_P^{-1})^m$ *is a sufficient statistic for all test problems* $(\mathrm{P}_{\Gamma_0}, \mathrm{P}_{\Gamma\setminus\Gamma_0})$ *with* $\mathrm{P}_\Gamma \subseteq \mathcal{M}^1_{\mathrm{O}(n)}(\mathrm{GL}(n))^m$.

Proof. The statements (i), (ii) and (v) follow from 4.77 and 4.68. It hence remains to prove (iii) and (iv). Clearly, $\mathrm{GL}(n) \times \mathrm{GL}(n) \to \mathrm{GL}(n)$, $(\boldsymbol{M}, \boldsymbol{N}) \mapsto \boldsymbol{MN}$ and $\mathrm{Pos}(n) \times \mathrm{GL}(n) \to \mathrm{Pos}(n)$, $(\boldsymbol{P}, \boldsymbol{N}) \mapsto \boldsymbol{N}^t\boldsymbol{PN}$ define $\mathrm{GL}(n)$ right actions (cf. 4.77(ii)). A simple calculation shows $q^*(\boldsymbol{MN}) = \boldsymbol{N}^t q^*(\boldsymbol{M})\boldsymbol{N}$, i.e. q^* is equivariant with respect to these $\mathrm{GL}(n)$-actions. As in the proof of 2.14(i) one immediately obtains

$$q_*^{-1}(\boldsymbol{N}^t B\boldsymbol{N}) = \{\boldsymbol{M} \in \mathrm{GL}(n) \mid \boldsymbol{M}^t\boldsymbol{M} \in \boldsymbol{N}^t B\boldsymbol{N}\}$$
$$= \{\boldsymbol{M} \in \mathrm{GL}(n) \mid (\boldsymbol{M}\boldsymbol{N}^{-1})^t(\boldsymbol{M}\boldsymbol{N}^{-1}) \in B\} = q_*^{-1}(B)\boldsymbol{N}$$

for all $\boldsymbol{N} \in \mathrm{GL}(n)$ and $B \in \mathcal{B}(\mathrm{Pos}(n))$. As $\mathrm{GL}(n)$ is unimodular $\mu_{\mathrm{GL}(n)}$ is right invariant and its image measure $\mu_{\mathrm{GL}(n)}^{q^*}$ is $\mathrm{GL}(n)$-invariant, too. On the other hand, from (4.136) we conclude $q^* = \chi_S \circ \mathrm{pr}_2 \circ \varphi_P^{-1}$. As χ_S is a diffeomorphism $\mu_{\mathrm{GL}(n)}^{q^*} = (\Phi^{-1}(\mu_{\mathrm{GL}(n)}))^{\chi_S} \in \mathcal{M}(\mathrm{Pos}(n))$ and hence $\mu_{\mathrm{GL}(n)}^{q^*} = \mu_{(\mathrm{P}os(n))}$ if the constant c'_μ (cf. 4.77(ii)) is chosen suitably. From the transformation theorem on $\mathrm{Pos}(n)$ we conclude that $\Phi^{-1}(\mu_{\mathrm{GL}(n)})$ has a $\lambda_{\mathrm{P}os(n)}$-density. Applying (ii) with $\nu = \mu_{\mathrm{GL}(n)}$ and $\eta := h_{\mu,n}^{-1} \cdot \nu = \lambda_{\mathrm{GL}(n)}$ (cf. (4.116)) we obtain $\lambda_{\mathrm{GL}(n)} = \big(\mu_{\mathrm{O}(n)} \otimes \mathrm{h}_{\mathrm{P}os(n)} \cdot \lambda_{\mathrm{P}os(n)}\big)^{\varphi_P}$ for a particular density $\mathrm{h}_{\mathrm{P}os(n)}$. If $g\colon \mathrm{Pos}(n) \to \mathbb{R}$ is $\lambda_{\mathrm{P}os(n)}$-integrable the transformation theorem implies

$$\int_B \frac{g(\sqrt{\boldsymbol{P}})}{|\det D\chi_S(\sqrt{\boldsymbol{P}})|}\lambda_{\mathrm{P}os(n)}(\mathrm{d}\boldsymbol{P})$$
$$= \int_B g(\chi_S^{-1}(\boldsymbol{P}))\left|\det\left(D\chi_S^{-1}\right)(\boldsymbol{P})\right|\lambda_{\mathrm{P}os(n)}(\mathrm{d}\boldsymbol{P}) = \int_{\chi_S^{-1}(B)} g(\boldsymbol{S})\,\lambda_{\mathrm{P}os(n)}(\mathrm{d}\boldsymbol{S}).$$

In other words: The mapping χ_S transforms any $\lambda_{\mathrm{P}os(n)}$-density g into the $\lambda_{\mathrm{P}os(n)}$-density g^{χ_S} which is given by $g^{\chi_S}(\boldsymbol{P}) := g(\sqrt{\boldsymbol{P}})/|\det D\chi_S(\sqrt{\boldsymbol{P}})|$. Now assume $\nu = f \cdot \lambda_{\mathrm{GL}(n)}$ with $\mathrm{O}(n)$-invariant density f. From (ii) $\nu = \big(\mu_{\mathrm{O}(n)} \otimes f_{|\mathrm{P}os(n)}\mathrm{h}_{\mathrm{P}os(n)} \cdot \lambda_{\mathrm{P}os(n)}\big)^{\varphi_P}$ and

$$\nu^{q_*} = (\Phi^{-1}(\nu))^{\chi_S} = f^{q_*} \cdot \lambda_{\mathrm{P}os(n)} \qquad \text{with } f^{q_*}(\boldsymbol{P}) := \frac{h_\lambda(\sqrt{\boldsymbol{P}}) \cdot f(\sqrt{\boldsymbol{P}})}{|\det D\chi_S(\sqrt{\boldsymbol{P}})|}.$$

Recall that $\mu_{\mathrm{GL}(n)}^{q^*} = \mu_{(\mathrm{P}os(n))}$. Especially, for $f = h_{\mu;n}$ equation (4.134) implies

$$\frac{c_\mu}{(\det \boldsymbol{P})^{(n+1)/2}} = \frac{h_\lambda(\sqrt{\boldsymbol{P}})(\det \boldsymbol{P})^{-n/2}}{2^n(\det \boldsymbol{P})^{1/2}\prod_{1\le i<j\le n}\left(\rho_i(\sqrt{\boldsymbol{P}}) + \rho_j(\sqrt{\boldsymbol{P}})\right)}$$

which yields $h_\lambda(\sqrt{\boldsymbol{P}}) = 2^n c_\mu \prod_{1\le i<j\le n}(\rho_i(\sqrt{\boldsymbol{P}}) + \rho_j(\sqrt{\boldsymbol{P}}))$. This proves (4.139) with $c_{\mathrm{P};n} = 2^n c_\mu$. Inserting the expressions for $\mathrm{h}_{\mathrm{P}os(n)}$ and $|\det D\chi_S|$ in the above formula for f^{q^*} proves (4.141), and this completes the proof of this theorem. □

Remark 4.79. Theorem 4.78(iii) provides $\mathrm{h}_{\mathrm{P}os(n)}(\boldsymbol{P})$ as a function of the eigenvalues of $\boldsymbol{P}$. For integration problems on $\mathrm{Pos}(n)$, however, one usually needs an expression in the components of $\boldsymbol{P}$. Clearly,

$$\mathrm{h}_{\mathrm{P}os(n)}(\boldsymbol{P}) = c_{P;2} \cdot \mathrm{tr}\boldsymbol{P} \qquad \text{for } n = 2. \tag{4.142}$$

For the general case $n \geq 2$ recall that $D\chi_S(\boldsymbol{P})(\boldsymbol{S}) = \boldsymbol{PS} + \boldsymbol{SP}$. Hence $\det D\chi_S(\boldsymbol{P})$ is a polynomial in the components of $\boldsymbol{P}$. (Note that $\det D\chi_S$ is positive on $\mathrm{Pos}(n)$ by (4.138).) Inserting this expression yields a representation of $\mathrm{h}_{\mathrm{P}os(n)}$ as a polynomial in the components of its argument. As $\rho_i(c\boldsymbol{P}) = c\rho_i(\boldsymbol{P})$ for each scalar $c > 0$ and $D\chi_S(c\boldsymbol{P})(\boldsymbol{S}) = c(\boldsymbol{PS} + \boldsymbol{SP})$ this polynomial is homogeneous of degree $n(n+1)/2$. More precisely, in [64], pp. 108, it is verified that $\mathrm{h}_{\mathrm{P}os(n)}(\boldsymbol{P})$ is a scalar multiple of $\det(\boldsymbol{F}_1 + \boldsymbol{F}_2)$ for $\boldsymbol{F}_1, \boldsymbol{F}_2 \in \mathrm{Mat}(n,n)$ with

$$(\boldsymbol{F}_1)_{(i-1)n+j,(k-1)n+l} = \begin{cases} 1 & \text{if } i = j \text{ and } (k,l) = (i,i) \\ \frac{1}{2} & \text{if } i \neq j \text{ and } ((k,l) = (i,j) \text{ or } (k,l) = (j,i)) \\ 0 & \text{else} \end{cases}$$

$$(\boldsymbol{F}_2)_{(i-1)n+j,(k-1)n+l} = \begin{cases} \frac{1}{2}p_{lj} & \text{if } k = i \text{ and } l \neq i \\ -\frac{1}{2}p_{kj} & \text{if } l = i \text{ and } k \neq i \\ 0 & \text{else.} \end{cases} \tag{4.143}$$

We mention that $\boldsymbol{F}_1 + \boldsymbol{F}_2$ has merely $2n^3 - 2n^2 + n$ nonzero entries. Adding its $((l-1)n+k)^{th}$ column to the $((k-1)n+l)^{th}$ column $(1 \leq k < l \leq n)$ and cancelling the $((i-1)n+i)^{th}$ row and column $(1 \leq i \leq n)$ simplifies the computation of the determinant additionally. For $n = 3$ an elementary but careful computation yields

$$\begin{aligned} \det(\boldsymbol{F}_1 + \boldsymbol{F}_2) = \frac{1}{8}\big[&(p_{11}+p_{22})(p_{11}+p_{33})(p_{22}+p_{33}) - (p_{11}+p_{22})p_{21}^2 \\ &-(p_{11}+p_{33})p_{31}^2 - (p_{22}+p_{33})p_{32}^2 - p_{21}p_{31}p_{32}\big]. \end{aligned} \tag{4.144}$$

Theorem 4.78 is an analogon of 4.72, and hence one might expect similar benefits for applications. Unfortunately, at least for $n > 2$ Theorem 4.78 does only support very specific integration problems since $\mathrm{Pos}(n)$ is a rather unsuitable domain of integration for both exact and numerical computations. Unlike for $c_{Q;n}$ (cf. Section 4.7) it even seems to be very tough to give a closed expression for the constant $c_{\mathrm{P};n}$ for all n. For any fixed n, of course, it can be determined by the evaluation of a reference integral. On $\mathrm{Pos}(n)$ numerical methods may be applied. For $n = 2$ a careful computation yields

$$c_{\mathrm{P};2} = 4\pi \tag{4.145}$$

([64], p. 115).

The relevance of the polar decomposition for applications (in connection with our symmetry concept) will become clear in Section 4.10 where we combine Theorem 4.78 with the results from the following section. This will provide useful insights in the class of biinvariant measures on GL(n) which are of particular significance for a number of applications. In the remainder of this section we study a 'by-product' of Theorem 4.78. To be precise, 4.78(iv) can be used for the evaluation of specific integrals and it enables the efficient simulation of a specific class of unbounded Lebesgue densities on $\mathrm{Pos}(n)$ or $\mathbb{R}^{n(n+1)/2}$, resp.

For the vast majority of $\mathrm{O}(n)$-invariant integration and simulation problems on $\mathrm{GL}(n)$ it is clearly advisable to transfer the necessary computations to $\mathrm{O}(n)$ and $\mathrm{R}(n)$, or $\mathrm{O}(n)$ and $\mathrm{Pos}(n)$, resp. As $\Phi^{-1}\colon \mathcal{M}_{\mathrm{O}(n)}(\mathrm{GL}(n)) \to \mathcal{M}(\mathrm{Pos}(n))$ and the square mapping $\chi_S\colon \mathrm{Pos}(n) \to \mathrm{Pos}(n)$ are bijective the mapping

$$\mathcal{M}_{\mathrm{O}(n)}(\mathrm{GL}(n)) \to \mathcal{M}(\mathrm{Pos}(n)), \qquad \nu \mapsto \nu^{q_*} \tag{4.146}$$

is bijective, too. This observation can be used to transfer the esssential parts of specific integration or simulation problems from $\mathrm{Pos}(n)$ to $\mathrm{GL}(n)$. In particular, for each $B \in \mathrm{Pos}(n)$ and any ν^{q^*}-integrable function $g\colon \mathrm{Pos}(n) \to \mathbb{R}$ we have

$$\int_B g(\boldsymbol{P})\, \nu^{q^*}(\mathrm{d}\boldsymbol{P}) = \int_{q^{*-1}(B)} g \circ q^*(\boldsymbol{M})\, \nu(\mathrm{d}\boldsymbol{M}) \qquad \text{for } \nu \in \mathcal{M}_{\mathrm{O}(n)}(\mathrm{GL}(n)). \tag{4.147}$$

Principally, the essential part of the simulation of any distribution $\eta \in \mathcal{M}(\mathrm{Pos}(n))$ could be transferred from $\mathrm{Pos}(n)$ to $\mathrm{GL}(n)$. In a final step the generated pseudorandom elements on $\mathrm{GL}(n)$ could be mapped into $\mathrm{Pos}(n)$ with q^*. Obviously, this violates our general rule which says that concrete computations and simulations should be transferred from the high-dimensional space to the low-dimensional one but not vice versa. However, this inverse approach may only be useful for a specific class of measures on $\mathrm{Pos}(n)$. Canonical candidates are Borel measures $\nu^{q_*} = f^{q_*} \cdot \lambda_{\mathrm{P}os(n)} \in \mathcal{B}(\mathrm{Pos}(n))$ with unbounded density f^{q_*} for which $f\colon \mathrm{GL}(n) \to \mathbb{R}$ is very appropriate for integration problems or simulation purposes, e.g. because f is bounded or has a high degree of symmetry.

In general it should be very costly to determine the pre-image $q_*^{-1}(B)$. In specific situations, however, this is extremely simple. For example, it is $\det(\boldsymbol{M}^t\boldsymbol{M}) = (\det \boldsymbol{M})^2$ and $\mathrm{tr}(\boldsymbol{M}^t\boldsymbol{M}) = \|\boldsymbol{M}\|_{\mathrm{F}}^2$ which implies

$$q_*^{-1}\left(\{\boldsymbol{P} \in \mathrm{Pos}(n) \mid \mathrm{tr}\boldsymbol{P} \le b\}\right) = \{\boldsymbol{M} \in \mathrm{GL}(n) \mid \|\boldsymbol{M}\|_{\mathrm{F}}^2 \le b\}. \tag{4.148}$$

Of particular practical importance are $\lambda_{\mathrm{G}L(n)}$-densities $f(\boldsymbol{M}) = \bar{f}(\|\boldsymbol{M}\|_{\mathrm{F}})$ with measurable $\bar{f}\colon \mathbb{R} \to \mathbb{R}$. As the left multiplication with an orthogonal matrix remains the lengths of the columns unchanged f is $\mathrm{O}(n)$-invariant. With regard to (4.141) we note that

$$f\left(\sqrt{\boldsymbol{P}}\right) := \bar{f}\left(\|\sqrt{\boldsymbol{P}}\|_{\mathrm{F}}\right) = \bar{f}\left(\sqrt{\mathrm{tr}(\sqrt{\boldsymbol{P}}^{\,t}\sqrt{\boldsymbol{P}})}\right) = \bar{f}\left(\sqrt{\mathrm{tr}\boldsymbol{P}}\right) \tag{4.149}$$

for $\boldsymbol{P} \in \mathrm{Pos}(n)$. In particular, we can determine $f(\sqrt{\boldsymbol{P}})$ without computing $\sqrt{\boldsymbol{P}}$ explicitly.

We identify $\mathrm{GL}(2)$ and $\mathrm{Pos}(2)$ with open subsets of $\mathbb{R}^4$ or $\mathbb{R}^3$, resp., via $\boldsymbol{M} \mapsto (m_{11}, m_{21}, m_{12}, m_{22})$ and $\boldsymbol{P} \mapsto (p_{11}, p_{12}, p_{22})$. Table 4.4 provides pairs of densities (f, f^{q^*}) for $n = 2$ where $g\colon \mathbb{R} \to \mathbb{R}$ denotes any measurable function. As $c_{\mathrm{P};2} = 4\pi$ equation (4.141) implies

$$f^{q_*}(\boldsymbol{P}) = \frac{\pi f(\sqrt{\boldsymbol{P}})}{\sqrt{\det \boldsymbol{P}}} \qquad \text{for } n = 2. \tag{4.150}$$

Table 4.4. Density pairs

$f(x_1, x_2, x_3, x_4)$	$f^{q_*}(a, b, c)$
1	$\frac{\pi}{\sqrt{ac-b^2}}$
$g(x_1^2 + x_2^2 + x_3^2 + x_4^2)$	$\frac{\pi g(a+c)}{\sqrt{ac-b^2}}$
$g(x_1x_4 - x_2x_3)$	$\frac{\pi g\left(\sqrt{ac-b^2}\right)}{\sqrt{ac-b^2}}$
$g(x_1^2 + x_2^2)$	$\frac{\pi g(a)}{\sqrt{ac-b^2}}$
$g(x_1x_3 + x_2x_4)$	$\frac{\pi g(b)}{\sqrt{ac-b^2}}$
$g(x_3^2 + x_4^2)$	$\frac{\pi g(c)}{\sqrt{ac-b^2}}$

Identifying $\mathbb{R}^4$ with Mat(2, 2) the terms $x_1x_4 - x_2x_3$ and $x_1x_3 + x_2x_4$, resp., correspond with the determinant and the scalar product of the first and second column. On the boundary of Pos(2) the denominators of the right-hand densities are constant zero. In contrast the left-hand densities are bounded if g is bounded. Although $4 = \dim(\mathrm{GL}(2)) > \dim(\mathrm{Pos}(2)) = 3$ the left-hand densities are much more suitable for integration and simulation purposes than the right-hand densities. In stochastic simulations the mapping $q_*(\cdot)$ has to be evaluated for each pseudorandom element on GL(2). As this requires only three additions and six multiplications this is of subordinate significance. We point out that the situation for $n > 2$ is quite similar. This section terminates with an integration problem.

