Springer Monographs in Mathematics

Joachim Hilgert
Karl-Hermann Neeb

Structure and Geometry of Lie Groups

Springer

S*pringer* **M***onographs* in **M***athematics*

For further volumes:
www.springer.com/series/3733

Joachim Hilgert • Karl-Hermann Neeb

Structure and Geometry of Lie Groups

Springer

Joachim Hilgert
Mathematics Institute
University of Paderborn
Warburgerstr. 100
33095 Paderborn
Germany
hilgert@math.uni-paderborn.de

Karl-Hermann Neeb
Department of Mathematics
Friedrich-Alexander Universität
Erlangen-Nürnberg
Cauerstrasse 11
91054 Erlangen
Germany
neeb@mi.uni-erlangen.de

ISSN 1439-7382 Springer Monographs in Mathematics
ISBN 978-0-387-84793-1 e-ISBN 978-0-387-84794-8
DOI 10.1007/978-0-387-84794-8
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2011942060

Mathematics Subject Classification (2010): 7Bxx, 22Exx, 22Fxx

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Nowadays there are plenty of textbooks on Lie groups to choose from, so we feel we should explain why we decided to add another one to the row. Most of the readily available books on Lie groups either aim at an elementary introduction mostly restricted to matrix groups, or else they try to provide the background on semisimple Lie groups needed in harmonic analysis and unitary representation theory with as little general theory as possible. In [HN91], we tried to exhibit the basic principles of Lie theory rather than specific material, stressing the exponential function as the means of translating problems and solutions between the global and the infinitesimal level. In that book, written in German for German students who typically do not know differential geometry but are well versed in advanced linear algebra, we avoided abstract differentiable manifolds by combining matrix groups with covering arguments. Having introduced the basic principles, we demonstrated their power by proving a number of standard and not so standard results on the structure of Lie groups. The choice of results included owed a lot to Hochschild's book [Ho65], which even then was not so easy to come by.

This book builds on [HN91], but after twenty years of teaching and research in Lie theory we found it indispensable to also have the differential geometry of Lie groups available. Even though this is not apparent from the text, the reason for this is the large number of applications and further developments of Lie theory in which differential manifolds are essential. Moreover, we decided to include a number of structural results we found to be useful in the past but not readily available in the textbook literature. The basic line of thought now is:

- Simple examples: Matrix groups
- Tools from algebra: Lie algebras
- Tools from geometry: Smooth manifolds
- The basic principles: Lie groups, their Lie algebras, and the exponential function
- Structure theory: General Lie groups and special classes
- Testing methods on examples: The topology of classical groups
- A slight extension: Several connected components

While this book offers plenty of tested material for various introductory courses such as *Matrix Groups*, *Lie Groups*, *Lie Algebras*, or *Differentiable*

Manifolds, it is not a textbook to follow from A to Z (see page 6 for teaching suggestions). Moreover, it contains advanced material one would not typically include in a first course. In fact, some of the advanced material has not appeared in any monograph before. This and the fact that we wanted the book to be self contained is the reason for its considerable length. In order to still keep the work within reasonable limits, for some topics which are well covered in the textbook literature, we decided to include only what was needed for the further developments in the book. This applies, e.g., to the standard structure and classification theories of semisimple Lie algebras. Thus we do *not* want to suggest that this book can replace previous textbooks. It is meant rather to be a *true addition* to the existing textbook literature on Lie groups.

As was mentioned before, we are well aware of the fact that modern mathematics abounds with applications of Lie theory while this book hardly mentions any of them. The reason is that most applications require additional knowledge of the field in which these applications occur, so describing them would have meant either extensive storytelling or else a considerable expansion in length of this book. Neither option seemed attractive to us, so we leave it to future books to give detailed accounts of the beautiful ways in which Lie theory enters different fields of mathematics.

Even though there was a forerunner book and many lecture notes produced for various courses over the years, in compiling this text we produced many typos and made some mistakes. Many of those were shown to us by a small army of enthusiastic proofreaders to whom we are extremely grateful: Hanno Becker, Jan Emonds, Hasan Gündogan, Michael Klotz, Stéphane Merigon, Norman Metzner, Wolfgang Palzer, Matthias Peter, Niklas Schaeffer, Henrik Seppänen, and Stefan Wagner read major parts of the manuscript, and there were others who looked at particular sections. Of course, we know that the final version of the book will also contain mistakes, and we assume full responsibility for those.

We also would like to thank Ilka Agricola and Thomas Friedrich for some background information on the early history of Lie theory.

Paderborn *Joachim Hilgert*
Erlangen *Karl-Hermann Neeb*

Table of Contents

Part V Appendices

1 Introduction

To locate the theory of Lie groups within mathematics, one can say that Lie groups are groups with some additional structure that permits us to apply analytic techniques such as differentiation in a group theoretic context.

In the elementary courses on calculus of one variable, one studies functions on three levels:

(1) abstract functions between sets,
(2) continuous functions, and
(3) differentiable functions.

Going from level (1) to level (3), we refine the available tools at each step. At level (1), we have no structure at all to do anything; at level (2), we obtain results like the Intermediate Value Theorem or the Maximal Value Theorem saying that each continuous function on a compact interval takes a maximal value. The latter result is a useful existence theorem, but it provides no help at all to calculate maximal values. For that we need refined tools such as the derivative of a function and a transformation mechanism between properties of a function and its derivative. The situation is quite similar when we study groups. There is a level (1) consisting of abstract group theory which is particularly interesting for finite groups because the finiteness assumption is a powerful tool in the structure theory of finite groups. For infinite groups G, it is good to have a topology on G which is compatible with the group structure in the sense that the group operations are continuous, so that we are at level (2), and G is called a *topological group*. If we want to apply calculus techniques to study a group, we need *Lie groups*[1], i.e., groups which at the same time are differentiable manifolds such that the group operations are smooth.

For Lie groups we also need a translation mechanism telling us how to pass from group theoretic properties of G to properties of its "derivative" $\mathbf{L}(G)$, which in technical terms is the tangent space $T_{\mathbf{1}}(G)$ of G at the identity.

[1] The Norwegian mathematician Marius Sophus Lie (1842–1899) was the first to study differentiability properties of groups in a systematic way. In the 1890s, Sophus Lie developed his theory of differentiable groups (called continuous groups at a time when the concept of a topological space was not yet developed) to study symmetries of differential equations.

J. Hilgert, K.-H. Neeb, *Structure and Geometry of Lie Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-0-387-84794-8_1,

We think of $\mathbf{L}(G)$ as a "linear" object attached to the "nonlinear" object G because $\mathbf{L}(G)$ is a vector space endowed with an additional algebraic structure $[\cdot,\cdot]$, the *Lie bracket*, turning it into a *Lie algebra*. This algebraic structure is a bilinear map $\mathbf{L}(G)\times\mathbf{L}(G)\to\mathbf{L}(G)$, satisfying the axioms

$$[x,x]=0 \quad \text{and} \quad \big[x,[y,z]\big]+\big[y,[z,x]\big]+\big[z,[x,y]\big]=0 \quad \text{for} \quad x,y,z\in\mathbf{L}(G),$$

which can be considered as linearized versions of the group axioms. The connecting element between the group and its Lie algebra is the exponential function

$$\exp_G\colon\ \mathbf{L}(G)\to G$$

for which we have the Product Formula

$$\exp_G(x+y)=\lim_{k\to\infty}\big(\exp_G\big(\tfrac{1}{k}x\big)\exp_G\big(\tfrac{1}{k}y\big)\big)^k$$

and the Commutator Formula

$$\exp_G\big([x,y]\big)=\lim_{k\to\infty}\Big(\exp_G\big(\tfrac{1}{k}x\big)\exp_G\big(\tfrac{1}{k}y\big)\exp_G\big(-\tfrac{1}{k}x\big)\exp_G\big(-\tfrac{1}{k}y\big)\Big)^{k^2}$$

connecting the algebraic operations (addition and Lie bracket on $\mathbf{L}(G)$) to the group operations (multiplication and commutator). For the important class of matrix groups $G\subseteq\mathrm{GL}_n(\mathbb{R})$, the Lie algebra $\mathbf{L}(G)$ is a set of matrices and the exponential function is simply given by the power series $\exp_G(x)=\sum_{n=0}^{\infty}\frac{x^n}{n!}$.