Example 4.80. Let $I_n := \lambda_{\mathrm{P}os(n)}(\{\boldsymbol{P} \in \mathrm{Pos}(n) \mid \mathrm{tr}\boldsymbol{P} \le 1\})$.
We first note that

$$q_*^{-1}(\{\boldsymbol{P} \in \mathrm{Pos}(n) \mid \mathrm{tr}\boldsymbol{P} \le 1\}) = \{\boldsymbol{M} \in \mathrm{GL}(n) \mid \|\boldsymbol{M}\|_F^2 \le 1\}) = K_{\mathrm{GL}(n)}(1). \tag{4.151}$$

The $\lambda_{\mathrm{GL}(n)}$-density $f\colon \mathrm{GL}(n) \to \mathbb{R}$, $f(\boldsymbol{M}) := 2^n |\det \boldsymbol{M}| / c_{\mathrm{P};n}$ is O(n)-invariant. From (4.141) we obtain $f^{q^*}(\boldsymbol{P}) = 1$. Putting the pieces together we obtain from 4.74

$$\begin{aligned} I_n &= \int_{\mathrm{tr}\mathbf{P} \le 1} 1 \, \lambda_{\mathrm{P}os(n)}(\mathrm{d}\boldsymbol{P}) = \frac{2^n}{c_{\mathrm{P};n}} \int_{K_{\mathrm{GL}(n)}(1)} |\det \boldsymbol{M}| \, \lambda_{\mathrm{GL}(n)}(\mathrm{d}\boldsymbol{M}) \\ &= \frac{2^n 1^{n^2+n} \pi^{\frac{n^2}{2}} \Gamma\left(\frac{n+1}{2}\right)}{c_{\mathrm{P};n} \, \Gamma\left(\frac{n^2+n+2}{2}\right) \Gamma\left(\frac{1}{2}\right)}. \end{aligned} \tag{4.152}$$

In particular, $I_2 = \pi/12$. We point out that we made use of the QR-decomposition for the computation of the right-hand integral.

4.9 $\mathrm{O}(n)$-invariant Borel measures on $\mathrm{Pos}(n)$

For each real symmetric $(n \times n)$-matrix $\boldsymbol{S}$ there exists a $\boldsymbol{T} \in \mathrm{O}(n)$ such that $\boldsymbol{D} := \boldsymbol{T}^t \boldsymbol{S} \boldsymbol{T}$ is diagonal. In particular, the diagonal entries of $\boldsymbol{D}$ are the eigenvalues of $\boldsymbol{S}$ (principal axis transformation). In this section we restrict our attention to $\mathrm{Pos}(n)$, the set of all positive definite symmetric $(n \times n)$-matrices.

Clearly,

$$\Theta\colon \mathrm{O}(n) \times \mathrm{Pos}(n) \to \mathrm{Pos}(n), \qquad \Theta(\boldsymbol{T}, \boldsymbol{P}) := \boldsymbol{T}\boldsymbol{P}\boldsymbol{T}^t \tag{4.153}$$

defines an $\mathrm{O}(n)$-action on $\mathrm{Pos}(n)$. The focus of our investigations is $\mathcal{M}_{\mathrm{O}(n)}(\mathrm{Pos}(n))$, the set of all Borel measures on $\mathrm{Pos}(n)$ which are invariant under this $\mathrm{O}(n)$-action. Again, we apply our symmetry concept. We begin with some definitions.

Definition 4.81. *Let* $\mathrm{D}(n)$ *denote the vector space of all real* $(n \times n)$*-diagonal matrices and* $\mathrm{D}_+(n) \subseteq \mathrm{D}(n)$ *the subset of all diagonal matrices with positive diagonal elements. Similarly,* $\mathrm{D}_{+\geq}(n) := \{\boldsymbol{D} \in \mathrm{D}_+(n) \mid d_{11} \geq d_{22} \geq \cdots \geq d_{nn}\}$ *and* $\mathrm{D}_{+>}(n) := \{\boldsymbol{D} \in \mathrm{D}_+(n) \mid d_{11} > d_{22} > \cdots > d_{nn}\}$.

For further definitions we refer to 4.70. We recall that $\rho_1(\boldsymbol{P}), \ldots, \rho_n(\boldsymbol{P})$ denote the eigenvalues of $\boldsymbol{P} \in \mathrm{Pos}(n)$ where we always assume $\rho_1(\boldsymbol{P}) \geq \cdots \geq \rho_n(\boldsymbol{P})$. As eigenvalues are invariant under conjugation

$$\varphi_{\mathrm{D}}\colon \mathrm{O}(n) \times \mathrm{D}_+(n) \to \mathrm{Pos}(n), \qquad \varphi_{\mathrm{D}}(\boldsymbol{T}, \boldsymbol{D}) := \boldsymbol{T}\boldsymbol{D}\boldsymbol{T}^t \tag{4.154}$$

defines a smooth surjective mapping. The orthogonal group $\mathrm{O}(n)$ acts on itself by left multiplication. As $\varphi_{\mathrm{D}}(\boldsymbol{T}_1\boldsymbol{T}, \boldsymbol{D}) = \boldsymbol{T}_1\varphi_{\mathrm{D}}(\boldsymbol{T}, \boldsymbol{D})\boldsymbol{T}_1^t$ the mapping φ_{D} is equivariant with respect to the $\mathrm{O}(n)$-actions on $\mathrm{O}(n) \times \mathrm{Pos}(n)$ (given by $\boldsymbol{T}_1(\boldsymbol{T}, \boldsymbol{D}) \mapsto (\boldsymbol{T}_1\boldsymbol{T}, \boldsymbol{D})$) and $\mathrm{Pos}(n)$.

Lemma 4.82. *(i) The* 5*-tupel* $(\mathrm{O}(n), \mathrm{O}(n), \mathrm{D}_+(n), \mathrm{Pos}(n), \varphi_{\mathrm{D}})$ *has Property (*).*
(ii) $E_{\mathrm{O}(n)}(\boldsymbol{P}) = \{\boldsymbol{P}' \in \mathrm{Pos}(n) \mid \rho_j(\boldsymbol{P}') = \rho_j(\boldsymbol{P})$ *for* $1 \leq j \leq n\}$.

Proof. $\mathrm{D}_+(n)$ is an open subset of $\mathrm{D}(n)$. Equipped with the induced topology $\mathrm{D}_+(n)$ hence is a second countable locally compact space. The remaining conditions -a-, -b-, -c-, and -d- concerning $\mathrm{O}(n)$ and $\mathrm{Pos}(n)$ are also fulfilled whereas conditions -e- and -f- have already been verified above. This proves (i). As the eigenvalues are invariant under conjugation the inclusion '$\subseteq$' in (ii) is obvious. If $\boldsymbol{P}_1$ is contained in the right-hand side then there exist $\boldsymbol{T}, \boldsymbol{T}_1 \in \mathrm{O}(n)$ and $\boldsymbol{D}, \boldsymbol{D}_1 \in \mathrm{D}_+(n)$ with $\boldsymbol{D} = \boldsymbol{T}\boldsymbol{P}\boldsymbol{T}^t$ and $\boldsymbol{D}_1 = \boldsymbol{T}_1\boldsymbol{P}_1\boldsymbol{T}_1^t$.

Up to a permutation of their diagonal entries $\boldsymbol{D}$ and $\boldsymbol{D}_1$ are equal. Hence there exists a permutation matrix $\boldsymbol{U}$ such that $\boldsymbol{D}_1 = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^t$, i.e. $\boldsymbol{P}_1 = (\boldsymbol{T}_1^t\boldsymbol{U}\boldsymbol{T})\boldsymbol{P}(\boldsymbol{T}_1^t\boldsymbol{U}\boldsymbol{T})^t \in E_{\mathrm{O}(n)}(\boldsymbol{P})$ which completes the proof of (ii). □

Theorem 4.83. *Let* $\Phi\colon \mathcal{M}^{\sigma}(\mathrm{D}_+(n)) \to \mathcal{M}^{+}_{\mathrm{O}(n)}(\mathrm{Pos}(n)), \Phi(\tau) := \big(\mu_{\mathrm{O}(n)} \otimes \tau\big)^{\varphi_{\mathrm{D}}}$. *Then the following statements are valid.*
(i) $\Phi(\mathcal{M}(\mathrm{D}_+(n))) = \mathcal{M}_{\mathrm{O}(n)}(\mathrm{Pos}(n))$ *and* $\Phi^{-1}(\mathcal{M}_{\mathrm{O}(n)}(\mathrm{Pos}(n))) = \mathcal{M}(\mathrm{D}_+(n))$.
(ii) *Let* $\nu \in \mathcal{M}_{\mathrm{O}(n)}(\mathrm{Pos}(n))$ *and* $\eta = f \cdot \nu \in \mathcal{M}(\mathrm{Pos}(n))$ *where* f *is invariant under the conjugation with orthogonal matrices. Then* $\eta \in \mathcal{M}_{\mathrm{O}(n)}(\mathrm{Pos}(n))$, *and* $\Phi(\tau_\nu) = \nu$ *implies* $\Phi(f_{|\mathrm{D}_+(n)} \cdot \tau_\nu) = \eta$.
(iii) $\lambda_{\mathrm{P}os(n)} \in \mathcal{M}_{\mathrm{O}(n)}(\mathrm{Pos}(n))$.
(iv) *The subset* $\mathrm{D}_{+\geq}(n) \subseteq \mathrm{D}_+(n)$ *is a measurable section (with respect to* $\sim$*). For each* $\nu \in \mathcal{M}_{\mathrm{O}(n)}(\mathrm{Pos}(n))$ *there exists a unique* $\tau_{\nu;0} \in \Phi^{-1}(\nu)$ *with* $\tau_{\nu;0}(\mathrm{D}_+(n) \setminus \mathrm{D}_{+\geq}(n)) = 0$. *To be precise*

$$\begin{aligned} \tau_{\nu;0}(B) :&= \nu\left(\bigcup_{\mathbf{T}\in\mathrm{O}(n)} \boldsymbol{T}\,(B \cap \mathrm{D}_{+\geq}(n))\,\boldsymbol{T}^t\right) \\ &= \nu\left(E_{\mathrm{O}(n)}\,(B \cap \mathrm{D}_{+\geq}(n))\right) \qquad \text{for each } B \in \mathrm{D}_+(n). \end{aligned} \tag{4.155}$$

Moreover, this section induces the mapping

$$\psi_{\mathrm{D}}\colon \mathrm{Pos}(n) \to \mathrm{D}_{+\geq}(n), \qquad \psi_{\mathrm{D}}(\boldsymbol{P}) := \mathbf{diag}(\rho_1(\boldsymbol{P}), \ldots, \rho_n(\boldsymbol{P})). \tag{4.156}$$

The mapping ψ_{D} *is measurable.*
(v) *The mapping* ψ_{D}^m *is a sufficient statistic for all test problems* $(\mathrm{P}_{\Gamma_0}, \mathrm{P}_{\Gamma\setminus\Gamma_0})$ *with* $\mathrm{P}_\Gamma \subseteq \mathcal{M}_{\mathrm{O}(n)}(\mathrm{Pos}(n))^m$.

Proof. Suppose that $B \in \mathcal{B}(\mathrm{Pos}(n))$ is bounded. Then there exists a real number M such that the matrix components of each $\boldsymbol{P} \in B$ lie in the interval $[-M, M]$. Let $\mathbf{v} \in \mathbb{R}^n$ be an eigenvector to the eigenvalue $\rho_1(\boldsymbol{P})$ with $\|\mathbf{v}\| = 1$. Then $\rho_1(\boldsymbol{P}) = \mathbf{v}^t\boldsymbol{P}\mathbf{v} = \sum_{i,j=1}^n \mathbf{v}_i p_{ij} \mathbf{v}_j \leq Mn^2$. Consequently, the pre-image $(\varphi_D \circ \iota)^{-1}(B) \subseteq \{\boldsymbol{D} \in \mathrm{D}_+(n) \mid d_{ii} \leq Mn^2\}$ is bounded and hence relatively compact. Hence (i) follows from 4.8(vi). From 4.82(ii) the pre-image $(\varphi_{\mathrm{D}} \circ \iota)^{-1}(E_{\mathrm{O}(n)}(\boldsymbol{P}))$ equals $\{\boldsymbol{D} \in \mathrm{D}(n) \mid d_{11}, \ldots, d_{nn}$ is a permutation of $\rho_1(\boldsymbol{P}), \ldots, \rho_n(\boldsymbol{P})\}$, and hence $\mathrm{D}_{+\geq}(n)$ is a measurable section. As $\mathrm{D}_{+\geq}(n)$ is locally compact in the induced topology the remainder of (iv) follows from Theorem 4.11. As $E_T(\boldsymbol{D}) = E_{\mathrm{O}(n)}(\boldsymbol{D}) \cap \mathrm{D}_+(n)$ assertion (ii) is an immediate consequence from 4.8(vii). In particular $\mu_{(\mathrm{P}os(n))} \in \mathcal{M}_{\mathrm{O}(n)}(\mathrm{Pos}(n))$ by 4.77(ii). Applying (ii) with $\nu = \mu_{(\mathrm{P}os(n))}$ and $f(\boldsymbol{P}) := (\det \boldsymbol{P})^{(n+1)/2}/c_\mu$ proves (iii). Finally, (v) follows from (iv) and Theorem 4.14. □

At first sight the usefulness of the mapping φ_D and the invariant Borel measures $\mathcal{M}_{\mathrm{O}(n)}(\mathrm{Pos}(n))$ for applications might seem to be marginal. In fact, in this section we abstain from examples. Their real significance will become clear in the following section when we combine results from the present and

the preceding section. For this purpose the knowledge of a pre-image $\tau_\lambda \in \Phi^{-1}(\lambda_{\mathrm{P}os(n)})$ will turn out to be very useful. Unfortunately, formula (4.155) is rather unhandy for applications. Instead, we need an explicit expression for $\tau_{\lambda;0}$ which is the unique measure in $\Phi^{-1}(\lambda_{\mathrm{P}os(n)})$ which has its total mass concentrated on the section $\mathrm{D}_{+\geq}(n)$. Therefore, we follow essentially the approach from [69], Sect. 4.

Definition 4.84. *The Lebesgue measure on* $\mathrm{D}(n)$ *is given by*

$$\lambda_{\mathrm{D}(n)}\left(\{\boldsymbol{D} \in \mathrm{D}(n) \mid a_j \leq d_{jj} < b_j \text{ for } 1 \leq j \leq n\}\right) = \prod_{j \leq n}(b_j - a_j). \quad (4.157)$$

In analogy to Definition 4.70 we denote the restrictions of $\lambda_{\mathrm{D}(n)}$ *to* $\mathrm{D}_+(n)$, $\mathrm{D}_{+\geq}(n)$ *and* $\mathrm{D}_{+>}(n)$, *resp., with* $\lambda_{\mathrm{D}+(n)}$, $\lambda_{\mathrm{D}+\geq(n)}$ *and* $\lambda_{\mathrm{D}+>(n)}$, *resp.*

Lemma 4.85. (i) $\tau_{\lambda;0}(\mathrm{D}_{+\geq}(n) \setminus \mathrm{D}_{+>}(n)) = 0$.
(ii) *For each* $(\boldsymbol{T}_0, \boldsymbol{D}_0) \in \mathrm{O}(n) \times \mathrm{D}_{+>}(n)$ *there exists a neighbourhood* $W_0 \subseteq \mathrm{O}(n) \times \mathrm{D}_{+>}(n)$ *for which the restriction* $\varphi_{D|W_0}\colon W_0 \to \mathrm{Pos}(n)$ *is a diffeomorphism onto its image.*
(iii) *Let* $J := \{\boldsymbol{D}' \in \mathrm{D}(n) \mid d_{ii} \in \{1, -1\} \text{ for } i \leq n\}$ *and* $\boldsymbol{P} = \boldsymbol{T}\boldsymbol{D}\boldsymbol{T}^t$. *If* $\boldsymbol{P} \in \mathrm{Pos}(n)$ *has mutually distinct eigenvalues then*

$$\varphi_D^{-1}\left(\{\boldsymbol{P}\}\right) \cap \left(\mathrm{O}(n) \times \mathrm{D}_{+>}(n)\right) = \boldsymbol{T}J \times \{\boldsymbol{D}\}. \quad (4.158)$$

(iv) *For each* $\boldsymbol{T}_0 \in \mathrm{O}(n)$ *there exists a neighbourhood* $U_0 \subseteq \mathrm{O}(n)$ *such that the restriction* $\varphi_D\colon U_0 \times \mathrm{D}_{+>}(n) \to \mathrm{Pos}(n)$ *is a diffeomorphism onto its image* $\varphi_D(U_0 \times \mathrm{D}_{+>}(n))$. *In particular,*

$$\left(\mu_{\mathrm{O}(n)|U_0} \otimes \tau_{\lambda;0}\right)^{\varphi_{\mathrm{D}}} = 2^{-n}\lambda_{\mathrm{P}os(n)|\varphi_{\mathrm{D}}(U_0 \times \mathrm{D}_{+>}(n))}. \quad (4.159)$$

(v) *Let* (U, χ) *denote a chart on* $\mathrm{O}(n)$ *with* $\boldsymbol{T}_0 \in U \subseteq U_0$ *and let*

$$h_*\colon \mathrm{D}_{+>}(n) \to \mathbb{R}, \qquad h_*(\boldsymbol{D}) := \left|\det D(\varphi_{\mathrm{D}} \circ (\chi^{-1} \times \mathrm{id}_{\mathrm{D}_{+>}(n)}))\,(\chi(\boldsymbol{T}_0), \boldsymbol{D})\right|. \quad (4.160)$$

Then

$$(\tau_{\lambda;0})_{|\mathrm{D}_{+>}(n)} = c_* h_* \cdot \lambda_{\mathrm{D}+>(n)} \qquad \text{for a suitable constant } c_* > 0. \quad (4.161)$$

Proof. Lemma 4.82(ii) implies $\varphi_D\left(\mathrm{O}(n) \times (\mathrm{D}_{+\geq}(n) \setminus \mathrm{D}_{+>}(n))\right) = \mathrm{Pos}(n)_=$ $:= \{\boldsymbol{P} \in \mathrm{Pos}(n) \mid \boldsymbol{P} \text{ has a multiple eigenvalue}\}$. Of course, $\boldsymbol{P} \in \mathrm{Pos}(n)_=$ iff the characteristic polynomial $p(x) := \det(\boldsymbol{P} - x1_n)$ and its derivative $p'(x)$ have a common zero. This in turn is equivalent to saying that the resultant of $p(x)$ and $p'(x)$ is zero ([49], p. 203 (Corollary 8.4)). The resultant is a polynomial in the components of $\boldsymbol{P}$ and hence 4.21(v) implies $\lambda_{\mathrm{P}os(n)}(\mathrm{Pos}(n)_=) = 0$ which proves (i). In a neighbourhood of $\boldsymbol{T}_0$ we have

$$\varphi_D = C_0 \circ \varphi_D \circ (\exp_{so(n)} \times \mathrm{id}) \circ (\exp_{so(n)}^{-1} \times \mathrm{id}) \circ (L_0^{-1} \times \mathrm{id})$$

with $C_0\colon \mathrm{Pos}(n) \to \mathrm{Pos}(n)$, $C_0(\boldsymbol{P}) := \boldsymbol{T}_0\boldsymbol{P}\boldsymbol{T}_0^t$ and $L_0\colon \mathrm{O}(n) \to \mathrm{O}(n)$, $L_0(\boldsymbol{T}_1) := \boldsymbol{T}_0\boldsymbol{T}_1$. As C_0 and L_0 are bijective linear mappings it suffices to show that $|\det D(\varphi_D \circ (\exp_{so(n)} \times \mathrm{id}))(0, \boldsymbol{D})| > 0$. As ΔF is skew-symmetric

$$\begin{aligned}\varphi_D(\exp_{so(n)}(\Delta\boldsymbol{F}), \boldsymbol{D} + \Delta\boldsymbol{D}) - \varphi_D(\exp_{so(n)}(0), \boldsymbol{D})\\ = (1 + \Delta\boldsymbol{F} + O(\|\Delta\boldsymbol{F}\|^2))(\boldsymbol{D} + \Delta\boldsymbol{D})(1 - \Delta\boldsymbol{F} + O(\|\Delta\boldsymbol{F}\|^2)) - \boldsymbol{D}\\ = \Delta\boldsymbol{F}\boldsymbol{D} - \boldsymbol{D}\Delta\boldsymbol{F} + \Delta\boldsymbol{D} + O(\|\Delta\boldsymbol{F}\|^2) + O(\|\Delta\boldsymbol{F}\|\,\|\Delta\boldsymbol{D}\|).\end{aligned}$$