An important property of the Lie algebra $\mathbf{L}(G)$ is that we can extend $\mathbf{L}$ to a smooth homomorphism $\varphi\colon G_1\to G_2$ of Lie groups by putting $\mathbf{L}(\varphi):=T_{\mathbf{1}}(\varphi)$ (the tangent map at $\mathbf{1}$) to obtain the so-called *Lie functor*, assigning to Lie groups Lie algebras and to group homomorphisms of Lie algebras. The compatibility of all that with the exponential function is encoded in the commutativity of the diagram

$$\begin{tikzcd}
\mathbf{L}(G_1) \arrow[r, "\mathbf{L}(\varphi)"] \arrow[d, "\exp_{G_1}"] & \mathbf{L}(G_2) \arrow[d, "\exp_{G_2}"] \\
G_1 \arrow[r, "\varphi"] & G_2.
\end{tikzcd}$$

The exponential function of a Lie group always maps sufficiently small 0-neighborhoods U in $\mathbf{L}(G)$ diffeomorphically to identity neighborhoods in G, so that the local structure of G is completely encoded in the multiplication

$$x*y:=(\exp_G|_U)^{-1}(\exp_G x\exp_G y),$$

which turns out to be given by a universal power series,

$$x*y=x+y+\frac{1}{2}[x,y]+\frac{1}{12}\big[x,[x,y]\big]+\frac{1}{12}\big[y,[y,x]\big]+\cdots.$$

Its summands are obtained by iterated Lie brackets whose precise structure we know after fundamental work of H.F. Baker, J.E. Campbell, E.B. Dynkin, and F. Hausdorff.

The basic philosophy of Lie theory now is that the local structure of the group G is determined by its Lie algebra $\mathbf{L}(G)$, and that the description of the global structure of a Lie group requires additional information that can be obtained in topological terms involving covering theory.

In Part I of this book, we approach the general concept of a Lie group by first discussing certain groups of matrices and groups arising in geometric contexts (Chapter 2). All these groups will later turn out to be Lie groups. In Chapter 3, we study the central tool in the theory of matrix groups that permits us to reverse the differentiation process from a Lie group G to its Lie algebra $\mathbf{L}(G)$: the exponential function $\exp_G\colon \mathbf{L}(G) \to G$, which is obtained by restriction from the matrix exponential function used in the theory of linear differential equations with constant coefficients. Chapter 4 treats Lie algebras of matrix groups and provides methods to calculate these Lie algebras effectively.

In Part II, we study Lie algebras as independent algebraic structures. We start in Chapter 5 by working out the standard approach: What are the substructures? Under which conditions does a substructure lead to a quotient structure? What are the simple structures? Does one have composition series? This leads to concepts like Lie subalgebras and ideals, nilpotent, solvable, and semisimple Lie algebras. In Chapter 6, we introduce Cartan subalgebras and the associated root and weight decompositions as tools to study the structure of (semi)simple Lie algebras. Further, we define *abstract root systems* and associated Weyl groups. Even though representation theory is not in the focus of this book, we provide in Chapter 7 the basic theory as it repeatedly plays an important role in structural questions. In particular, we introduce the universal enveloping algebra and prove the Poincaré–Birkhoff–Witt (PBW) Theorem on the structure of the enveloping algebra which implies in particular that each finite-dimensional Lie algebra sits in an associative algebra which has the same modules. From this we derive Serre's Theorem on the presentation of semisimple Lie algebras in terms of generators and relations, the Highest Weight Theorem on the classification of the simple finite dimensional modules, and Ado's Theorem on the existence of a faithful finite-dimensional representation of a finite-dimensional Lie algebra. Finally, we introduce basic cohomology theory for Lie algebras and describe extensions of Lie algebras.