In particular

$$D(\varphi_D \circ (\exp_{so(n)} \times \mathrm{id}))(0, \boldsymbol{D})\,(\Delta\boldsymbol{F}, \Delta\boldsymbol{D}) = \Delta\boldsymbol{F}\boldsymbol{D} - \boldsymbol{D}\Delta\boldsymbol{F} + \Delta\boldsymbol{D}.$$

As it will be needed later we compute the determinant of this differential explicitly. Clearly, $(\Delta\boldsymbol{F}\boldsymbol{D} - \boldsymbol{D}\Delta\boldsymbol{F} + \Delta\boldsymbol{D})_{ij} = \Delta f_{ij}(d_{ii} - d_{jj})$ for $i \neq j$ and $= \Delta d_{ii}$ if $i = j$. This induces a linear map $\mathbb{R}^{n(n+1)/2} \to \mathbb{R}^{n(n+1)/2}$, $(\Delta f_{12}, \ldots, \Delta f_{n-1,n}, \Delta d_{11}, \ldots, \Delta d_{nn}) \mapsto (\Delta f_{12}(d_{11} - d_{22}), \ldots,$ $\Delta f_{n-1,n}(d_{n-1,n-1} - d_{nn}), \Delta d_{11}, \ldots, \Delta d_{nn})$. With respect to the standard basis the matrix representation is diagonal and

$$\det D(\varphi_D \circ (\exp_{so(n)} \times \mathrm{id}))(0, \boldsymbol{D}) = \prod_{i<j}(d_{ii} - d_{jj}). \tag{4.162}$$

To prove (iii) we first show that J is the isotropy group for each $\boldsymbol{D} \in \mathrm{D}_{+>}(n) \subseteq \mathrm{Pos}(n)$ under the conjugation with orthogonal matrices. The inclusion $J \leq \mathrm{O}(n)_{\mathbf{D}}$ is obvious. On the other hand $d_{11} = \sum_{k=1}^n t_{1k}^2 d_{kk} \leq \sum_{k=1}^n t_{1k}^2 d_{11} = d_{11}$ with equality iff $t_{11}^2 = 1$. By induction follows $\boldsymbol{T} \in J$. As $\mathrm{D}_{+\geq}(n)$ is a section $\boldsymbol{P} = \varphi_D(\boldsymbol{T}_1, \boldsymbol{D}_1) = \varphi_D(\boldsymbol{T}_2, \boldsymbol{D}_2)$ implies $\boldsymbol{D}_1 = \boldsymbol{D}_2$ and hence $\boldsymbol{T}_1^t\boldsymbol{T}_2 \in J$ which proves (iii). Let W_0 defined as in (ii). As $\mathrm{O}(n)$ is a topological group there exists an open neighbourhood U' of $1_n \in \mathrm{O}(n)$ with $\boldsymbol{T}_0U'\boldsymbol{S}_i \cap \boldsymbol{T}_0U'\boldsymbol{S}_j = \emptyset$ for different $\boldsymbol{S}_i, \boldsymbol{S}_j \in J$. We claim that the restriction of φ_D to $(U_0 := \boldsymbol{T}_0U') \times \mathrm{D}_{+>}(n)$ is a diffeomorphism onto its image. We already know from (ii) that the differential $D\varphi_D$ is bijective on this domain. Further, $\boldsymbol{T}_1\boldsymbol{D}_1\boldsymbol{T}_1^t = \boldsymbol{T}_2\boldsymbol{D}_2\boldsymbol{T}_2^t$ implies $\boldsymbol{D}_1 = \boldsymbol{D}_2$ and hence $\boldsymbol{T}_1\boldsymbol{T}_2^t \in J$. Moreover, if $\boldsymbol{T}_1, \boldsymbol{T}_2 \in U_0$ we conclude $\boldsymbol{T}_1 \in U_0(\boldsymbol{T}_2^t\boldsymbol{T}_1)$ which proves the first assertion of (iv). Let $B_1 \in \mathcal{B}(U_0)$ and $B_2 \in \mathcal{B}(\mathrm{D}_{+>}(n))$. By (4.158) we have $\varphi_D^{-1}(\varphi_D(B_1 \times B_2)) = \sum_{\mathbf{S}_j \in J} U_0\boldsymbol{S}_j \times B_2$, and since $\mu_{\mathrm{O}(n)}$ is right invariant

$$\mu_{\mathrm{O}(n)} \otimes \tau_{\lambda;0}(B_1 \times B_2) = 2^{-n}\lambda_{\mathrm{P}os(n)}(\varphi_D(B_1 \times B_2)).$$

As B_1 and B_2 were arbitrary this proves the remainder of (iv). As the restriction $\varphi_D \circ (\chi^{-1} \times 1_n)_{|\chi(U_0)\times \mathrm{D}+>(n)}$ is a diffeomorphism ${\mu_{\mathrm{O}(n)}}_{|U_0}{}^{\chi} \otimes \tau_{\lambda;0} = f \cdot \lambda_{n(n-1)/2} \otimes \lambda_{\mathrm{D}+>(n)}$ with a smooth positive density $f\colon \chi(U) \times \mathrm{D}_{+>}(n) \to \mathbb{R}$. As ${\mu_{\mathrm{O}(n)}}_{|U_0}{}^{\chi} \otimes \tau_{\lambda;0}$ is a product measure $f = f_1h_*$ with continuous positive densities $f_1\colon \chi(U) \to \mathbb{R}$ and $h_*\colon \mathrm{D}_{+>}(n) \to \mathbb{R}$. The transformation theorem and (4.159) finally imply $f_1(\chi(\boldsymbol{T}))h_*(\boldsymbol{D}) = 2^{-n}|\det D(\varphi_D \circ (\chi^{-1} \times \mathrm{id}))(\chi(\boldsymbol{T}), \boldsymbol{D})|$ which completes the proof of this lemma as $f_1(\chi(\boldsymbol{T}_0)) > 0$. □

Due to 4.85(i) we may neglect the difference set $\mathrm{D}_{+\geq}(n) \setminus \mathrm{D}_{+>}(n)$ for the computation of $\tau_{\lambda;0}$. Consequently, the density $h_n\colon \mathrm{D}_{+\geq}(n) \to \mathbb{R}$ in 4.86 can arbitrarily be defined on $\mathrm{D}_{+\geq}(n) \setminus \mathrm{D}_{+>}(n)$.

Theorem 4.86. *Let*

$$h_n\colon \mathrm{D}_{+\geq}(n) \to \mathbb{R}, \qquad h_n(\boldsymbol{D}) := \prod_{1\leq i<j\leq n} (d_{ii} - d_{jj})\,. \tag{4.163}$$

Then

$$\tau_{\lambda;0} = c_n h_n \cdot \lambda_{\mathrm{D}+\geq(n)} \qquad \textit{for a suitable constant } c_n > 0. \tag{4.164}$$

Proof. Theorem 4.86 is an immediate consequence of 4.85(iv),(v) with $\chi^{-1} := \exp_{so(n)}$ and $\boldsymbol{T}_0 := 1_n$ (cf. (4.162)). □

Remark 4.87. (i) For an alternate proof of Theorem 4.86 using differential forms we refer the interested reader to [81], Examples 8.1 and 8.7.
(ii) Let for the moment $\boldsymbol{U}$ denote the diagonal matrix with $u_{11} = -1$ and $u_{ii} = 1$ for $i \geq 2$. Then $\varphi_D(\boldsymbol{T}, \boldsymbol{D}) = \varphi_D(\boldsymbol{TU}, \boldsymbol{D})$ for all $(\boldsymbol{T}, \boldsymbol{D}) \in \mathrm{O}(n) \times \mathrm{D}(n)$. In particular, Theorems 4.83 and 4.86 remain valid if we replace $\mathrm{O}(n)$ by $\mathrm{SO}(n)$ and $\varphi_D\colon \mathrm{O}(n) \times \mathrm{D}(n) \to \mathrm{Pos}(n)$ by its restriction $\varphi_{D|SO(n)\times \mathrm{D}(n)}$, resp.

4.10 Biinvariant Borel Measures on $\mathrm{GL}(n)$

In Sections 4.7 and 4.8 we considered Borel measures on $\mathrm{GL}(n)$ which are invariant under the left multiplication with orthogonal matrices (one-sided $\mathrm{O}(n)$- symmetry). In the present section we investigate *biinvariant* Borel measures on $\mathrm{GL}(n)$ which by definition are invariant under both the left and the right multiplication with orthogonal matrices. Biinvariant measures are of particular importance for various applications. Possible applications range from the simplification of analytical and numerical evaluations of integrals to a priori error bounds for complex numerical operations.

Definition 4.88. *We call* $\nu \in \mathcal{M}(\mathrm{GL}(n))$ biinvariant *if* $\nu(B) = \nu(\boldsymbol{T}B) = \nu(B\boldsymbol{T})$ *for all* $(\boldsymbol{T}, B) \in \mathrm{O}(n) \times \mathcal{B}(\mathrm{GL}(n))$. *The set of all biinvariant Borel measures on* $\mathrm{GL}(n)$ *is denoted with* $\mathcal{M}_{\mathrm{bi}}(\mathrm{GL}(n))$. *Similarly, a function* $f\colon \mathrm{GL}(n) \to \mathbb{R}$ *is said to be* biinvariant *if* $f(\boldsymbol{M}) = f(\boldsymbol{TM}) = f(\boldsymbol{MT})$ *for all* $(\boldsymbol{T}, \boldsymbol{M}) \in \mathrm{O}(n) \times \mathrm{GL}(n)$, *and* $\mathcal{F}_{\mathrm{bi}}(\mathrm{GL}(n)) := \{f\colon \mathrm{GL}(n) \to \mathbb{R} \mid f$ *is biinvariant*$\}$. *The* spectral norm *on* $\mathrm{Mat}(n,n)$ *is given by* $\|\boldsymbol{M}\|_2 := \sqrt{\rho_1(\boldsymbol{M}^t\boldsymbol{M})}$ *where* $\sigma_j(\boldsymbol{M}) := \sqrt{\rho_j(\boldsymbol{M}^t\boldsymbol{M})}$ *denotes the* j^{th} singular value *of* $\boldsymbol{M}$. *(Note that* $\boldsymbol{M}^t\boldsymbol{M}$ *is symmetric, and by definition 4.70* $\rho_i(\boldsymbol{M}^t\boldsymbol{M}) \geq \rho_j(\boldsymbol{M}^t\boldsymbol{M})$ *if* $i < j$.*) In analogy to* $\mathrm{D}_{+\geq}(n)$ *we define* $\mathrm{R}^n_{+\geq} := \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_1 \geq \cdots \geq x_n > 0\}$.

With regard to the preceding sections it is near at hand to consider the set of Borel measures

$$\mathcal{M}_{\Delta} := \{(\mu_{\mathrm{O}(n)} \otimes (\mu_{\mathrm{O}(n)} \otimes \tau)^{\varphi_{\mathrm{D}}})^{\varphi_{\mathrm{P}}} \mid \tau \in \mathcal{M}(\mathrm{D}_+(n))\} \tag{4.165}$$

and to apply the respective Theorems from Sections 4.8 and 4.9. However, the definition of $\mathcal{M}_{\Delta}$ is unhandy to decide whether a particular $\nu \in \mathcal{M}(\mathrm{GL}(n))$ belongs to $\mathcal{M}_{\Delta}$. Fortunately, the elements of $\mathcal{M}_{\Delta}$ can be characterized in a more convenient fashion.

We first note that the mapping

$$\varphi_{\mathrm{b}}\colon \mathrm{O}(n) \times \mathrm{O}(n) \times \mathrm{D}_+(n) \to \mathrm{GL}(n), \qquad \varphi_{\mathrm{b}}(\boldsymbol{S}, \boldsymbol{T}, \boldsymbol{D}) := \boldsymbol{S}\boldsymbol{D}\boldsymbol{T}^t \tag{4.166}$$

is smooth. The product group $\mathrm{O}(n) \times \mathrm{O}(n)$ acts on itself and on $\mathrm{GL}(n)$, resp., via $((\boldsymbol{S}_1, \boldsymbol{T}_1), (\boldsymbol{S}, \boldsymbol{T})) \mapsto (\boldsymbol{S}_1\boldsymbol{S}, \boldsymbol{T}_1\boldsymbol{T})$ and $((\boldsymbol{S}_1, \boldsymbol{T}_1), \boldsymbol{M}) \mapsto \boldsymbol{S}_1\boldsymbol{M}\boldsymbol{T}_1^t$, resp.

Lemma 4.89. *(i) The* 5*-tupel* $(\mathrm{O}(n) \times \mathrm{O}(n), \mathrm{O}(n) \times \mathrm{O}(n), \mathrm{D}_+(n), \mathrm{GL}(n), \varphi_{\mathrm{b}})$ *has Property (*).*
(ii) $\mathcal{M}_{\mathrm{b}i}(\mathrm{GL}(n)) = \{(\mu_{\mathrm{O}(n)} \otimes \kappa)^{\varphi_{\mathrm{P}}} \mid \kappa \in \mathcal{M}_{\mathrm{O}(n)}(\mathrm{Pos}(n))\} = \mathcal{M}_{\Delta}$.
(iii) $(\mu_{\mathrm{O}(n)} \otimes \mu_{\mathrm{O}(n)} \otimes \tau)^{\varphi_b} = (\mu_{\mathrm{O}(n)} \otimes (\mu_{\mathrm{O}(n)} \otimes \tau)^{\varphi_D})^{\varphi_P}$ *for each* $\tau \in \mathcal{M}(\mathrm{D}_+(n))$.
(iv) The following statements are equivalent:

(α) $f \in \mathcal{F}_{\mathrm{b}i}(\mathrm{GL}(n))$
(β) $f(\boldsymbol{D}) = f(\boldsymbol{T}\boldsymbol{D}\boldsymbol{S})$ *for all* $\boldsymbol{S}, \boldsymbol{T} \in \mathrm{O}(n)$, $\boldsymbol{D} \in \mathrm{D}_+(n)$
(γ) There exists a function $f^*\colon \mathrm{R}^n_{+\,\geq} \to \mathbb{R}$ *with* $f(\boldsymbol{M}) = f^*(\sigma_1(\boldsymbol{M}), \ldots, \sigma_n(\boldsymbol{M}))$.

(v)

$$\iota_n\colon \mathbb{R}^n \to \mathbb{R}^n, \qquad \iota_n(x_1, \ldots, x_n) := (x_1 - x_2, x_2 - x_3, \ldots, x_{n-1} - x_n, x_n) \tag{4.167}$$

is a linear isomorphism with determinant 1*. Further,*

$$\iota_n^{-1}(y_1, \ldots, y_n) = (y_1 + \cdots + y_n, y_2 + \cdots + y_n, \ldots, y_n) \quad \textit{and} \tag{4.168}$$

$$\iota_n(\mathrm{R}^n_{+\,\geq}) = [0, \infty)^{n-1} \times (0, \infty) \tag{4.169}$$

Proof. As $\boldsymbol{S}_1\varphi_{\mathrm{b}}(\boldsymbol{S}, \boldsymbol{T}, \boldsymbol{D})\boldsymbol{T}_1^t = \boldsymbol{S}_1\boldsymbol{S}\boldsymbol{D}\boldsymbol{T}^t\boldsymbol{T}_1^t = \varphi_{\mathrm{b}}(\boldsymbol{S}_1\boldsymbol{S}, \boldsymbol{D}, \boldsymbol{T}_1\boldsymbol{T})$ the mapping φ_{b} is $\mathrm{O}(n) \times \mathrm{O}(n)$-equivariant. As $\varphi_{\mathrm{b}}(\boldsymbol{S}\boldsymbol{T}, \boldsymbol{T}, \boldsymbol{D}) = \varphi_{\mathrm{P}}(\boldsymbol{S}, \varphi_{\mathrm{D}}(\boldsymbol{T}, \boldsymbol{D}))$ the mapping φ_{b} is surjective. Equipped with the induced topology the set $\mathrm{D}_+(n)$ is a second countable locally compact space (cf. the proof of 4.82(i)). The remaining conditions in -a-, -b-, -c-, and -d- concerning $\mathrm{O}(n) \times \mathrm{O}(n)$ and $\mathrm{GL}(n)$ are also fulfilled which completes the proof of (i). The second equation of (ii) follows from 4.83(i). Suppose that $\nu \in \mathcal{M}_{\Delta}$. Then $\nu = \mu_{\mathrm{O}(n)} \otimes \kappa$ for a suitable $\kappa \in \mathcal{M}_{\mathrm{O}(n)}(\mathrm{Pos}(n))$, and

$$\nu(\boldsymbol{S}B_1B_2\boldsymbol{T}^t) = \nu(\boldsymbol{S}B_1\boldsymbol{T}^t\boldsymbol{T}B_2\boldsymbol{T}^t) = \mu_{\mathrm{O}(n)} \otimes \kappa(\boldsymbol{S}B_1\boldsymbol{T}^t \times \boldsymbol{T}B_2\boldsymbol{T}^t)$$
$$\mu_{\mathrm{O}(n)}(B_1)\kappa(B_2) = \nu(B_1B_2)$$

for all $B_1 \in \mathcal{B}(\mathrm{O}(n))$ and $B_2 \in \mathcal{B}(\mathrm{Pos}(n))$. As $\varphi_P\colon \mathrm{O}(n) \times \mathrm{Pos}(n) \to \mathrm{GL}(n)$ is in particular a homeomorphism this implies $\nu \in \mathcal{M}_{\mathrm{b}i}(\mathrm{GL}(n))$. Now assume $\nu' \in \mathcal{M}_{\mathrm{b}i}(\mathrm{GL}(n))$. By 4.78(i) there exists a $\kappa' \in \mathcal{M}(\mathrm{Pos}(n))$ with $\nu' = \mu_{\mathrm{O}(n)} \otimes \kappa'$, and

$$\begin{aligned}\kappa'(\boldsymbol{T}B_2\boldsymbol{T}^t) &= \mu_{\mathrm{O}(n)} \otimes \kappa'(\mathrm{O}(n) \times \boldsymbol{T}B_2\boldsymbol{T}^t) = \nu'(\mathrm{O}(n)\boldsymbol{T}B_2\boldsymbol{T}^t)\\ &= \nu'(\mathrm{O}(n)B_2\boldsymbol{T}^t) = \nu'(\mathrm{O}(n)B_2) = \kappa'(B_2)\end{aligned}$$

for each $B_2 \in \mathcal{B}(\mathrm{Pos}(n))$. Hence $\kappa' \in \mathcal{M}_{\mathrm{O}(n)}(\mathrm{Pos}(n))$ which completes the proof of (ii). Let for the moment $j\colon \mathrm{O}(n) \times \mathrm{O}(n) \times \mathrm{D}_+(n) \to \mathrm{O}(n) \times \mathrm{O}(n) \times \mathrm{D}_+(n)$ be given by $j(\boldsymbol{S}, \boldsymbol{T}, \boldsymbol{D}) := (\boldsymbol{ST}, \boldsymbol{T}, \boldsymbol{D})$. Note that $(\mu_{\mathrm{O}(n)} \otimes \mu_{\mathrm{O}(n)} \otimes \tau)^j = (\mu_{\mathrm{O}(n)} \otimes \mu_{\mathrm{O}(n)} \otimes \tau)$ and $\varphi_{\mathrm{b}} \circ j = \varphi_{\mathrm{P}} \circ (\mathrm{id}_{\mathrm{O}(n)} \times \varphi_{\mathrm{D}})$ which completes the proof of (iii). Assertion (iv)(α) implies (iv)(β) by the definition of $\mathcal{F}_{\mathrm{b}i}(\mathrm{GL}(n))$. As φ_{b} is surjective each $\boldsymbol{M} \in \mathrm{GL}(n)$ can be represented as a product $\boldsymbol{TDS}$ with $\boldsymbol{T}, \boldsymbol{S} \in \mathrm{O}(n)$ and $\boldsymbol{D} \in \mathrm{D}_+(n)$. Condition ($\beta$) hence implies $f(\boldsymbol{M}) = f(\boldsymbol{D}) = f^*(d_{11}, \ldots, d_{nn})$. As $\boldsymbol{D}$ is conjugate to $\sqrt{\boldsymbol{M}^t\boldsymbol{M}}$ we have $d_{ii} = \sigma_i(\boldsymbol{D}) = \sigma_i(\boldsymbol{M})$ which implies (γ). As $\sigma_i(\boldsymbol{SMT}) = \sigma_i(\boldsymbol{M})$ condition (γ) implies (α). The verification of (v) is straight-forward. □