In Part III, we provide an introduction to Lie groups based on the theory of smooth manifolds. The basic concepts and results from differential geometry needed for this are introduced in Chapter 8. In particular, we discuss vector fields on smooth manifolds and their integration to local flows. Chapter 9 is devoted to the subject proper of this book—Lie groups, defined as smooth manifolds with group structure such that all structure maps are smooth. Here we introduce the key tools of Lie theory. The Lie functor which associates a Lie algebra with a Lie group and the exponential function from the Lie algebra to the Lie group. They provide the means to translate global problems into infinitesimal ones and to lift infinitesimal solutions to local and, with the help of some additional topology, global ones. As a first set of applications of these methods, we identify the Lie group structures of closed subgroups of Lie groups and show how to construct Lie groups from local and infinitesimal data. Further, we explain covering theory for Lie groups. Finally, we prove Yamabe's Theorem asserting that any arcwise connected subgroup of a Lie group carries a natural Lie group structure, and this allows us to equip *any* subgroup of a Lie group with a canonical Lie group structure.

As we have explained before, a key method in Lie theory is to study the structure of Lie groups by translating group theoretic problems into linear algebra problems via the Lie functor, solving these problems, and translating the solutions back using the exponential function. In Part IV, we illustrate this general method by deriving a number of important structural results about Lie groups. Since in practice Lie groups often occur as symmetry groups which are not connected but have a finite number of connected components, we prove the results in this generality whenever it is possible without too much extra effort.

We start with quotient structures in Chapter 11, which also leads to homogeneous spaces, semidirect products, and eventually to a complete description of connected nilpotent and 1-connected solvable Lie groups.

In Chapter 12, we turn our attention to compact Lie groups and their covering groups. Again we first study the corresponding Lie algebras which are, by abuse of terminology, called *compact*. Then we prove Weyl's Theorem saying that the simply connected covering of a semisimple compact Lie group is compact. Further, we prove the important fact that a compact connected Lie group is the union of its maximal tori and show that such a Lie group is the semidirect product of its (semisimple) commutator subgroup and a torus subgroup. We also show that each compact Lie group is linear, i.e., can be realized as a closed subgroup of some $\mathrm{GL}_n(\mathbb{R})$. It is possible to describe the fundamental group in terms of the Lie algebra and the exponential map. In this context, we introduce the analytic Weyl group and a number of relevant lattices (i.e., discrete additive subgroups of maximal rank) in the Lie algebra $\mathfrak{t}$ of a maximal torus T and its dual $\mathfrak{t}^*$. The techniques are finally extended a little to prove that fixed point sets of automorphisms of simply connected groups are connected, a fact that is very useful, e.g., in the study of symmetric spaces.

Chapter 13 is devoted to the Cartan and the Iwasawa decomposition of noncompact semisimple Lie groups. These two decompositions are really only the starting point for a very rich structure theory which, in contrast to some other topics we present in this book, is very well covered in the existing literature (see, e.g., [Wa88] and [Kn02]). Therefore, we decided to keep this chapter brief.

In Chapter 14, we return to the general structure theory and show that each Lie group with finitely many connected components admits a maximal compact subgroup which is unique up to conjugation. In fact, it turns out that the group is diffeomorphic to a product of the maximal compact subgroup and a finite dimensional vector space (the Manifold Splitting Theorem 14.3.11). In particular, the topology of a Lie group with finitely many connected components is completely determined by any of its maximal compact subgroups. Before we can prove that we have to characterize the center of a connected Lie group as a certain subset of the exponential image (Theorem 14.2.8). The techniques developed for the proof of the Manifold Splitting Theorem also allow us to prove Dixmier's Theorem which characterizes the 1-connected solvable Lie groups for which the exponential function is a diffeomorphism. Finally, we study in detail under which circumstances one finds integral subgroups which are *not* closed, i.e., proper dense subgroups of their closure. In particular, we give a series of verifiable sufficient conditions for an integral subgroup to be closed. These results build on the classification of finitely generated abelian groups for which we provide a proof in Appendix 14.6.

In Chapter 15, we explain how to complexify Lie groups. It turns out that each Lie group G has a universal complexification $G_{\mathbb{C}}$, but G does in general not embed into $G_{\mathbb{C}}$. If G is compact, however, it does embed into its universal complexification, and this gives rise to the class of *linearly complex reductive Lie groups*. They can be characterized by the existence of a holomorphic faithful representation and the fact that all holomorphic representations are completely reducible, hence the name (see Theorem 15.3.11). On the way to this characterization, we study *abelian* complex connected Lie groups in some detail and introduce the *linearizer* of a complex group which measures how far the group is from being complex linear.

In the literature, one finds a lot of different notions of *reductive groups*, for many of which one imposes extra linearity properties. This is why in Chapter 16 we take a closer look at the structural implications of the existence of a faithful continuous finite dimensional representation of a Lie group. In particular, we introduce a *real linearizer* and the notion of *linearly real reductive groups*. Combining these notions with suitable Levi complements, we obtain a characterization of connected Lie groups which admit such faithful representations (see Theorems 16.2.7 and 16.2.9). The results of this chapter rely heavily on the results of Chapter 15. Conversely, we use the results of Chapter 16 to complete the discussion of the existence of faithful holomorphic representations in Section 16.3.

In Chapter 17, we apply the general results to compact and noncompact classical groups in order to provide explicit structural and topological information. In particular, we determine connected components and fundamental groups. Moreover, we include a rather detailed discussion of spin groups which builds on the material on Clifford algebras and related groups presented in Appendix B.3. Here we also explain a number of isomorphisms of low-dimensional groups. This discussion, as well as the one on conformal groups in Section 17.4 exemplifies the way Lie theory can be used to study groups defined in geometric terms. For a more detailed information of this kind, we refer to [GW09] for the classical and to [Ad96] for the exceptional Lie groups.

The examples from Chapter 17 show that many geometrically defined Lie groups have several connected components. While only the connected component of the identity is accessible to the methods built on the exponential function, there are still tools to analyze nonconnected Lie groups. In Chapter 18, we present some of these tools. The key notion is that of an extension of a discrete group by a (connected) Lie group. We explain how to classify such extensions in terms of group cohomology and apply this result to characterize those Lie groups with a finite number of connected components which admit a simply connected covering group.

1.1 Teaching Suggestions

For a one-semester course on *Lie algebras*, one could use Chapter 5 with possible additions from Sections 6.1 or 7.1. In a two-semester course on *Lie algebras*, one can cover most of the material in Chapters 5, 6, and 7.

A one-semester course on the Lie theory of *Matrix groups* can be drawn from Chapters 2 to 4. Chapter 8, together with Sections 10.2 and 10.3, makes a one-semester course on *Calculus on manifolds*. Building on such a course one can use Chapter 9 to teach a one-semester course on *Lie groups*.

Combining Chapters 9, 10, and 11, one obtains the material for a two-semester course on *Lie groups* with an emphasis on general structure theory using the exponential map. If one wants a two-semester course on *Lie groups* with an emphasis on semisimple and reductive groups, one should rather combine Chapters 9, 12, and 13, adding the necessary bits from Sections 10.4 and 11.1.

The remaining Chapters 14 through 18 are not interdependent, so they can be used to teach various different topics courses on *Lie groups*.

1.2 Fundamental Notation

Throughout this book $\mathbb{K}$ denotes either the field $\mathbb{R}$ of the real numbers or the field $\mathbb{C}$ of the complex numbers. All vector spaces will be $\mathbb{K}$-vector spaces if not otherwise specified. We write $M_n(\mathbb{K})$ for the ring of $(n \times n)$-matrices with entries in $\mathbb{K}$, $\mathbf{1}$ for the identity matrix, and $\mathrm{GL}_n(\mathbb{K})$ for its group of units, the general linear group. Further, we write $\mathbb{N} := \{1, 2, \dots\}$ for the set of the natural numbers and denote (half-)open intervals as $]a, b] := \{x \in \mathbb{R} \colon a < x \leq b\}$, $[a, b[:= \{x \in \mathbb{R} \colon a \leq x < b\}$, and $]a, b[:= \{x \in \mathbb{R} \colon a < x < b\}$.