Definition 4.90. *Let $\nu \in \mathcal{M}_{\mathrm{O}(n)}(\mathrm{GL}(n))$ and $\kappa \in \mathcal{M}_{\mathrm{O}(n)}(\mathrm{Pos}(n))$. To simplify the notation we denote the unique Borel measures $\kappa' \in \mathcal{M}(\mathrm{Pos}(n))$ and $\tau \in \mathcal{M}(\mathrm{D}_{+\geq}(n))$ satisfying $(\mu_{\mathrm{O}(n)} \otimes \kappa')^{\varphi_P} = \nu$ and $(\mu_{\mathrm{O}(n)} \otimes \tau)^{\varphi_{D|\mathrm{O}(n)\otimes \mathrm{D}_{+\geq}(n)}} = \kappa$, resp., with $m_{\mathrm{P}}(\nu)$ and $m_{\mathrm{D}}(\kappa')$, resp. The term $\boldsymbol{diag}(x_1, \ldots, x_n)$ stands for the $(n \times n)$-diagonal matrix with diagonal entries $x_1, \ldots, x_n$, and $s_D\colon \mathrm{D}_{+\geq}(n) \to \mathrm{D}_{+\geq}(n)$, $s_D(\boldsymbol{D}) := \boldsymbol{D}^2$ denotes the square mapping on $\mathrm{D}_{+\geq}(n)$.*

Theorem 4.91. *Let $\Phi\colon \mathcal{M}^\sigma(\mathrm{D}_+(n)) \to \mathcal{M}^+_{\mathrm{O}(n)\times\mathrm{O}(n)}(\mathrm{GL}(n))$, $\Phi(\tau) := \big(\mu_{\mathrm{O}(n)} \otimes \mu_{\mathrm{O}(n)} \otimes \tau\big)^{\varphi_{\mathrm{b}}}$. Then the following statements are valid:*
(i) *$\Phi(\mathcal{M}(\mathrm{D}_+(n))) = \mathcal{M}_{\mathrm{b}i}(\mathrm{GL}(n))$ and $\Phi^{-1}(\mathcal{M}_{\mathrm{b}i}(\mathrm{GL}(n))) = \mathcal{M}(\mathrm{D}_+(n))$.*
(ii) *Let $\nu \in \mathcal{M}_{\mathrm{b}i}(\mathrm{GL}(n))$ and $\eta = f \cdot \nu \in \mathcal{M}(\mathrm{GL}(n))$ with $f \in \mathcal{F}_{\mathrm{b}i}(\mathrm{GL}(n))$. Then $\eta \in \mathcal{M}_{\mathrm{b}i}(\mathrm{GL}(n))$, and $\Phi(\tau_\nu) = \nu$ implies $\Phi(f_{|\mathrm{D}_+(n)} \cdot \tau_\nu) = \eta$.*
(iii) *$\lambda_{\mathrm{G}L(n)} \in \mathcal{M}_{\mathrm{b}i}(\mathrm{GL}(n))$. Let $h_n\colon \mathrm{D}_{+\geq}(n) \to \mathbb{R}$ be defined as in Theorem 4.86. Then*

$$m_{\mathrm{D}}(m_{\mathrm{P}}(\lambda_{\mathrm{G}L(n)})) = b_n(h_n \circ s_D) \cdot \lambda_{\mathrm{D}+\geq(n)}. \tag{4.170}$$

for a suitable constant $b_n > 0$. To be precise

$$b_n = \frac{2^{n(n+3)/2}\pi^{\frac{n^2-1}{2}} \cdot \Gamma\left(\frac{n+1}{2}\right)}{\int_0^\infty \cdots \int_0^\infty \prod_{1\leq i\leq j\leq n-1} \sum_{r=i}^{j} x_r \, \mathrm{e}^{-\frac{1}{2}(x_1+2x_2+\ldots+nx_n)} \, \mathrm{d}x_1 \cdots \mathrm{d}x_n}. \tag{4.171}$$

In particular $p_n(x_1, \ldots, x_n) = \prod_{1\leq i\leq j\leq n-1} \sum_{r=i}^{j} x_r$ is a homogeneous polynomial of degree $(n^2 - n)/2$. It can be written in the form $\sum_{j_1+\cdots+j_{n-1}=(n^2-n)/2} c_{j_1,\ldots,j_{n-1}} x_1^{j_1} \cdots x_{n-1}^{j_{n-1}}$ with integer constants $c_{j_1,\ldots,j_{n-1}}$ Using this notation (4.171) simplifies to

$$b_n = \frac{2^n n! \cdot \pi^{\frac{n^2-1}{2}} \cdot \Gamma\left(\frac{n+1}{2}\right)}{\sum\limits_{j_1+\cdots+j_{n-1}=(n^2-n)/2} c_{j_1,\ldots,j_{n-1}} \prod_{r=1}^{n-1} \frac{j_r!}{r^{j_r}}}. \tag{4.172}$$

In particular $b_2 = 4\pi^2$, $b_3 = 32\pi^4$ *and* $b_4 = 64\pi^8$.
(iv) *The subset* $\mathrm{D}_{+\geq}(n) \subseteq \mathrm{D}_+(n)$ *is a measurable section (with respect to* $\sim$*). For each* $\nu \in \mathcal{M}_{\mathrm{b}i}(\mathrm{GL}(n))$ *there exists a unique* $\tau_{\nu;b} \in \Phi^{-1}(\nu)$ *with* $\tau_{\nu;b}(\mathrm{D}_+(n) \setminus \mathrm{D}_{+\geq}(n)) = 0$. *Moreover, this section induces the mapping*

$$\psi_{\mathrm{b}}\colon \mathrm{GL}(n) \to \mathrm{D}_{+\geq}(n), \qquad \psi_{\mathrm{b}}(\boldsymbol{M}) := \mathbf{diag}(\sigma_1(\boldsymbol{M}), \ldots, \sigma_n(\boldsymbol{M})). \tag{4.173}$$

The mapping ψ_{b} *is measurable.*
(v) *The mapping* ψ_{b}^m *is a sufficient statistic for all test problems* $(\mathrm{P}_{\Gamma_0}, \mathrm{P}_{\Gamma\setminus\Gamma_0})$ *with* $\mathrm{P}_\Gamma \subseteq \mathcal{M}^1_{\mathrm{b}i}(\mathrm{GL}(n))^m$.

Proof. Assertion (i) is an immediate consequence of 4.89(ii),(iii), 4.78(i) and 4.83(i). Theorem 4.8(vii) yields (ii). The first and the second assertion of (iii) follow from 4.89(iii), 4.83(ii), 4.86, 4.78(iii). (Note that $(d_{ii}+d_{jj})(d_{ii}-d_{jj}) = (d_{ii}^2 + d_{jj}^2)$.) Using (4.170) we obtain

$$\begin{aligned}
&\int_{GL(n)} \mathrm{e}^{-\frac{1}{2}\|\boldsymbol{M}\|_{\mathrm{F}}^2} |\det \boldsymbol{M}| \,\mathrm{d}\boldsymbol{M} = b_n \int_{\mathrm{D}_{+\geq}(n)} h_n(\boldsymbol{D}^2)\, \mathrm{e}^{-\frac{1}{2}\sum_{j=1}^n d_{jj}^2} \prod_{j=1}^n \mathrm{d}_{jj}\, \mathrm{d}\boldsymbol{D} \\
&= 2^{-n} b_n \int_{\mathrm{D}_{+\geq}(n)} h_n(\boldsymbol{D}) \mathrm{e}^{-\frac{1}{2}\sum_{j=1}^n d_{jj}}\, \mathrm{d}\boldsymbol{D} \\
&= 2^{-n} b_n \int_0^\infty \cdots \int_0^\infty \left(\prod_{1\leq i\leq j\leq n-1} \sum_{r=i}^j x_r \right) \mathrm{e}^{-\frac{1}{2}(x_1+2x_2+\ldots+nx_n)}\, \mathrm{d}x_n \cdots \mathrm{d}x_1.
\end{aligned}$$

Applying the transformation theorem to the square map $s_D\colon \mathrm{D}_{+\geq}(n) \to \mathrm{D}_{+\geq}(n)$, identifying $\mathrm{D}(n)$ with $\mathbb{R}^n$, and applying the linear isomorphism ι_n from Lemma 4.89(v) verifies the second and the third equation, resp. From Example 4.75 we obtain (4.171). Using the alternate representation of p_n the denominator of (4.171) equals a sum of n-fold products of one-dimensional integrals. Elementary but careful computations finally yield (4.172). Clearly, $\mathrm{D}_{+\geq}(n)$ is a measurable section as $d_{jj} = \sigma_j(\boldsymbol{D}) = \sigma_j(\boldsymbol{S}\boldsymbol{D}\boldsymbol{T}^t)$ for each $\boldsymbol{D} \in \mathrm{D}_+(n)$ and all $\boldsymbol{S}, \boldsymbol{T} \in \mathrm{O}(n)$. The second and the final assertion of (iv) follow from 4.11(ii),(iv) as $\mathrm{D}_{+\geq}(n)$ is locally compact in the induced topology, and the third is obvious. Statement (v) is an immediate consequence of (iv) and Theorem 4.14(i). □

If the integrand and the measure are biinvariant integrals on $\mathrm{Mat}(n,n)$ and $\mathrm{GL}(n)$ can be reduced to integrals on $\mathbb{R}^n$ (cf. (3.3)).

Corollary 4.92. *Assume that* $\nu \in \mathcal{M}_{\mathrm{b}i}(\mathrm{GL}(n))$. *If* $f \in \mathcal{F}_{\mathrm{b}i}(\mathrm{GL}(n))$ *is* ν*-integrable then*

$$\int_{\mathrm{GL}(n)} f(\boldsymbol{M})\,\nu(\mathrm{d}\boldsymbol{M}) = \int_{\mathrm{D}_{+\geq}(n)} f(\boldsymbol{D})\,\nu^{\psi_{\mathrm{b}}}(\mathrm{d}\boldsymbol{D}). \tag{4.174}$$

The dimension of the domain of integration shrinks from n^2 to n. Additionally, in many cases the integrand simplifies considerably. The spectral norm of $\boldsymbol{D} \in \mathrm{D}_+(n)$, for example, equals its maximal diagonal element, the spectral norm of $\boldsymbol{D}^{-1}$ is given by the multiplicative inverse of the minimal diagonal element of $\boldsymbol{D}$. On $\mathrm{GL}(n)$, $\mathrm{R}(n)$ and $\mathrm{Pos}(n)$ the respective terms are very unhandy for explicit computations.

Corollary 4.93. *Let*

$$f(\boldsymbol{M}) = \overline{f}(\|\boldsymbol{M}\|_{\mathrm{F}}, \|\boldsymbol{M}^{-1}\|_{\mathrm{F}}, \|\boldsymbol{M}\|_2, \|\boldsymbol{M}^{-1}\|_2, |\det \boldsymbol{M}|) \tag{4.175}$$

for any function $\overline{f}\colon \mathbb{R}^5 \to \mathbb{R}$. *Then* $f \in \mathcal{F}_{\mathrm{b}i}(\mathrm{GL}(n))$ *and*

$$\begin{aligned} &\int_{\mathrm{GL}(n)} f(\boldsymbol{M})\,\nu(\mathrm{d}\boldsymbol{M}) \\ &= \int_{\mathrm{D}_{+\geq}(n)} \bar{f}\left(\sqrt{\sum_{j=1}^{n} d_{jj}^2}, \sqrt{\sum_{j=1}^{n} d_{jj}^{-2}}, d_{11}, d_{nn}^{-1}, \prod_{j=1}^{n} d_{jj}\right) \nu^{\psi_{\mathrm{b}}}(\mathrm{d}\boldsymbol{D}). \end{aligned} \tag{4.176}$$

Depending on the concrete integration problem (4.176) either enables its exact solution or it at least simplifies numerical computations considerably. The benefit of Corollary 4.93 should be enourmous anyway.

Remark 4.94. (i) The image measure $\nu^{\psi_{\mathrm{b}}} = m_{\mathrm{D}} \circ m_{\mathrm{P}}(\nu)$ contains the overall information on the distribution of the singular value vector $(\sigma_1(\boldsymbol{M}), \ldots, \sigma_n(\boldsymbol{M}))$.
(ii) Due to 4.91(i) a random variable Z on $\mathrm{GL}(n)$ is biinvariant iff Z can be represented as a three-fold product XVY of independent random variables where $X, Y \sim \mu_{\mathrm{O}(n)}$ and $V \sim \tau$ for a particular $\tau \in \mathcal{M}^1(\mathrm{D}_+(n))$. We point out that the same result is shown in [26], p. 95, with a direct computation. Note that 'spherical' in [26] corresponds with 'biinvariant' in our terminology.

Remark 4.94(ii) is very useful for stochastic simulations as the simulation of any biinvariant distribution on $\mathrm{GL}(n)$ can be decomposed into three independent simulation problems on $\mathrm{O}(n)$ and $\mathrm{D}_{+\geq}(n)$. For the special case $n = 3$ the results from Section 4.5 enable a further decomposition of the simulation problem. Suppose that $\widetilde{V}_1, \widetilde{V}_2$ and $\widetilde{W}_1, \widetilde{W}_2$, resp., denote $\mu_{(\mathrm{S}^2)}$-distributed pseudorandom vectors and $h_3 \cdot \lambda_{[0,2\pi)}$-distributed pseudorandom numbers, resp. Further, $\widetilde{U}$ denotes a standard random number and $\widetilde{Y}$ a $\nu^{\psi_{\mathrm{b}}}$-distributed pseudorandom matrix. Note that $\mu_{\mathrm{O}(n)|SO(n)} = 0.5 \cdot \mu_{SO(n)}$.

Algorithm (Biinvariant distributions on GL(3))

1.) Generate independently $\widetilde{V}_1, \widetilde{V}_2, \widetilde{W}_1, \widetilde{W}_2, \widetilde{Y}$ and $\widetilde{U}$.

2.) $\widetilde{S} := \varphi_3(\widetilde{V}_1, \widetilde{W}_1)$, $\widetilde{T} := \varphi_3(\widetilde{V}_2, \widetilde{W}_2)$.
3.) If $\widetilde{U} > 0.5$ then $\widetilde{S} := \boldsymbol{diag}(-1,1,1)\widetilde{S}$.
If $\widetilde{U} \in [0.25, 0.5)$ or $\widetilde{U} \in [0.75, 1)$ then $\widetilde{T} := \boldsymbol{diag}(-1,1,1)\widetilde{T}$.
4.) $\widetilde{Z} := \widetilde{T}\widetilde{Y}\widetilde{S}^t$.

The efficiency of this algorithm clearly depends on the distribution $\nu \in \mathcal{M}_{\mathrm{b}i}(\mathrm{GL}(n))$. If, for example, the total mass of ν is concentrated on $\{\boldsymbol{M} \mid \|\boldsymbol{M}\|_2 \leq 1\}$ then the domain of simulation in $\mathrm{D}_{+\geq}(3)$ equals $\{\boldsymbol{D} \in \mathrm{D}_{+\geq}(3) \mid d_{11} \leq 1\}$. The latter is a cone which is very suitable concrete simulations.

We have already mentioned the outstanding property of biinvariant distributions, namely that $\nu^{\psi_{\mathrm{b}}} = m_D(m_P(\nu))$ contains the overall information on the distribution of the singular values of a ν-distributed random variable. In other words, ν is uniquely determined by $\nu^{\psi_{\mathrm{b}}}$. In Example 4.97 we compute the volume of the unit ball in $\mathrm{Mat}(n,n)$ with respect to the spectral norm. The convergence rate and round off errors bounds of various numerical algorithms in linear algebra depend essentially on the singular values of the input matrices. In Example 4.98 we combine Theorem 4.91 with a well-known theorem from perturbation theory. We point out that Examples 4.96 to 4.98 coincide essentially with the Examples 5.2 to 5.4 in [69].

In the following we identify $\mathrm{D}(n)$ with $\mathbb{R}^n$ and $\mathrm{D}_{+\geq}(n)$ with $\mathrm{R}^n_{+\geq}$ via $\boldsymbol{D} \mapsto (d_{11}, \ldots, d_{nn})$. The mappings $h'_n \colon \mathrm{R}^n_{+\geq} \to \mathbb{R}$ and $s_R \colon \mathrm{R}^n_{+\geq} \to \mathrm{R}^n_{+\geq}$ defined below correspond with $h_n \colon \mathrm{D}_{+\geq}(n) \to \mathbb{R}$ and $s_D \colon \mathrm{D}_{+\geq}(n) \to \mathrm{D}_{+\geq}(n)$. In particular, h'_n is a homogeneous polynomial of degree $(n^2-n)/2$.