Part I

Matrix Groups

2 Concrete Matrix Groups

In this chapter, we mainly study the general linear group $\mathrm{GL}_n(\mathbb{K})$ of invertible $n\times n$-matrices with entries in $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and introduce some of its subgroups. In particular, we discuss some of the connections between matrix groups and also introduce certain symmetry groups of geometric structures like bilinear or sesquilinear forms. In Section 2.3, we introduce also groups of matrices with entries in the quaternions $\mathbb{H}$.

2.1 The General Linear Group

We start with some notation. We write $\mathrm{GL}_n(\mathbb{K})$ for the group of invertible matrices in $M_n(\mathbb{K})$ and note that

$$\mathrm{GL}_n(\mathbb{K}) = \big\{g \in M_n(\mathbb{K})\colon \big(\exists h \in M_n(\mathbb{K})\big)\, hg = gh = \mathbf{1}\big\}.$$

Since the invertibility of a matrix can be tested with its determinant,

$$\mathrm{GL}_n(\mathbb{K}) = \big\{g \in M_n(\mathbb{K})\colon \det g \neq 0\big\}.$$

This group is called the *general linear group*.

On the vector space $\mathbb{K}^n$, we consider the *euclidian norm*

$$\|x\| := \sqrt{|x_1|^2 + \cdots + |x_n|^2}, \quad x \in \mathbb{K}^n,$$

and on $M_n(\mathbb{K})$ the corresponding *operator norm*

$$\|A\| := \sup\big\{\|Ax\|\colon x \in \mathbb{K}^n, \|x\| \leq 1\big\}$$

which turns $M_n(\mathbb{K})$ into a Banach space. On every subset $S \subseteq M_n(\mathbb{K})$, we shall always consider the subspace topology inherited from $M_n(\mathbb{K})$ (otherwise we shall say so). In this sense, $\mathrm{GL}_n(\mathbb{K})$ and all its subgroups carry a natural topology.

Lemma 2.1.1. *The group* $\mathrm{GL}_n(\mathbb{K})$ *has the following properties:*

(i) $\mathrm{GL}_n(\mathbb{K})$ *is open in* $M_n(\mathbb{K})$.

J. Hilgert, K.-H. Neeb, *Structure and Geometry of Lie Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-0-387-84794-8_2, © Springer Science+Business Media, LLC 2012

(ii) *The multiplication map* $m\colon \mathrm{GL}_n(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K}) \to \mathrm{GL}_n(\mathbb{K})$ *and the inversion map* $\eta\colon \mathrm{GL}_n(\mathbb{K}) \to \mathrm{GL}_n(\mathbb{K})$ *are smooth and in particular continuous.*

Proof. (i) Since the determinant function

$$\det\colon M_n(\mathbb{K}) \to \mathbb{K}, \quad \det(a_{ij}) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$$

is continuous and $\mathbb{K}^\times := \mathbb{K} \setminus \{0\}$ is open in $\mathbb{K}$, the set $\mathrm{GL}_n(\mathbb{K}) = \det^{-1}(\mathbb{K}^\times)$ is open in $M_n(\mathbb{K})$.

(ii) For $g \in \mathrm{GL}_n(\mathbb{K})$, we define $b_{ij}(g) := \det(g_{mk})_{m \neq j, k \neq i}$. According to Cramer's Rule, the inverse of g is given by

$$\left(g^{-1}\right)_{ij} = \frac{(-1)^{i+j}}{\det g} b_{ij}(g).$$

The smoothness of the inversion therefore follows from the smoothness of the determinant (which is a polynomial) and the polynomial functions b_{ij} defined on $M_n(\mathbb{K})$.

For the smoothness of the multiplication map, it suffices to observe that

$$(ab)_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$$

is the (ik)-entry in the product matrix. Since all these entries are quadratic polynomials in the entries of a and b, the product is a smooth map. □

Definition 2.1.2. A *topological group* G is a Hausdorff space G, endowed with a group structure, such that the multiplication map $m_G\colon G \times G \to G$ and the inversion map $\eta\colon G \to G$ are continuous, when $G \times G$ is endowed with the product topology.

Lemma 2.1.1(ii) says in particular that $\mathrm{GL}_n(\mathbb{K})$ is a topological group. It is clear that the continuity of group multiplication and inversion is inherited by every subgroup $G \subseteq \mathrm{GL}_n(\mathbb{K})$, so that every subgroup G of $\mathrm{GL}_n(\mathbb{K})$ also is a topological group.

We write a matrix $A = (a_{ij})_{i,j=1,\ldots,n}$ also as (a_{ij}) and define

$$A^\top := (a_{ji}), \quad \overline{A} := (\overline{a_{ij}}), \quad \text{and} \quad A^* := \overline{A}^\top = (\overline{a_{ji}}).$$

Note that $A^* = A^\top$ is equivalent to $\overline{A} = A$, which means that all entries of A are real. Now we can define the most important classes of matrix groups.

Definition 2.1.3. We introduce the following notation for some important subgroups of $\mathrm{GL}_n(\mathbb{K})$:

(1) The *special linear group*: $\mathrm{SL}_n(\mathbb{K}) := \{g \in \mathrm{GL}_n(\mathbb{K}) \colon \det g = 1\}$.
(2) The *orthogonal group*: $\mathrm{O}_n(\mathbb{K}) := \{g \in \mathrm{GL}_n(\mathbb{K}) \colon g^\top = g^{-1}\}$.
(3) The *special orthogonal group*: $\mathrm{SO}_n(\mathbb{K}) := \mathrm{SL}_n(\mathbb{K}) \cap \mathrm{O}_n(\mathbb{K})$.
(4) The *unitary group*: $\mathrm{U}_n(\mathbb{K}) := \{g \in \mathrm{GL}_n(\mathbb{K}) \colon g^* = g^{-1}\}$. Note that $\mathrm{U}_n(\mathbb{R}) = \mathrm{O}_n(\mathbb{R})$, but $\mathrm{O}_n(\mathbb{C}) \neq \mathrm{U}_n(\mathbb{C})$.
(5) The *special unitary group*: $\mathrm{SU}_n(\mathbb{K}) := \mathrm{SL}_n(\mathbb{K}) \cap \mathrm{U}_n(\mathbb{K})$.

One easily verifies that these are indeed subgroups. One simply has to use that $(ab)^\top = b^\top a^\top$, $\overline{ab} = \overline{a}\overline{b}$ and that

$$\det \colon \mathrm{GL}_n(\mathbb{K}) \to (\mathbb{K}^\times, \cdot)$$

is a group homomorphism.

We write $\mathrm{Herm}_n(\mathbb{K}) := \{A \in M_n(\mathbb{K}) \colon A^* = A\}$ for the set of *hermitian matrices*. For $\mathbb{K} = \mathbb{C}$, this is not a complex vector subspace of $M_n(\mathbb{K})$, but it is always a real subspace. A matrix $A \in \mathrm{Herm}_n(\mathbb{K})$ is called *positive definite* if for each $0 \neq z \in \mathbb{K}^n$ we have $\langle Az, z\rangle > 0$, where

$$\langle z, w\rangle := \sum_{j=1}^n z_j \overline{w_j}$$

is the natural scalar product on $\mathbb{K}^n$. We write $\mathrm{Pd}_n(\mathbb{K}) \subseteq \mathrm{Herm}_n(\mathbb{K})$ for the subset of positive definite matrices.

Lemma 2.1.4. *The groups*

$$\mathrm{U}_n(\mathbb{C}), \quad \mathrm{SU}_n(\mathbb{C}), \quad \mathrm{O}_n(\mathbb{R}), \quad \textit{and} \quad \mathrm{SO}_n(\mathbb{R})$$

are compact.

Proof. Since all these groups are subsets of $M_n(\mathbb{C}) \cong \mathbb{C}^{n^2}$, by the Heine–Borel Theorem we only have to show that they are closed and bounded.