Definition 4.95. *For the remainder of this section we define* $h'_n \colon \mathrm{R}^n_{+\geq} \to \mathbb{R}$, $h'_n(x_1, \ldots, x_n) := \prod_{1 \leq i < j \leq n}(x_i - x_j)$ *and* $s_R \colon \mathrm{R}^n_{+\geq} \to \mathrm{R}^n_{+\geq}$, $s_R(x_1, \ldots, x_n) := (x_1^2, \ldots, x_n^2)$. *For each* $f \in \mathcal{F}_{\mathrm{b}i}(\mathrm{GL}(n))$ *the function* $f_R \colon \mathrm{R}^n_{+\geq} \to \mathbb{R}$, *is defined by* $f_R(x_1, \ldots, x_n) := f(\boldsymbol{diag}(x_1, \ldots, x_n))$.

Example 4.96. The first example may be artificial. However, it demonstrates the usefulness of Theorem 4.91 for integration problems on $\mathrm{GL}(n)$ and $\mathrm{Mat}(n,n)$. Theorem 4.91 and the transformation theorem (applied to the mapping $\boldsymbol{D} \to \boldsymbol{D}^2$ on $\mathrm{D}_{+\geq}(n)$) yield

$$\int_{\mathrm{GL}(n)} \|\boldsymbol{M}^{-1}\|_{\mathrm{F}}^2 \sin\left(\frac{n+1}{\|\mathbf{M}^{-1}\|_2^2}\right) |\det \boldsymbol{M}|^3 \mathrm{e}^{-\|\mathbf{M}\|_{\mathrm{F}}^2} \mathrm{e}^{-\|\mathbf{M}\|_2^2} \,\mathrm{d}\boldsymbol{M} \tag{4.177}$$

$$= 2^{-n} b_n \int_{\mathrm{D}_{+\geq}(n)} h_n(\boldsymbol{D}) \left(\sum_{k=1}^{n} \prod_{r \neq k} d_{rr}\right) \sin((n+1)d_{nn}) \mathrm{e}^{-\left(\sum_{j=1}^{n} d_{jj} + d_{11}\right)} \,\mathrm{d}\boldsymbol{D}.$$

Identifying $\mathrm{D}_{+\geq}(n)$ with $\mathrm{R}^n_{+\geq}$ and applying the linear isomorphism ι_n from 4.89(v) yields

$$= 2^{-n} b_n \int_{[0,\infty)^n} p_n(x_1, \ldots, x_n) \sum_{k=1}^{n} \prod_{r \neq k} x_r \mathrm{e}^{-\sum_{j=1}^{n}(j+1)x_j} \sin((n+1)x_n) \,\mathrm{d}x_1 \cdots \mathrm{d}x_n$$

$$= 2^{-n} b_n \sum_{j_1+\cdots+j_{n-1}=(n^2-n)/2} c_{j_1,\ldots,j_{n-1}} I(j_1,\ldots,j_{n-1}) \tag{4.178}$$

where we used the same polynomial representation as in 4.91(iii). The term $I(j_1,\ldots,j_{n-1})$ abbreviates

$$\sum_{k=1}^{n-1} \prod_{\substack{r=1\\ r\neq k}}^{n-1} \int_0^\infty y^{j_r+1} \mathrm{e}^{-(r+1)y}\,\mathrm{d}y \int_0^\infty y^{j_k} \mathrm{e}^{-(k+1)y}\,\mathrm{d}y \int_0^\infty y^1 \mathrm{e}^{-(n+1)y} \sin((n+1)y)\,\mathrm{d}y$$
$$+ \prod_{r=1}^{n-1} \int_0^\infty y^{j_r+1} \mathrm{e}^{-(r+1)y}\,\mathrm{d}y \int_0^\infty y^0 \mathrm{e}^{-(n+1)y} \sin((n+1)y)\,\mathrm{d}y.$$

Using formulas 1.1.3.4.1 in [14] and 3.944.9 in [30] (the latter with $p = 2$ and $p = 1$, $q = n+1$ and $t = \pi/4$) we obtain the values of these five one-dimensional integrals, namely $(j_r+1)!/(r+1)^{j_r+2}$, $(j_k)!/(k+1)^{j_k+1}$ $1/2(n+1)^2$, $(j_r+1)!/(r+1)^{j_r+2}$ or $1/2(n+1)^1$, resp. Altogether, this implies

$$I(j_1,\ldots,j_{n-1}) = \frac{1}{2\left((n+1)!\right)^2} \prod_{r=1}^{n-1} \frac{(j_r+1)!}{(r+1)^{j_r}} \left(\sum_{k=1}^{n-1} \frac{k+1}{j_k+1} + (n+1) \right). \tag{4.179}$$

While it seems to be impossible to solve the left-hand integral from (4.177) 'directly' on $\mathrm{GL}(n)$ Theorem 4.91 reduces the integration problem to finitely many additions and multiplications. For $n = 2$ and $n = 3$ the integral equals $\pi^2/18$ and $61\pi^4/1728$, resp.

Example 4.97. The $\|\cdot\|_2$-unit ball in $\mathrm{Mat}(n,n)$, i.e. $\{\boldsymbol{M} \in \mathrm{Mat}(n,n) \mid \|\boldsymbol{M}\|_2 \leq 1\}$, has Lebesgue measure

$$\begin{aligned} &\int_{\mathrm{GL}(n)} 1_{\{\|\mathbf{M}\|_2 \leq 1\}}(\boldsymbol{M})\,\mathrm{d}\boldsymbol{M} = b_n \int_{\mathrm{D}_{+\geq}(n)} h_n(\boldsymbol{D}^2) 1_{\{\mathbf{D} \mid d_{11} \leq 1\}}(\boldsymbol{D})\,\mathrm{d}\boldsymbol{D} \\ &= b_n \int_0^1 \int_0^{x_1} \int_0^{x_2} \cdots \int_0^{x_{n-1}} h_n'(x_1^2,\ldots,x_n^2)\,\mathrm{d}x_n \cdots \mathrm{d}x_1 \\ &= b_n \sum_{j_1+\cdots+j_n=(n^2-n)/2} c^*_{j_1,\ldots,j_n} \frac{1}{\prod_{s=1}^n \left(2\sum_{t=s}^n j_t + (n-s+1)\right)} \end{aligned} \tag{4.180}$$

for suitable integer constants $c^*_{j_1,\ldots,j_n}$. For $n = 2$ and $n = 3$, for instance, this sum equals $2\pi^2/3$ or $8\pi^4/45$, resp., while the unit balls in $\mathrm{Mat}(2,2)$ and $\mathrm{Mat}(3,3)$ with respect to the Frobenius norm have volumes $\pi^2/2$ and $32\pi^4/945$ (4.73(ii)). We remark that in the final step we used the formula

$$\int_0^1 \int_0^{y_1} \cdots \int_0^{y_{n-1}} x_1^{k_1} \cdots x_n^{k_n}\,\mathrm{d}x_n \cdots \mathrm{d}x_1 = \frac{1}{\prod_{s=1}^n \left(\sum_{t=s}^n k_t + (n-(s-1))\right)}$$

which can be verified by induction (cf. Example 5.3 in [69]).

Example 4.98. To simplify the notation within this example we introduce the abbreviation

$$\psi_{P;1} := \mathrm{pr}_1 \circ \varphi_P^{-1} \colon \mathrm{GL}(n) \to \mathrm{O}(n), \qquad \psi_{P;1}(\boldsymbol{M}) := \boldsymbol{M}\sqrt{\boldsymbol{M}^t\boldsymbol{M}}^{-1}, \tag{4.181}$$

i.e. $\psi_{P;1}$ denotes the projection of the polar decomposition onto the first component. Further let

$$L(\boldsymbol{M}) := \sup_{\mathbf{Z}\in Mat(n,n)\setminus\{\mathbf{0}\}} \frac{\| \left(D\psi_{P;1}(\boldsymbol{M})\right)(\mathbf{Z})\|_{\mathrm{F}}}{\|\mathbf{Z}\|_{\mathrm{F}}} = \frac{2}{\sigma_{n-1}(\boldsymbol{M}) + \sigma_n(\boldsymbol{M})}. \tag{4.182}$$

The image $\psi_{P;1}(\boldsymbol{M})$ can be computed with an algorithm of Golub and Reinsch which grounds on the singular value decomposition. In this case we have the error bound

$$\|\psi_{P;1}(\boldsymbol{M}) - (\psi_{P;1}(\boldsymbol{M}))_C\|_{\mathrm{F}} \le c\varepsilon\|\boldsymbol{M}\|_2 L(\boldsymbol{M})\sqrt{n} = \frac{2c\varepsilon\sqrt{n}\sigma_1(\boldsymbol{M})}{\sigma_{n-1}(\boldsymbol{M}) + \sigma_n(\boldsymbol{M})} \tag{4.183}$$

where the index 'C' indicates the computer output. Further, c denotes a positive constant near 1, and ε stands for the relative machine precision, i.e. for the maximal relative error when storing a real number. For further details and proofs of the preceding statements we refer the interested reader to [45], pp. 491f., and [37], pp. 1164 ff. We point out that the second reference also considers various applications of the polar decomposition in applied sciences, e.g. applications in psychometrics, computations for aerospace systems and optimization problems.

Clearly, the error $\|\psi_{P;1}(\boldsymbol{M}) - (\psi_{P;1}(\boldsymbol{M}))_C\|_{\mathrm{F}}$ should be kept small since otherwise possible conclusions derived from $(\psi_{P;1}(\boldsymbol{M}))_C$ may lack of reliability. Consider, for example, a test problem $(\mathrm{P}_{\Gamma_0}, \mathrm{P}_{\Gamma\setminus\Gamma_0})$ with $\mathrm{P}_\Gamma \subseteq \mathcal{M}^1(\mathrm{GL}(n))^m$ and $\mathrm{P}_{\Gamma_0} \subseteq \mathcal{M}^1_{\mathrm{O}(n)}(\mathrm{GL}(n))^m$ and ${p_\gamma}^{\psi^m_{P;1}} \ne \mu^m_{\mathrm{O}(n)}$ for each $\gamma \in \Gamma \setminus \Gamma_0$. Depending on the concrete problem it may be reasonable to compute the values $(\psi_{P;1}(\boldsymbol{M}_1))_C, \ldots, (\psi_{P;1}(\boldsymbol{M}_m))_C$ for a given sample $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_m \in \mathrm{GL}(n)$ and to apply a statistical test on the orthogonal components with null hypothesis ${p_\gamma}^{\psi^m_{P;1}} = \mu^m_{\mathrm{O}(n)}$. Clearly, applying (4.183) to a fixed matrix $\boldsymbol{M}_j \in \mathrm{GL}(n)$ the error term $\|\psi_{P;1}(\boldsymbol{M}_j) - (\psi_{P;1}(\boldsymbol{M}_j))_C\|_{\mathrm{F}}$ can be guaranteed to be smaller than any positive bound s_0 by decreasing ε, that is, by increasing the number of digits in the floating point representation of real numbers.

This approach has two disadvantages. At first, different input matrices $\boldsymbol{M}_i \ne \boldsymbol{M}_j$ may require different machine precision, and the singular values of $\boldsymbol{M}_j$ are not known before $\psi_{P;1}(\boldsymbol{M}_j)_C$ is computed. Two unpleasant scenarios are conceivable. Maybe the computation of $\psi_{P;1}(\boldsymbol{M}_j)_C$ often has to restarted with smaller machine precision ε, or ε might have been chosen unnecessarily small. In the following we derive a formula to determine ε so that

$$\mathrm{Prob}_\nu\left(\|\psi_{P;1}(Z) - (\psi_{P;1}(Z))_C\|_{\mathrm{F}} \le s_0\right) \ge 1 - \beta \tag{4.184}$$

if Z is a biinvariant ν-distributed random variable. We recall that the image measure $\nu^{\psi_{\mathrm{b}}} = m_{\mathrm{D}}(m_{\mathrm{P}}(\nu))$ gives the distribution of the singular value vector $(\sigma_1(Z), \dots, \sigma_n(Z))$. Our goal is to determine the cumulative distribution function $G_{\nu;2c\varepsilon\sqrt{n}}(z)$ of the random variable

$$2c\varepsilon\sqrt{n}\Lambda(Z) := 2c\varepsilon\sqrt{n}\frac{\sigma_1(Z)}{\sigma_{n-1}(Z) + \sigma_n(Z)}. \tag{4.185}$$

As $G_{\nu;2c\varepsilon\sqrt{n}}(2c\varepsilon\sqrt{n}z) = G_{\nu;1}(z)$ it essentially suffices to determine the cumulative distribution function $G_{\nu;1}$ of $\Lambda(Z)$. Obviously, $G_{\nu;1}(z) = 0$ for $z < 1/2$. As $\sigma_j(\boldsymbol{M}) = \sigma_j(\psi_b(\boldsymbol{M}))$ the random variables $\Lambda(Z)$ and $\Lambda(\psi_{\mathrm{b}}(Z))$ are identically distributed. This observation can be used to transfer the computation of $G_{\nu;1}$ from $\mathrm{GL}(n)$ to $\mathrm{D}_{+\geq}(n)$. In particular, if ν has a biinvariant $\lambda_{\mathrm{G}L(n)}$-density f the term $b_n(h_n' \circ s_R)f_R$ equals the n-dimensional Lebesgue density of the singular value vector. Then $G_{\nu;1}(z) = \int_{1/2}^{z} g_\nu(t)\,\mathrm{d}t$ where g_ν stands for the Lebesgue density of $\Lambda(Z)$. this density can be determined with standard methods described below. We mention that there is a simple way to determine the cumulative distribution function $G_{\nu;1} = G_{\nu^{\psi_b};1}$ at least approximately. First, one evaluates $\Lambda(\boldsymbol{diag}(\mathbf{x}_1)), \dots, \Lambda(\boldsymbol{diag}(\mathbf{x}_N))$ on a lattice $\{\mathbf{x}_1, \dots, \mathbf{x}_N\} \subseteq \mathrm{R}^n_{+\geq}$ and stores the 'weighted' values $b_n(h_n' \circ s_R)f_R(\mathbf{x}_j)\Lambda(\boldsymbol{diag}(\mathbf{x}_j))$. These data yield an empirical cumulative distribution function which serves as an approximation of $G_{\nu;1}$. Next, we explain how the density g_ν can be computed exactly.

Within this example we define the diffeomorphism

$$\chi_n \colon (0,\infty)^n \to (0,\infty)^n, \quad \chi_n(x_1, \dots, x_n) := \left(\frac{x_1}{x_{n-1} + x_n}, x_2, \dots, x_n\right). \tag{4.186}$$

For $n \geq 3$ we obtain $\chi_n^{-1}(y_1, \dots, y_n) = (y_1(y_{n-1} + y_n), y_2, \dots, y_n)$ and $\det D\chi_n^{-1}(y_1, \dots, y_n) = y_{n-1} + y_n$. Further, $\chi_n(\mathrm{R}^n_{+\geq}) = \{(y_1, \dots, y_n) \in (0,\infty)^n \mid y_2 \geq y_3 \geq \cdots \geq y_n;\, y_1 \geq y_2/(y_{n-1} + y_n)\}$. The transformation theorem yields

$$\begin{aligned} &\int_{\mathrm{R}^n_{+\geq}} b_n f_R(\mathbf{x})(h_n' \circ s_R(\mathbf{x}))\,\lambda_n(\mathrm{d}\mathbf{x}) \\ &\quad = \int_{\chi_n\left(\mathrm{R}^n_{+\geq}\right)} b_n f_R(\mathbf{y})(h_n' \circ s_R(\mathbf{y})) \circ \chi_n^{-1}(\mathbf{y}) \left|\det D\chi_n^{-1}(\mathbf{y})\right| \,\lambda_n(\mathrm{d}\mathbf{y}), \end{aligned} \tag{4.187}$$

and the searched density g_ν is given by the one-dimensional marginal density of the right-hand integrand with respect to the first component. To be precise, for $n \geq 3$ we obtain

$$g_\nu(y_1) = b_n \int_0^\infty \int_0^{y_2} \cdots \int_0^{y_{n-2}} \int_{\mathrm{mix}}^{y_{n-1}} h_n'(z^2, y_2^2, \dots, y_n^2)(y_{n-1} + y_n) f_R(z, \dots, y_n)\,\mathrm{d}y_n \cdots \mathrm{d}y_2 \tag{4.188}$$

with $\mathrm{mix} := \min\{y_{n-1}, \max\{0, -y_{n-1} + y_2/y_1\}\}$ and $z := y_1(y_{n-1} + y_n)$. Similarly, for $n = 2$ we obtain

$$g_\nu(y_1) = \frac{4\pi^2(2y_1 - 1)}{(1 - y_1)^4} \int_0^\infty y_2^3 f_R\left(\frac{y_1 y_2}{1 - y_1}, y_2\right) \mathrm{d}y_2 \qquad \text{for } \frac{1}{2} \le y_1 \le 1. \tag{4.189}$$

Assume for the moment that $\theta\colon \mathrm{GL}(n) \to \mathrm{GL}(n)$, $\theta(\boldsymbol{M}) := \boldsymbol{M}/\theta_*(\boldsymbol{M})$ where θ_* is any measurable positive function on GL(n). Clearly, as $\sigma_j(\theta(\boldsymbol{M})) = \sigma_j(\boldsymbol{M})/\theta_*(\boldsymbol{M})$ the random variables $\Lambda(Z)$ and $\Lambda(\theta(Z))$ are identically distributed. Consequently, if $\nu_2 = {\nu_1}^\theta$ then $G_{\nu_1;1} = G_{\nu_2;1}$. Using this observation $\mathcal{M}^1_{\mathrm{b}i}(\mathrm{GL}(n))$ can be divided into equivalence classes so that the random variable $\Lambda(Z)$ is identically distributed on each equivalence class. Consequently, we can use the same machine precision ε for all members of the same equivalence class. In particular, we only have to compute one distribution function $G_{\cdot;1}$ per equivalence class.

Note that $(\boldsymbol{M}, \boldsymbol{L}) := \mathrm{tr}(\boldsymbol{M}^t \boldsymbol{L})$ defines a scalar product on $\mathrm{Mat}(n, n)$ which coincides with the standard scalar product on $\mathbb{R}^{n^2}$ if we identify $\mathrm{Mat}(n, n)$ and $\mathbb{R}^{n^2}$ via $\boldsymbol{M} \mapsto (m_{11}, \ldots, m_{nn})$. Consequently, we call $\eta \in \mathcal{M}(\mathrm{Mat}(n, n))$ strictly radially symmetric if it is invariant under orthogonal transformations and $\eta(\{\mathbf{0}\}) = \mathbf{0}$. (Example: $NI(0, 1)^n_n$ from Corollary 4.23). Clearly, $\boldsymbol{M} \mapsto \boldsymbol{M}/\|\boldsymbol{M}\|_{\mathrm{F}}$ maps all strictly radially symmetric distributions onto the normalized geometric surface measure $\mu_{(\|\mathbf{M}\|_{\mathrm{F}}=1)}$ on the unit sphere $\{\boldsymbol{M} \in \mathrm{Mat}(n, n) \mid \|\boldsymbol{M}\|_{\mathrm{F}} = 1\}$. As $\det(\boldsymbol{M}) = 0$ iff $\det(\boldsymbol{M}/\|\boldsymbol{M}\|_{\mathrm{F}}) = 0$ we conclude $\nu(\mathrm{Mat}(n, n) \setminus \mathrm{GL}(n)) = \mu_{(\|\mathbf{M}\|_{\mathrm{F}}=1)}(\mathrm{Mat}(n, n) \setminus \mathrm{GL}(n)) = NI(0, 1)^n_n(\mathrm{Mat}(n, n) \setminus \mathrm{GL}(n)) = 0$ for each strictly radially symmetric distribution $\nu \in \mathcal{M}^1(\mathrm{Mat}(n, n))$. Clearly, $\nu_{|GL(n)} \in \mathcal{M}^1_{\mathrm{b}i}(\mathrm{GL}(n))$ as $(\boldsymbol{MT}, \boldsymbol{LT}) = (\boldsymbol{TM}, \boldsymbol{TL}) = (\boldsymbol{M}, \boldsymbol{L})$ for each $\boldsymbol{T} \in \mathrm{O}(n)$. We point out that all strictly radially symmetric distributions belong to the same equivalence class. To avoid clumsy formulations we call the restriction $\nu_{|\mathrm{GL}(n)}$ strictly radially symmetric, too. For a concrete computation of $G_{\cdot;1}$ we recommend to use $\nu := (\Gamma(n^2/2 + 1)/\pi^{n^2/2}) 1_{\|\mathbf{M}\|_{\mathrm{F}} \le 1} \cdot \lambda_{\mathrm{G}L(n)}$.

The i^{th} row of Table 4.5 provides those arguments for which $G_{\nu_i;1}(\cdot)$ equals the value on top of the respective column (e.g. $G_{\nu_2;1}(2.30) = 0.96$). We have $\nu_1 = (945/32\pi^4) 1_{\|\mathbf{M}\|_{\mathrm{F}} \le 1} \cdot \lambda_{\mathrm{G}L(3)}$, $\nu_2 = (45/8\pi^4) 1_{\|\mathbf{M}\|_2 \le 1} \cdot \lambda_{\mathrm{G}L(3)}$, $\nu_3 = (\mu_{\mathrm{O}(n)} \otimes \mu_{\mathrm{O}(n)} \otimes \varepsilon_{\boldsymbol{diag}(3,2.8,2.1)})^{\varphi_{\mathrm{b}}}$. Recall that the first row is valid for all strictly radially symmetric distributions on GL(3). The normalizing constant of ν_2 has already been computed in Example 4.97.

We point out that the table values have not been determined exactly by evaluating the density g_{ν_i} but approximately as described above. We mention that $\Lambda(Z) \equiv 3/(2.8 + 2.1) = 0.61$ for ν_3. To ensure

$$\mathrm{Prob}_{\nu_j}\left(\|\psi_{P;1}(Z) - (\psi_{P;1}(Z))_C\|_{\mathrm{F}} \le s_0\right) \ge 1 - \beta \tag{4.190}$$

we have to use a machine precision $\varepsilon_{j;1-\beta} \le s_0/(2c\sqrt{3} z_{j;1-\beta})$ where $z_{j;1-\beta}$ is implicitly defined by $G_{\nu_j;1}(z_{j;1-\beta}) = 1 - \beta$. The strictly radially symmetric

Table 4.5. $n = 3$, Cumulative distribution function $G_{\nu_i;1}$

	0.95	0.96	0.97	0.98	0.99
ν_1	3.17	3.37	3.66	4.08	4.91
ν_2	2.17	2.30	2.48	2.76	3.30
ν_3	0.61	0.61	0.61	0.61	0.61

distributions require the smallest machine precision ε among the distribution classes considered in Table 4.5. For ν_1 we need three mantissa bits more than for ν_3 to ensure that $\|\psi_{P;1}(Z) - (\psi_{P;1}(Z))_C\|_{\mathrm{F}} \leq s_0$ with a probability of at least 0.99.

In order to determine an absolute value of ε instead of a value relative to a reference distribution (e.g. relative to $\nu = \nu_1$ and $1 - \beta = 0.99$) one has to estimate the constant c (which is near 1). This might be done as follows: Generate pseudorandom matrices $(\widetilde{\boldsymbol{T}_j}, \widetilde{\boldsymbol{S}_j}, \widetilde{\boldsymbol{D}_j}) \in \mathrm{O}(n) \times \mathrm{O}(n) \times \mathrm{D}_{+\geq}(n)$ and consider the inequation

$$c \geq \frac{\widetilde{d}_{j;n-1,n-1} + \widetilde{d}_{j;nn}}{2\varepsilon\sqrt{n}\,\widetilde{d}_{j;11}} \left\| \widetilde{\boldsymbol{T}_j} - (\psi_{P;1}(\widetilde{\boldsymbol{T}_j}\widetilde{\boldsymbol{S}_j}\widetilde{\boldsymbol{D}_j}\widetilde{\boldsymbol{S}_j}^t))_C \right\|_{\mathrm{F}}. \tag{4.191}$$

Table 4.6 considers the case $n = 4$. In analogy to Table 4.5 we have chosen $\nu_4 = (8!/\pi^8)1_{\|\mathbf{M}\|_{\mathrm{F}} \leq 1} \cdot \lambda_{\mathrm{G}L(4)}$ and $\nu_5 = (1575/4\pi^8)1_{\|\mathbf{M}\|_2 \leq 1} \cdot \lambda_{\mathrm{G}L(4)}$. Qualitatively, the results are similar to that for $n = 3$. Quantitatively the differences between ν_4 and ν_5 are larger than between ν_1 and ν_2. To ensure a similar error bound for $\|\psi_{P;1}(Z) - (\psi_{P;1}(Z))_C\|_{\mathrm{F}}$ a further mantissa bit is required .

Table 4.6. $n = 4$, Cumulative distribution function $G_{\nu_i;1}$

	0.95	0.96	0.97	0.98	0.99
ν_4	4.84	5.16	5.59	6.24	7.50
ν_5	3.11	3.31	3.57	3.98	4.77

Example 4.99. In Example 4.98 the distribution of the round-off error bound (4.183) for a particular algorithm from Golub and Reinsch was considered. Besides, there are a number of other numerical algorithms used in linear algebra for which the respective error bounds or the speed of convergence depend on the singular values of the input matrix. This is the case for some iterative algorithms used for the computation of the singular value decomposition, for instance, where the speed of convergence depends on the ratio of subsequent singular values (cf. [78], pp. 377). Clearly, if the random variable Z is

(biinvariantly) ν-distributed then the distribution of the $(n-1)$-dimensional vector $(\sigma_1(Z)/\sigma_2(Z), \ldots, \sigma_{n-1}(Z)/\sigma_n(Z))$ depends on Z only through $\psi_b(Z)$. If $\nu = f \cdot \lambda_{\mathrm{GL}(n)}$ with biinvariant density f then this vector has a λ_{n-1}-density $\overline{g}_\nu$ which can be derived similarily as g_ν in the preceding example.

In place of χ_n we use the diffeomorphism

$$\overline{\chi}_n \colon (0,\infty)^n \to (0,\infty)^n, \qquad \overline{\chi}_n(x_1,\ldots,x_n) := \left(\frac{x_1}{x_2}, \frac{x_2}{x_3}, \ldots, \frac{x_{n-1}}{x_n}, x_n\right). \tag{4.192}$$

By induction, one verifies easily

$$\overline{\chi}_n^{-1}(y_1,\ldots,y_n) = \left(\prod_{j=1}^n y_j, \prod_{j=2}^n y_j, \ldots, y_{n-1}y_n, y_n\right) \tag{4.193}$$

and $\det D\,\overline{\chi}_n^{-1}(y_1,\ldots,y_n) = \prod_{j=2}^n y_j^{j-1}$. Moreover,

$$\overline{\chi}_n\left(\mathrm{R}^n_{+\geq}\right) = \{(y_1,\ldots,y_n) \in (0,\infty)^n \mid y_1,\ldots,y_{n-1} \geq 1; y_n > 0\}. \tag{4.194}$$

From the transformation theorem we conclude that the searched density $\overline{g_\nu}$ equals the marginal density of $b_n(h_n' \circ s_R) f_R \circ \overline{\chi}_n^{-1} |\det D\overline{\chi}_n^{-1}|$ with respect to the first $(n-1)$ components. To be precise

$$\begin{aligned}\overline{g}_\nu(y_1,\ldots,y_{n-1}) &= b_n C_\nu(y_1,\ldots,y_{n-1}) h_n'\left(\prod_{j=1}^{n-1} y_j^2, \prod_{j=2}^{n-1} y_j^2, \ldots, y_{n-1}^2, 1\right) \prod_{j=2}^{n-1} y_j^{j-1} \\ \text{with } C_\nu(y_1,\ldots,y_{n-1}) &:= \int_0^\infty y_n^{n^2-1} f_R\left(\prod_{j=1}^n y_j, \ldots, y_n\right) \mathrm{d}y_n \end{aligned} \tag{4.195}$$

where we made use of the fact that $h_n' \circ s_R$ is a homogeneous polynomial of degree (n^2-n). As in the preceding example the biinvariant distributions fall into equialence classes so that $\overline{g}_{\nu_1} = \overline{g}_{\nu_2}$ if ν_1 and ν_2 belong to the same equivalence class. Clearly, for the evaluation of C_ν one may choose any distribution ν' which is contained in the same class as ν.

Example 4.100. For a number of numerical algorithms the error bounds depend on the Frobenius or the spectral norm. If the random variable Z is $\nu = f \cdot \lambda_{\mathrm{GL}(n)}$-distributed with biinvariant density f then

$$\begin{aligned} F_\nu(z) := \mathrm{Prob}(\|X\|_{\mathrm{F}}/\|X\|_2 \leq z) &= \int_{\mathrm{GL}(n)} 1_{\{\|\mathbf{M}\|_{\mathrm{F}}^2/\|\mathbf{M}\|_2^2 \leq z^2\}}(\boldsymbol{M}) f(\boldsymbol{M})\,\mathrm{d}\boldsymbol{M} \\ &= b_n \int_{\mathrm{D}_{+\geq}(n)} 1_{\left\{\mathbf{D} \mid \sum_{j=1}^n d_{jj}^2/d_{11}^2 \leq z^2\right\}}(\mathbf{D}) h_n(\boldsymbol{D}^2) f(\mathbf{D})\,\mathrm{d}\boldsymbol{D}. \end{aligned} \tag{4.196}$$

Remark 4.101. In the preceding examples we did only consider biinvariant distributions. Principally, the Examples 4.98 to 4.100 could be extended to random matrices on $\mathrm{GL}(n)$ which are not biinvariantly distributed as the singular values and the Frobenius norm are constant on each $\mathrm{O}(n) \times \mathrm{O}(n)$-orbit. To each $\eta \in \mathcal{M}^1(\mathrm{GL}(n))$ there exists a biinvariant $\eta^* \in \mathcal{M}^1_{bi}(\mathrm{GL}(n))$ which coincides with η on the sub-σ-algebra $\mathcal{B}_{\mathrm{O}(n)\times\mathrm{O}(n)}(\mathrm{GL}(n)) \subseteq \mathcal{B}(\mathrm{GL}(n))$ (Theorem 4.7(ii)). Recall that the cumulative distribution functions $G_{\cdot;1}$ in Example 4.98, for instance, does only depend on the distribution of the singular values of the random matrices. Consequently, $G_{\eta;1} = G_{\eta^*;1}$. However, the precomputations needed for an explicit representation of η^* may be very costly. If $\nu \in \mathcal{M}_{\mathrm{bi}}(\mathrm{GL}(n))$ and $\eta = f \cdot \nu \in \mathcal{M}(\mathrm{GL}(n))$ then clearly $f^*(\boldsymbol{M}) = f^*(\boldsymbol{D}_{\mathbf{M}})$ for $\boldsymbol{D}_{\mathbf{M}} := \boldsymbol{diag}(\sigma_1(\boldsymbol{M}), \ldots, \sigma_n(\boldsymbol{M}))$ as $\boldsymbol{M}$ and $\boldsymbol{D}_{\mathbf{M}}$ lie on the same $\mathrm{O}(n) \times \mathrm{O}(n)$-orbit.

4.11 Symmetries on Finite Spaces

The final section deals with symmetries on finite spaces. The terms M and G denote a finite set and a finite group, resp. As usually,

$$\Theta\colon G \times M \to M, \qquad \Theta(g, m) := gm \tag{4.197}$$

denotes a group action.

Definition 4.102. *It denotes $R \subseteq M$ a section with respect to the G-orbits on M, i.e. R contains exactly one element of each G-orbit, and the mapping $\varphi\colon G \times R \to M$ is given by $\varphi(g, r) := gr$ (group action). The mapping $\psi\colon M \to R$ is a selection, i.e. it maps $m \in M$ onto the unique element in $Gm \cap R$. A measure $\eta \in \mathcal{M}(M)$ is said to be* equidistributed *if $\eta(\{m\}) = |M|^{-1}$. It is equidistributed on a non-empty subset $M_1 \subseteq M$ if $\eta(\{m_1'\}) = \eta(\{m_1''\})$ for all $m_1', m_2'' \in M_1$.*

Note that the Haar probability measure μ_G on G and the unique G-invariant probability measure $\mu_{(S)}$ on a homogeneous G-space S equal the equidistribution on G or S, resp., i.e. $\mu_G(\{g\}) = 1/|G|$ and $\mu_{(S)}(\{s\}) = 1/|S|$. The group G acts onto itself by left multiplication and hence in particular is a homogeneous G-space.

Lemma 4.103. *(i) $|Gm| = |G|/|G_m|$.*
(ii) The 5-tupel (G, G, R, M, φ) has Property $(*)$.

Proof. Statement (i) is shown in [49], p. 28 (Proposition 5.1). From the definition of φ and R we immediately obtain $g_1\varphi(g, m) = g_1gm = \varphi(g_1g, m)$ and $M = \sum_{r\in R} \varphi(G \times \{r\})$. As G, R and M are finite the remaining conditions of Property $(*)$ are trivially fulfilled if we equip all spaces with the discrete topology. □

Theorem 4.104 is an analogon to the respective theorems in the preceding sections. Clearly, as G and M are finite also a 'direct' proof which does not exploit the results from Chapter 2 is easy to give.

Theorem 4.104. *Let $\Phi\colon \mathcal{M}^1(R) \to \mathcal{M}^1(M)$, $\Phi(\tau) := (\mu_G \otimes \tau)^\varphi$. Then the following statements are valid.*
(i) *$\Phi(\mathcal{M}^1(R)) = \mathcal{M}^1_G(M)$. In particular, Φ is bijective.*
(ii) *The mapping ψ^m is a sufficient statistic for all test problems $(\mathrm{P}_{\Gamma_0}, \mathrm{P}_{\Gamma\setminus\Gamma_0})$ with $\mathrm{P}_\Gamma \subseteq \mathcal{M}^1_G(M)^m$.*

Proof. The first statement of (i) follows from 4.9(i), and the second from the fact that $R \subseteq M$ is a section. Assertion (ii) is an immediate consequence of 4.14(i) as $\varphi(s_0, \psi(m)) \in Gm$. □

We abstain from statistical and integration problems in this section since integrals over finite sets are no more than finite sums. We merely point out that

$$\sum_{m\in M} f(m)\nu(m) = \sum_{r\in R} f(r)\tau_\nu(m) \qquad \text{for } \nu \in \mathcal{M}^1_G(M) \tag{4.198}$$

if $f\colon M \to \mathbb{R}$ is G-invariant. Then $\tau_\nu(r) = \sum_{m\in Gr} \nu(m)$. Instead, we apply Theorem 4.104 for the efficient simulation of G-invariant distributions on M.

A near-at-hand method to simulate a distribution $\eta \in \mathcal{M}^1(M)$ which is not the equidistribution is to enumerate the elements of M and to initialize a table which contains the cumulative probabilities $F(j) := \eta(m_1)+\cdots+\eta(m_j)$ for $j = 1, \ldots, |M|$. To a standard random number $\widetilde{U}$ one assigns that $m_j \in M$ with $F(j-1) < \widetilde{U} \le F(j)$ (inversion method, [18], p. 85). However, this table requires memory, its initialization computation time, and for each pseudorandom element a table-look-up is necessary. We point out that there exist more sophisticated simulation algorithms (as the alias method described in [18], p. 107) which save the table search time. However, also these methods need memory and the initialization of the table costs time.

In contrast, for the simulation of the equidistribution no table is necessary. It suffices to determine an injective mapping $\chi\colon M \to \mathbb{Z}$ which transfers the essential part of the simulation from M to $\mathbb{Z}$. To be precise, one generates equidistributed pseudorandom numbers $\widetilde{V}_1, \widetilde{V}_2, \ldots$ on an appropriate finite subset $S \subseteq \mathbb{Z}$ which contains $\chi(M)$. If $\widetilde{V}_j \in \chi(M)$ then $\chi^{-1}(\widetilde{V}_j)$ is a pseudorandom element on M. If not, $\widetilde{V}_j$ will be rejected. Roughly speaking, the efficiency of this approach depends on three criteria. At first, the ratio $|M|/|S| \le 1$ should be possibly large so that not too many pseudorandom numbers are rejected in the first step. Moreover, and these criteria are much more important than the first, it should be easy to decide whether $\widetilde{V}_j \in \chi(M)$ and to compute the inverse $\chi^{-1}(\widetilde{V}_j)$ in this case.

As for continuous spaces Theorem 4.104 can be applied to decompose the simulation of $\nu \in \mathcal{M}^1_G(M)$ into two independent simulation problems on

G and R. To be precise, one has to simulate the equidistribution μ_G on G and τ_ν on R. In the continuous case the main advantage of this decomposition usually grounds on the fact that $\dim S, \dim T \leq \dim H$ (for the 5-tuple (G, S, T, H, φ)). Additionally, in most cases the spaces S and T are more suitable for concrete computations than H, and due to its outstanding symmetry the invariant measure $\mu_{(S)}$ can efficiently be simulated (see e.g. Section 4.5). In the finite case the cardinality of G is of subordinate meaning. For the simulation of μ_G the properties of the injective mapping $\chi: G \to S \subseteq \mathbb{Z}$, introduced above (with M instead of G) are much more important. The second criterion is, of course, the cardinality of R. Recall that we do not have to determine the G-orbits on M explicitly. If τ_ν is known even their cardinalities are not relevant.

Remark 4.105. If $|M|$ is moderate the cardinality of particular orbits can be estimated using the birthday's paradoxon. In fact, the random variables $X_1 m, X_2 m, \ldots$ are iid equidistributed on the orbit Gm if $X_1, X_2, \ldots$ are iid μ_G-distributed. This fact might be exploited as follows: Generate pseudorandom elements $\widetilde{X}_1 m, \widetilde{X}_2 m, \ldots$ until the first value $m' \in Gm$ is assumed for the second time. The expected number of pseudorandom elements until this happens is about $1.2\sqrt{|Gm|}$.

Algorithm (Simulation of $\nu \in \mathcal{M}_G^1(M)$)

1. Generate a τ_ν-distributed pseudorandom element $\widetilde{Y}$ on R.
2. Generate an equidistributed pseudorandom number $\widetilde{V}$ on $S \subseteq \mathbb{Z}$.
3. If $\widetilde{V} \notin \chi(M)$ goto Step 2.
4. $\widetilde{Z} := \varphi(\chi^{-1}(\widetilde{V}), \widetilde{Y})$.

We close this section with two examples. Lemma 4.107 provides two auxiliary results which will be needed in the first example.

Definition 4.106. *For the remainder of this section F denotes a finite field, $F^* := F \setminus \{0\}$, and* $\mathrm{GL}(n; F)$ *stands for the subgroup of all invertible $(n \times n)$-matrices over F. Let G' be a group and $g' \in G'$. Then $\langle g' \rangle$ denotes the group which is generated by g'. As usually,* $\deg(h)$ *stands for the degree of a polynomial h.*

Lemma 4.107. *(i) (Jordan normal form) Each $(n \times n)$-matrix $\boldsymbol{M}$ over F is conjugate to a block diagonal matrix $\boldsymbol{J}_{\mathbf{M}}$ (Jordan matrix) whose diagonal blocks are of the following type ('box matrices'):*

$$\boldsymbol{F}_{(a_0,\ldots,a_{k-1})} := \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & & & -a_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & & 1 & 0 & \\ 0 & \ldots & 0 & 1 & -a_{k-1} \end{pmatrix} \tag{4.199}$$

The polynomial $p(X) := X^k + a_{k-1}X^{k-1} + \cdots + a_0$ *is either irreducible or a power of an irreducible polynomial over* F*. Apart from their order the diagonal blocks are unique.*
(ii) $|\mathrm{GL}(n;F)| = \prod_{j=0}^{n-1} \left(|F|^n - |F|^j\right)$.

Proof. Statement (i) follows immediately from Theorem 35.6 and the respective remark after 35.8 in [48]. (We point out that the blocks in [48] are of type $\boldsymbol{F}^t_{(a_0,\ldots,a_{k-1})}$ which, however, is irrelevant.) Statement (ii) stems from [49], p. 546 (No. 15). □

Example 4.108. Clearly,

$$\Theta\colon \mathrm{GL}(n;F) \times \mathrm{GL}(n;F) \to \mathrm{GL}(n;F), \qquad \Theta(\boldsymbol{A},\boldsymbol{M}) := \boldsymbol{A}\boldsymbol{M}\boldsymbol{A}^{-1}. \quad (4.200)$$

defines a group action of $\mathrm{GL}(n;F)$ onto itself. Of course, the probability measure $\nu \in \mathcal{M}^1_{\mathrm{GL}(n;F)}(\mathrm{G}L(n;F))$ iff ν is equidistributed on each conjugacy class. To apply the simulation algorithm from above we need a representative system of the conjugacy classes on $\mathrm{GL}(n;F)$.

In a first step all irreducible polynomials over F of degree $\leq n$ have to be determined. Irreducible polynomials over prime fields $\mathrm{GF}(p) := \{0,1,\ldots,p-1\}$ (equipped with the addition and multiplication modulo p) or extension fields $\mathrm{GF}(2^e)$ can be found in appropriate tables unless p, e or n is too large. Alternatively, the irreducible polynomials can also be determined with a simple computer program. Obviously, $\det \boldsymbol{F}_{(a_0,\ldots,a_{k-1})} = (-1)^k a_0$ and $(\sum_{s=0}^t \alpha_s x^s)^j = \alpha_0^j + xA(x)$ for a polynomial $A(x)$. Further, $\det \boldsymbol{M} = \det \boldsymbol{A}\boldsymbol{J}_{\mathrm{M}}\boldsymbol{A}^{-1} = \det \boldsymbol{J}_{\mathrm{M}}$ which equals the product of the determinants of the diagonal blocks. Consequently, it suffices to consider the irreducible polynomials with constant term $a_0 \neq 0$. From the totality of box matrices one constructs a section R, i.e. a representative system of the conjugacy classes on $\mathrm{GL}(n;F)$. Any two of those block matrices are conjugate iff they consist of the same box matrices regardless of their order. Any section R consists of a maximal set of non-conjugate matrices in $\mathrm{GL}(n;F)$.

In the following we consider the special case $\mathrm{GL}(n;F) = \mathrm{GL}(6;\mathrm{GF}(2))$. Table 4.8 follows immediately from Table 4.7 and 4.107(i).

Table 4.7. Irreducible polynomials over GF(2) with constant term $a_0 = 1$

degree	1	2	3	4	5	6
number	1	1	2	3	6	9

In a final step $|R|$, i.e. the maximal number of non-conjugate block matrices in $\mathrm{GL}(6,\mathrm{GF}(2))$ has to be determined. A block matrix $\boldsymbol{J}$ is said to be

Table 4.8. Box matrices over GF(2) with determinant 1

dimension	1	2	3	4	5	6
number	1	2	3	5	7	13

Table 4.9. Number of non-conjugate block matrices in $\mathrm{GL}(6, \mathrm{GF}(2))$

type	$T(6)$	$T(5,1)$	$T(4,2)$	$T(4,1,1)$	$T(3,3)$	$T(3,2,1)$
number	13	7	10	5	6	6
type	$T(3,1,1,1)$	$T(2,2,2)$	$T(2,2,1,1)$	$T(2,1,1,1,1)$	$T(1,1,1,1,1,1)$	
number	3	4	3	2	1	

of type $T(j_1, \ldots, j_s)$ if its diagonal blocks have dimension $j_1, j_2 \ldots$ and j_s. Table 4.9 is an immedite consequence of Table 4.8.

Altogether, there exist only 60 conjugacy classes on $\mathrm{GL}(6; \mathrm{GF}(2))$. In contrast $|\mathrm{GL}(6; \mathrm{GF}(2))| = 63 \times 62 \times 60 \times 56 \times 48 \times 32 = 20\,158\,709\,760$ so that a 'naive' simulation of $\nu \neq \mu_{\mathrm{GL}(6;\mathrm{GF}(2))}$ is hardly practically feasible. A simulation of τ_ν on R, however, does not cause any problem. We point out that 4.103(i) can be applied to determine the cardinality of the particular conjugacy classes.

To generate equidistributed pseudorandom elements on $\mathrm{GL}(n; \mathrm{GF}(2))$ we use the mapping

$$\chi\colon \mathrm{Mat}(6,6;\mathrm{GF}(2)) \to \{0, \ldots, 2^{36}-1\}, \quad \chi(\boldsymbol{M}) := \sum_{1 \le i,j \le 6} m_{ij} 2^{6(i-1)+(j-1)} \tag{4.201}$$

where $\mathrm{Mat}(6,6;\mathrm{GF}(2))$ denotes the set of all (6×6)-matrices over GF(2). (The inverse mapping χ^{-1} interprets the digits of the binary representation of $v \in \{0, \ldots, 2^{36}-1\}$ as matrix entries.) Clearly, it is very easy to simulate the equidistribution on $\{0, \ldots, 2^{36}-1\}$, and $v \in \chi(\mathrm{GL}(6; \mathrm{GF}(2)))$ iff $\det \chi^{-1}(v) = 1$. Note that $|\chi(\mathrm{GL}(6; \mathrm{GF}(2)))|/2^{36} \approx 0.293$ which is very acceptable.

Example 4.109. The final example belongs to coding theory. A linear binary (n,k)-code C (i.e. its code words) is a k-dimensional subspace of $\mathrm{GF}(2)^n$. The code C is called *cyclic* if $(z_0, \ldots, z_{n-1})^t \in C$ implies $(z_{n-1}, z_0, \ldots, z_{n-2})^t \in C$. We apply our symmetry concept to determine the *weight distribution* of C, that is, to determine the number code words with Hamming weight $a \in \{0, \ldots, n\}$. These considerations could be relevant for applications in case it is not possible to compute the weight distribution with purely algebraic methods. For the definition and elementary properties of cyclic codes we refer the interested reader to [52], pp. 85 ff. We deviate from the common convention after which code words are represented as row vectors. Instead,

we represent code words as column vectors since otherwise the matrix group K defined below would not act from the left but from the right.

Assume that the left multiplication with $\boldsymbol{S} \in \mathrm{GL}(n, \mathrm{GF}(2))$ induces a cyclic shift by one position. To be precise,

$$\boldsymbol{S}(z_0, \ldots, z_{n-1})^t = (z_{n-1}, z_0, \ldots, z_{n-2})^t \quad \text{for } z_0, \ldots, z_{n-1} \in \mathrm{GF}(2)^n. \tag{4.202}$$

The finite group $K := \{\boldsymbol{S}^j \mid 0 \leq j \leq n-1\}$ acts on the vector space $\mathrm{GF}(2)^n$ by left multiplication, and by the definition of a cyclic code the subspace $C \leq \mathrm{GF}(2)^n$ is K-invariant. Clearly, the Hamming weight is constant on each K-orbit. Hence it is sufficient to consider a section $R \subseteq C$ with respect to the K-orbits and to count the Hamming weight of each $\mathbf{r} \in R$ with the cardinality of its orbit. The fundamental problem is to determine a section and the cardinalities of all K-orbits in C efficiently.

At first, we need a matrix representation $\boldsymbol{A}_{\mathbf{S}}$ of the restriction $\boldsymbol{S}_{|C}: C \to C \leq \mathrm{GF}(2)^n$ with respect to a suitable basis $\mathbf{w_1}, \ldots, \mathbf{w_k}$ of C. The generator polynomial $g(x) := g_0 + g_1 x + \ldots + g_{n-k} x^{n-k}$ of the cyclic code C divides x^n+1, i.e. there exists a polynomial $h(x) := h_0 + h_1 x + \ldots + h_k x^k$ with $x^n+1 = g(x)h(x)$ ([52], p. 89). Note that C can be identified with the set of polynomials $\{a(x)g(x) \mid \deg(a) < k\}$. In particular, $\mathbf{w_1} := (g_0, \ldots, g_{n-k}, 0, \ldots, 0)^t$, $\mathbf{w_2} := (0, g_0, \ldots, g_{n-k}, 0, \ldots, 0)^t, \ldots,$ $\mathbf{w_k} := (0, \ldots, 0, g_0, \ldots, g_{n-k})^t$ is a basis of the subspace C. From $x^k g(x) - (x^n + 1) = g(x)(x^k - h(x)) = g(x)(h_0 + h_1 x + \ldots + h_{k-1} x^{k-1})$ we obtain the $(k \times k)$-matrix

$$\boldsymbol{A}_{\mathbf{S}} := \begin{pmatrix} 0 & 0 & \cdots & 0 & h_0 \\ 1 & 0 & & & h_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & & 1 & 0 & \\ 0 & \ldots & 0 & 1 & h_{k-1} \end{pmatrix}. \tag{4.203}$$

By 4.107(ii) $\boldsymbol{A}_{\mathbf{S}}$ is conjugate to a Jordan matrix $\boldsymbol{J}_{\mathbf{A}}$, i.e. $\boldsymbol{A}_{\mathbf{S}} = \boldsymbol{B}\boldsymbol{J}_{\mathbf{A}}\boldsymbol{B}^{-1}$ for a suitable $\boldsymbol{B} \in \mathrm{GL}(k, \mathrm{GF}(2))$. Clearly, $\boldsymbol{A}_{\mathbf{S}}^j = \boldsymbol{B}\boldsymbol{J}_{\mathbf{A}}^j\boldsymbol{B}^{-1}$ for each $j \leq n-1$, and hence $\Theta(\boldsymbol{S}^j, \mathbf{x}) := \boldsymbol{J}_{\mathbf{A}}^j \mathbf{x}$ defines an action of the cyclic group K on $\mathrm{GF}(2)^k$. Assume that the columns of the $(n \times k)$-matrix $\boldsymbol{W}$ are given by the basis vectors $\mathbf{w_1}, \ldots, \mathbf{w_k} \in C$. From the construction of $\boldsymbol{A}_{\mathbf{S}}$ we obtain

$$\boldsymbol{S}(\boldsymbol{W}\boldsymbol{B})(\boldsymbol{B}^{-1}\mathbf{x}) = \boldsymbol{S}\boldsymbol{W}\mathbf{x} = \boldsymbol{W}\boldsymbol{A}_{\mathbf{S}}\mathbf{x} = (\boldsymbol{W}\boldsymbol{B})\boldsymbol{J}_{\mathbf{A}}(\boldsymbol{B}^{-1}\mathbf{x}) \tag{4.204}$$

for each $\mathbf{x} \in \mathrm{GF}(2)^k$. By induction we conclude $\boldsymbol{S}^j(\boldsymbol{W}\boldsymbol{B}) = (\boldsymbol{W}\boldsymbol{B})\boldsymbol{J}_{\mathbf{A}}^j$. Hence $\boldsymbol{W}\boldsymbol{B}: \mathrm{GF}(2)^k \to C$ is a K-equivariant vector space isomorphism. In particular, $\boldsymbol{W}\boldsymbol{B}$ maps K-orbits bijectively onto K-orbits. Consequently, the cardinality of the orbits can be determined on $\mathrm{GF}(2)^k$. Applying the linear isomorphism $\boldsymbol{W}\boldsymbol{B}$ yields the respective Hamming weight.

Of particular relevance for applications are code lengths of type $n = 2^s - 1$. In the field $\mathrm{GF}(2^s)$ the polynomial $x^{2^s-1} + 1$ falls into $2^s - 1$ mutually dis-

tinct linear factors. Consequently, $x^{2^s-1}+1$ is a product of mutually distinct irreducible factors over GF(2). (To be precise, $x^{2^s-1}+1$ is the product of all irreducible polynomials $\neq x$ whose degree divides s ([49], p. 254 (Nr. 22)). The matrix $\boldsymbol{A}_{\mathrm{S}}$ has the characteristic polynomial $h(x)$. In particular, $h(x) = h_1(x)\cdots h_t(x)$ with mutually distinct irreducible polynomials $h_1,\ldots,h_t$ and hence $h(x)$ is also the minimal polynomial of $\boldsymbol{A}_{\mathbf{S}}$. This yields the Jordan matrix $\boldsymbol{J}_{\mathbf{A}}$ without further computations. The matrix $\boldsymbol{B} \in GL(2)^k$ may be computed with an algorithm given in [48], pp. 263 ff. Alternatively, computer algebra systems may be used. Note that $\boldsymbol{B}$ only has to be computed once.

We identify the vector space $\mathrm{GF}(2)^k$ with the cartesian product $\prod_{j=1}^{t}\mathrm{GF}(2)^{\deg(h_j)}$ where the order of the particular factors fits to the block structure of $\boldsymbol{J}_{\mathbf{A}}$. We divide $\prod_{j=1}^{t}\mathrm{GF}(2)^{\deg(h_j)}$ into the $(k+1)$ disjoint subsets

$$M_1 := \left(\mathrm{GF}(2)^{\deg(h_1)}\right)^* \times \prod_{j=2}^{t}\mathrm{GF}(2)^{\deg(h_j)} \tag{4.205}$$

$$M_2 := \{\mathbf{0}\} \times \left(\mathrm{GF}(2)^{\deg(h_2)}\right)^* \times \prod_{j=3}^{t}\mathrm{GF}(2)^{\deg(h_j)},\ldots,$$

$$M_k := \{\mathbf{0}\}\times\cdots\times\{\mathbf{0}\}\times\left(\mathrm{GF}(2)^{\deg(h_k)}\right)^*, \qquad M_{k+1} := \{(\mathbf{0},\ldots,\mathbf{0})\}.$$

As $\boldsymbol{J}_{\mathbf{A}}$ is invertible $\boldsymbol{J}_{\mathbf{A};\mathbf{j}}$, the j^{th} diagonal block of $\boldsymbol{J}_{\mathbf{A}}$, is also invertible, and hence the sets $M_1,\ldots,M_{k+1}$ are K-invariant. As $K \to \mathrm{GL}(k;\mathrm{GF}(2))$, $\boldsymbol{S} \mapsto \boldsymbol{J}_{\mathbf{A}}$, is a group homomorphism the group order $|\langle \boldsymbol{J}_{\mathbf{A}}\rangle|$ divides $|\langle \boldsymbol{S}\rangle| = n$.

Now assume that $n = 2^s - 1$ is a prime (which is the case if $s \in \{5,7\}$, for instance). In particular, $\langle \boldsymbol{S}\rangle$ has prime order n, and this facilitates the determination of a section $R \subseteq C$ and the computation of the orbit lengths considerably. If $\mathbf{w} \in \{(0,\ldots,0),(1,\ldots,1)\}$ we have $|K\mathbf{w}| = 1$ while $|K\mathbf{w}| = |K| = n$ else since n is prime. Apart from the special case that all K-orbits in C are singleton (i.e. $C = \{(0,\ldots,0),(1,\ldots,1)\}$) we conclude $|\langle \boldsymbol{J}_{\mathbf{A}}\rangle| = n$. Similarly, as $\boldsymbol{J}_{\mathbf{A}} \mapsto \boldsymbol{J}_{\mathbf{A};\mathbf{j}}$ induces a group homomorphism and $\boldsymbol{W}$ has rank k also $|\langle \boldsymbol{J}_{\mathbf{A};\mathbf{j}}\rangle| = n$ if $\deg(h_j) > 1$. Apart from the linear factor $(x-1)$ the polynomial $x^{2^s-1}+1$ falls into irreducible polynomials of degree s. Hence $\deg(h_j) = s$ and $\langle \boldsymbol{J}_{\mathbf{A};\mathbf{j}}\rangle = n$. Note that $\langle \boldsymbol{J}_{\mathbf{A};\mathbf{j}}\rangle$ acts on $\mathrm{GF}(2)^{\deg(h_j)}$ which falls into the orbits $\{\mathbf{0}\}$ and $\left(\mathrm{GF}(2)^{\deg(h_j)}\right)^*$. If $\deg(h_j) = 1$ then $|\langle \boldsymbol{J}_{\mathbf{A};\mathbf{j}}\rangle| = 1$. (Note, however, that the K-action on $\mathrm{GF}(2)^k$ is not isomorphic to the action of the product group $\langle \boldsymbol{J}_{\mathbf{A};\mathbf{1}}\rangle \times\cdots\times \langle \boldsymbol{J}_{\mathbf{A};\mathbf{t}}\rangle$ on $\prod_{j=1}^{t}\mathrm{GF}(2)^{\deg(h_j)}$.) Let $\mathbf{e}_{\mathbf{1};\mathbf{j}} = (1,0\ldots,0) \in \mathrm{GF}(2)^{\deg(h_j)}$ and

$$R_1 := \{\mathbf{e}_{\mathbf{1};\mathbf{1}}\} \times \prod_{j=2}^{t}\mathrm{GF}(2)^{\deg(h_j)}, \tag{4.206}$$

$$R_2 := \{\mathbf{0}\} \times \{\mathbf{e_{1;2}}\} \times \prod_{i=3}^{t} \mathrm{GF}(2)^{\deg(h_j)}, \ldots,$$
$$R_k := \{\mathbf{0}\} \times \cdots \times \widetilde{R}, \qquad R_{k+1} := \{(\mathbf{0}, \ldots, \mathbf{0})\}$$

with $\widetilde{R}_k = \{e_{1;k}\}$ if $\deg(h_k) > 1$ and $\widetilde{R}_k = (\mathrm{GF}(2)^{\deg h_k})^*$ else. It is an immediate consequence of the preceding that $R := \sum_{j=1}^{k+1} R_j$ is a section. To compute the weight distribution of C one determines the Hamming weight of $(\boldsymbol{WB})\mathbf{r}$ for each $\mathbf{r} \in R$ and counts it with the multiplicity $|K\mathbf{r}|$. (Recall that $|K\mathbf{r}| = n = 2^s - 1$ for $\mathbf{r} \in R_j$ and $j < k$.)

The 'naive' approach is to determine the Hamming weight code word by code word which requires 2^k multiplications of the generator polynomial $g(x)$ with polynomials of degree $< k$ and the computation of 2^k many n-fold integer sums with $\{0, 1\}$-valued summands. Neglecting the single computation of $\boldsymbol{B}$ the symmmetry-based approach requires about $2^k/n$ matrix-vector multiplications with $(\boldsymbol{WB})$ (i.e. about 2^k many binary 'scalar products' of k-tuples) and the computation of $2^k/n$ many n-fold integer sums with $\{0, 1\}$-valued summands. Obviously, our approach is more efficient than the naive approach. Moreover, the specific form of the sets R_j can possibly be used for a more efficient combinatorial solution.

Note that the sets M_j are also disjoint and K-invariant if $n = 2^s - 1$ but n is not prime. If n is not prime, however, $1 < |\langle \boldsymbol{J_A} \rangle| < n$ and $1 < |\langle \boldsymbol{J_{A;j}} \rangle| < |\langle \boldsymbol{J_A} \rangle|$ is possible. Clearly, this complicates the determination of a section. Then the time needed to determine the weight distribution depends essentially on the time needed for the subset M_1. It hence may be recommendable to relabel the minimal polynomials $h_1, \ldots, h_k$ so that $\deg(h_1) = s$ and $|\langle \boldsymbol{J_{A;1}} \rangle| = 2^s - 1$ (provided that a polynomial h_j with these properties exists).

Of course, a similar approach is also possible for $n \neq 2^s - 1$. We point out that additional difficulties may occur. If any irreducible factor of $h(x)$ occurs twice it costs additional efforts to compute the Jordan matrix $\boldsymbol{J_A}$. Above all, the degrees of the irreducible factors may be very large which makes it more difficult to determine a section R. Over GF(2) the polynomial $x^{83} + 1$, for instance, falls into the two irreducible factors $(x+1)$ and $(x^{82} + x^{81} + \ldots + 1)$.

In particular cases the code C may be invariant even under the action of a supergroup K' of K. However, it will usually not be possible to transfer our symmetry-based approach from above to K'. We exploited the fact that $\mathbf{A_S}, \boldsymbol{A}_\mathbf{S}^2, \ldots, \boldsymbol{A}_\mathbf{S}^{n-1}$ can be transformed to Jordan matrices $\boldsymbol{J_A}, \boldsymbol{J}_\mathbf{A}^2, \ldots, \boldsymbol{J}_\mathbf{A}^{n-1}$ by the conjugation with the same matrix $\boldsymbol{B}$.

In this example we did not apply Theorem 4.104. We merely point out that the 5-Tupel (K, K, R, C, φ) has Property $(*)$ where $\varphi\colon K \times R \to C$ is defined by $\varphi(\boldsymbol{S}^j, \mathbf{r}) := (\boldsymbol{WB})\boldsymbol{J}_\mathbf{A}^j\mathbf{r}$.

References

1. Afflerbach, L. , Lehn, J. (eds.) (1986): Kolloquium über Zufallszahlen und Simulation. Teubner, Stuttgart
2. Anderson, S.A. (1982): Distributions of Maximal Invariants Using Quotient Measures. Ann. Stat., **10**, 955–961
3. Anderson, S.A., Brøns, H.K., Jensen, S.T. (1983): Distribution of Eigenvalues in Multivariate Statistical Analysis. Ann. Statist., **11**, 392–415
4. Arvo, J. (1992): Fast Random Matrices. In: Kirk, D. (ed.): Graphics Gems III. Academic Press, London, 117–120
5. Bauer, H. (1992): Maß- und Integrationstheorie. Zweite Auflage, De Gruyter, New York
6. Behrends, E. (1987): Maß- und Integrationstheorie. Springer, Berlin
7. Bierlein, D. (1961): Der Graph meßbarer Funktionen mit abstraktem Definitionsbereich. Math. Zeitschrift, **76**, 468–471
8. Bokowski, J., Richter, J., Schindler, W. (1992): On the Distribution of Order Types. Comput. Geom., **1**, 127–142
9. Bokowski, J., Sturmfels, B. (1989): Computational Synthetic Geometry. Springer, Lecture Notes in Mathematics, **1355**, Berlin
10. Bondar, J.V. (1976): Borel Cross Sections and Maximal Invariants. Ann. Math. Stat., **4**, 866–877
11. Bourbaki, N. (1963): Intégration (Chapitres 7 et 8). Herman, Paris
12. Bourbaki, N. (1958): Topologie Générale (Chapitre 9). Herman, Paris
13. Bröcker, T., tom Dieck, T. (1985): Representations of Compact Lie Groups. Springer, New York
14. Bronstein, I.N., Semendjajew, K.A. (1984): Taschenbuch der Mathematik. 21. Auflage, Harri Deutsch, Thun
15. Cordovil, R., Las Vergnas, M., Mandel, A. (1982): Euler's Relation, Möbius Functions and Matroid Identities. Geom. Dedicata, **12**, 147–162
16. Davidovicz, A.L. (1990): On the Lifting of Invariant Measures. Ann. Pol. Math., **51**, 137–139
17. Dellacherie, C. (1980): Un cours sur les ensembles analytiques. In: Rogers, C.A. et al. (eds): Analytic Sets. Academic Press, London, 183–316
18. Devroye, L. (1986): Non-Uniform Random Variate Generation. Springer, New York
19. Diaconis, P. (1988): Group Representations in Probability and Statistics. Institute of Mathematical Statistics, Hayward (Ca.), Lecture Notes Monograph Series **11**
20. Eaton, M.L. (1989): Group Invariance Applications in Statistics. Institute of Mathematical Statistics, Hayward (Ca)

21. Eaton, M.L., Sudderth, W.D. (1999): Group Invariant Inference and Right Haar Measures. To appear in J. Stat. Planning and Infer.
22. Eichenauer-Herrmann, J. (1993): The Lattice Structure of Nonlinear Congruential Pseudorandom Numbers. Metrika, **40**, 115–120
23. Engelking, R.D. (1989): General Topology (revised and completed). Helderman, Berlin
24. Fichtenholz, G.M. (1974): Differential- und Integralrechnung III. VEB, Berlin
25. Fang, K.T., Kotz, S., Ng, K.-W. (1989): Symmetric Multivariate and Related Distributions. Chapman & Hall, London
26. Fang, K.T., Zhang, Z.Y.-T. (1990): Generalized Multivariate Analysis. Springer, New York
27. Forster, O. (1984): Analysis 3. Dritte durchgesehene Auflage, Vieweg, Braunschweig
28. Gänssler, P., Stute, W. (1977): Wahrscheinlichkeitstheorie. Springer, Berlin
29. Giné, E. (1975): Invariant Tests for Uniformity on Compact Riemannian Manifolds Based on Sobolev Spaces. Ann. Statist., **3**, 1243–1266
30. Gradsteyn, I.S., Ryzhik, I.M. (1965): Table of Integrals, Series and Products. Fourth edition, Academic Press, New York
31. Gupta, A.K., Varga, T. (1993): Elliptically Contoured Models in Statistics. Kluwer Academic Publishers, Dordrecht
32. Gupta, A.K., Varga, T. (1997): Characterization of Matrix Variate Elliptically Contoured Distributions. In: Johnson, N.L. et al. (eds.): Advances in the Theory and Practice of Statistics. Wiley, New York
33. Halmos, P.R. (1950): Measure Theory. v. Nostrand, New York
34. Heiberger, R.M. (1978): Generation of Random Orthogonal Matrices. Appl. Stat., **27**, 199–206
35. Helgason, S. (1973): Functions on Symmetric Spaces. In: Moore, C.C. (ed.): Harmonic Analysis on Homogeneous Spaces. American Mathematical Society, Proc. Symposia Math. **26**, 101–146
36. Helland, I.S. (2001): Reduction of Regression Models under Symmetry. In: Viana, M., Richards, D. (eds): Algebraic Methods in Statistics. Contemporary Mathematics Series of the American Mathematical Society
37. Higham, N.J. (1986): Computing the Polar Decomposition – with Applications. SIAM J. Sci. Stat. Comput., **7**, 1160–1174
38. Hochschild, G. (1965): The Structure of Lie Groups. Holden Day, San Francisco
39. Hofmann, K.H., Terp, C. (1994): Compact Subgroups of Lie Groups and Locally Compact Groups. Proc. Amer. Math. Soc., **120**, 623–634
40. Humphreys, J.E. (1972): Introduction to Lie Algebras and Representation Theory. Springer, New York
41. Iwasawa, K. (1950): On Some Types of Topological Groups. Ann. Math., **50**, 507-558
42. Kasprowski, P. (1983): On the Existence of Invariant Measures for Piecewise Convex Transformations. Ann. Pol. Math., **40**, 179–184
43. Kelley, J.L. (1955): General Topology. van Nostrand, Princeton
44. Kelley, J.L., Srinivasan, T.P. (1988): Measure and Integral (Vol. 1). Springer, New York
45. Kenney, C., Laub, A.J. (1991): Polar Decomposition and Matrix Sign Function Condition Estimates. SIAM J. Sci. Stat. Comp., **12**, 488–504
46. Knuth, D.E. (1981): The Art of Computer Programming (Vol. 2). Addison-Wesley, London

47. Koecher, M. (1983): Lineare Algebra und analytische Geometrie. Springer, Berlin Heidelberg
48. Kowalsky, H.-J. (1979): Lineare Algebra. Neunte überarbeitete und erweiterte Auflage, de Gruyter, Berlin
49. Lang, S. (1993): Algebra. Third edition, Addison-Wesley, Reading
50. Las Vergnas, M. (1989): Convexity in Oriented Matroids. Journal of Comb. Geom., **A50**, 24–32
51. Lehmann, E.L. (1994): Testing Statistical Hypotheses. Second edition, Chapman & Hall, New York
52. Lin, S., Costello, D.J. (1983): Error Control Coding: Fundamentals and Applications. Prentice-Hall, Englewood Cliffs
53. Marsaglia, G. (1972): Choosing a Point from the Surface of a Sphere. Ann. Math. Stat., **43**, 645–646
54. Nachbin, L. (1965): The Haar Integral. van Nostrand, New York
55. Penconek, M., Zakrzewski, P. (1994): The Existence of Nonmeasurable Sets for Invariant Measures. Proc. Amer. Math. Soc., **121**, 579–583.
56. Pique, M.E. (1990): Rotation Tools. In: Glassner, A. (ed): Graphics Gem. Academic Press, London, 455–469
57. Pletincks, D. (1988): The Use of Quaternions for Animation, Modelling and Rendering. In: Magenat-Thalman, N., Thalman, D. (eds): New Trends in Computer Graphics. Springer, New York, 44–53
58. Promislow, D. (1983): Nonexistence of Invariant Measures. Proc. Am. Math. Soc., **88**, 89–92
59. v. Querenburg, B. (1979): Mengentheoretische Topologie. Second edition, Springer, New York
60. Reiffen, H.-J., Trapp, H.W. (1973): Einführung in die Analysis III. B.I. Hochschultaschenbuch **787**, B.I.-Verlag, Mannheim
61. Ringel, G. (1956): Teilungen der Ebene durch Geraden und topologische Geraden. Math. Z., **64**, 79–102
62. Roberts, J. (1975): Invariant Measures in Compact Hausdorff Spaces. Indiana Univ. Math. J., **24**, 691–718
63. Rudin, W. (1982): Functional Analysis. Seventh Reprinting, McGraw Hill, New Delhi
64. Schindler, W. (1991): Über das Erzeugen und Testen von Pseudozufallselementen. Dissertation (PhD Thesis), Department of Mathematics, TU Darmstadt, Germany
65. Schindler, W. (1993): Iwasawa's Theorem and Integrals on Lie Groups. Math. Nachr., **162**, 315–327
66. Schindler, W. (1994): A Sufficient Statistic for Con-invariant Test Problems. Math. Nachr., **169**, 243–265
67. Schindler, W. (1994): A Generalization of Weyl's Integration Theorem and its Meaning for Stochastic Simulations. Math. Oper. Res., **19**, 523–538
68. Schindler, W. (1994): Equivariant Mappings: A New Approach in Stochastic Simulations. Comput. Geom., **4**, 327–343
69. Schindler, W. (1995): Bi-invariant Integrals on GL(n) with Applications. Math. Nachr., **173**, 297–320
70. Schindler, W. (1997): On the Efficient Simulation of a Particular Class of Random Rotations Relevant to Computer Graphics. J. Comput. Appl. Math., **81**, 107–114

71. Schindler, W. (1998): Maße mit Symmetrieeigenschaften. Habilitationsschrift (postdoctoral thesis), Department of Mathematics, TU Darmstadt, Germany
72. Schindler, W. (2001): Efficient Online Tests for True Random Number Generators. In: Koç, Ç.K., Naccache, D., Paar, C. (eds.): Cryptographic Hardware and Embedded Systems — CHES 2001, Springer, Lecture Notes in Computer Science, **2162**, Berlin, 103–117
73. Schindler, W., Killmann, W. (2002): Evaluation Criteria for True (Physical) Random Number Generators Used in Cryptographic Applications. To appear in: Kaliski, B., Koç, Ç.K., Paar, C. (eds.): Cryptographic Hardware and Embedded Systems — CHES 2002, Springer, Lecture Notes in Computer Science, **2523**, Berlin, 431–449
74. Schreier, O., Sperner, E. (1959): Introduction to Modern Algebra and Matrix Theory. Second edition, Chelsea, New York.
75. Shoemake, K. (1985): Animating Rotations with Quaternion Curves. Computer Graphics, **19**, 245–254
76. Shoemake, K. (1992): Uniform Random Rotations. In: Kirk, D. (ed): Graphics Gems III. Academic Press, London, 124–132
77. Stewart, G.W. (1980): The Efficent Generation of Random Orthogonal Matrices with an Application to Condition Estimators. SIAM J. Numer. Anal., **17**, 403–409
78. Stoer, J., Bulirsch, R. (1980): Introduction to Numerical Analysis. Springer, New York
79. Tanner, M.A., Thisted, R.A. (1982): A Remark on AS127. Generation of Random Orthogonal Matrices. Appl. Stat., **31**, 190–192
80. Wijsman, R.A. (1986) Global Cross Sections as a Tool for Factorization of Measures and Distributions of Maximal Invariants. Sankhya Ser. A, **48**, 1–42
81. Wijsman, R.A. (1990): Invariant Measures on Groups and Their Use in Statistics. Institute of Mathematical Statistics, Hayward (Ca), Lecture Notes — Monograph Series **14**
82. Witting, H. (1985): Mathematische Statistik. Teubner, Stuttgart

Glossary

Numbers

$\mathbb{N}$	$\{1, 2, \ldots\}$	$\mathbb{R}$	real numbers
$\mathbb{N}_0$	$\mathbb{N} \cup \{0\}$	$\mathbb{R}^*$	$\mathbb{R} \setminus \{0\}$
$\mathbb{Z}$	integers	$\overline{\mathbb{R}}$	$\mathbb{R} \cup \{\infty\} \cup \{-\infty\}$
$\mathbb{Q}$	rational numbers	$\mathbb{C}$	complex numbers

The numbers that follow the symbols indicate the page numbers where their meanings are explained. Due to the conception of this book definitions may given more than once. Symbols which are only locally defined within particular proofs, paragraphs or examples are normally not included in the glossary. Occasionally, related symbols are subsumed under more general terms if these terms have already been defined in the same chapter. The list 'Measures' contains the term $\mathcal{M}_G(H)$, for example, but not $\mathcal{M}_{\mathrm{O}(n)}(\mathrm{Pos}(n))$ or $\mathcal{M}_{\mathrm{O}(n)}(\mathrm{GL}(n))$. The mapping Φ is multiply referenced as in the applications its general definition is adapted to the concrete situation. Densities are collected under the list 'Measures'.

Sets and Spaces

Mappings and Functions

φ_3	99	m_{P}	136
Eu	103	m_D	136
pr_2	116	s_D	136
Θ_r	123	ψ_{b}	137
χ_S	124	s_R	139
φ_P	124	$\psi_{P;1}$	141
q_*	125	$L(\cdot)$	141
ψ_D	131	χ_n	142
$\mathcal{F}_{\mathrm{b}i}(\mathrm{GL}(n))$	134	$\overline{\chi}_n$	145
φ_b	135		

σ-algebras

$\mathcal{B}(M)$	3, 13, 65	$\mathcal{B}_0(T)$	27, 68
$\mathcal{P}(\Omega)$	13, 64	$\mathcal{B}_E(T)$	27
$\mathcal{A}$	13, 64	$\mathcal{B}(H)_F$	31, 70
$\mathcal{A}_1 \otimes \mathcal{A}_2$	13, 64	$\mathcal{V}^m$	41
$\mathcal{B}_G(M)$	15, 66	$\mathcal{C}_F$	41
$\mathcal{A}_\eta$	25		

Measures

λ_2	3	μ_G	17, 74
$\lambda_{[0,2\pi)}$	3	λ	19, 76
$\lambda_{[0,\infty)}$	3	λ_n	19, 76
$\mu_{(\mathrm{S})}$	4, 26, 67	λ_C	19, 76
$\mathcal{M}^1(\Omega, \mathcal{A})$	13, 64	$N(\mu, \sigma^2)$	19, 76
$\mathcal{M}^+(\Omega, \mathcal{A})$	13, 64	ϵ_m	19, 76
$f \cdot \tau$	13, 64	$\mu_{(\mathrm{S}^2)}$	19, 99
η^φ	13, 64	τ^*	21, 67
$\mathcal{M}^\sigma(\Omega, \mathcal{A})$	13, 64	f^*	21, 67
$\eta_1 \otimes \eta_2$	14, 65	$\eta_1 \odot \eta_2$	21, 67
$\eta_{\mid \mathcal{A}_0}$	14, 65	η_v	25
$\mathcal{M}^1(M)$	14, 65	κ_*	27
$\mathcal{M}^\sigma(M)$	14, 65	f_T	28, 69
$\mathcal{M}^+(M)$	14, 65	ν_v	32
$\mathcal{M}(M)$	14, 65	P_Γ	40, 72
$\mathcal{M}_G^1(M)$	15, 66	P_{Γ_0}	40, 72
$\mathcal{M}_G(M)$	15, 66	$\mathrm{P}_{\Gamma \setminus \Gamma_0}$	40, 72
$\mathcal{M}_G^\sigma(M)$	15, 66	p_γ	40, 72
$\mathcal{M}_G^+(M)$	15, 66	$\mathcal{M}^1(V, \mathcal{V})^m$	41

Other Symbols

Index