Bounded: In view of

$$\mathrm{SO}_n(\mathbb{R}) \subseteq \mathrm{O}_n(\mathbb{R}) \subseteq \mathrm{U}_n(\mathbb{C}) \quad \text{and} \quad \mathrm{SU}_n(\mathbb{C}) \subseteq \mathrm{U}_n(\mathbb{C}),$$

it suffices to see that $\mathrm{U}_n(\mathbb{C})$ is bounded. Let $g_1, \ldots, g_n$ denote the rows of the matrix $g \in M_n(\mathbb{C})$. Then $g^* = g^{-1}$ is equivalent to $gg^* = \mathbf{1}$, which means that $g_1, \ldots, g_n$ form an orthonormal basis for $\mathbb{C}^n$ with respect to the scalar product $\langle z, w\rangle = \sum_{j=1}^n z_j \overline{w_j}$ which induces the norm $\|z\| = \sqrt{\langle z, z\rangle}$. Therefore, $g \in \mathrm{U}_n(\mathbb{C})$ implies $\|g_j\| = 1$ for each j, so that $\mathrm{U}_n(\mathbb{C})$ is bounded.

Closed: The functions

$$f, h \colon M_n(\mathbb{K}) \to M_n(\mathbb{K}), \quad f(A) := AA^* - \mathbf{1} \quad \text{and} \quad h(A) := AA^\top - \mathbf{1}$$

are continuous. Therefore, the groups

$$\mathrm{U}_n(\mathbb{K}) := f^{-1}(\mathbf{0}) \quad \text{and} \quad \mathrm{O}_n(\mathbb{K}) := h^{-1}(\mathbf{0})$$

are closed. Likewise $\mathrm{SL}_n(\mathbb{K}) = \det^{-1}(\mathbf{1})$ is closed, and therefore the groups $\mathrm{SU}_n(\mathbb{C})$ and $\mathrm{SO}_n(\mathbb{R})$ are also closed because they are intersections of closed subsets. □

2.1.1 The Polar Decomposition

Proposition 2.1.5 (Polar decomposition). *The multiplication map*

$$m\colon \mathrm{U}_n(\mathbb{K}) \times \mathrm{Pd}_n(\mathbb{K}) \to \mathrm{GL}_n(\mathbb{K}), \quad (u,p) \mapsto up$$

is a homeomorphism. In particular, each invertible matrix g can be written in a unique way as a product $g = up$ of a unitary matrix u and a positive definite matrix p.

Proof. We know from linear algebra that for each hermitian matrix A there exists an orthonormal basis $v_1, \ldots, v_n$ for $\mathbb{K}^n$ consisting of eigenvectors of A, and that all the corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$ are real (see [La93, Thm. XV.6.4]). From that it is obvious that A is positive definite if and only if $\lambda_j > 0$ holds for each j. For a positive definite matrix A, this has two important consequences:

(1) A is invertible, and A^{-1} satisfies $A^{-1}v_j = \lambda_j^{-1}v_j$.

(2) There exists a unique positive definite matrix B with $B^2 = A$ which will be denoted $\sqrt{A}$: We define B with respect to the basis $(v_1, \ldots, v_n)$ by $Bv_j = \sqrt{\lambda_j}v_j$. Then $B^2 = A$ is obvious and since all λ_j are real and the v_j are orthonormal, B is positive definite because

$$\left\langle B\Big(\sum_i \mu_i v_i\Big), \sum_j \mu_j v_j\right\rangle = \sum_{i,j} \mu_i\overline{\mu_j}\langle Bv_i, v_j\rangle = \sum_{j=1}^n |\mu_j|^2\sqrt{\lambda_j} > 0$$

for $\sum_j \mu_j v_j \neq 0$ and real coefficients μ_j. It remains to verify the uniqueness. So assume that C is positive definite with $C^2 = A$. Then $CA = C^3 = AC$ implies that C preserves all eigenspaces of A, so that we find an orthonormal basis $w_1, \ldots, w_n$ consisting of simultaneous eigenvectors of C and A (cf. Exercise 2.1.1). If $Cw_j = \alpha_j w_j$, we have $Aw_j = \alpha_j^2 w_j$, which implies that C acts on the λ-eigenspace of A by multiplication with $\sqrt{\lambda}$, which shows that $C = B$.

From (1) we derive that the image of the map m is contained in $\mathrm{GL}_n(\mathbb{K})$.

m **is surjective:** Let $g \in \mathrm{GL}_n(\mathbb{K})$. For $0 \neq v \in \mathbb{K}^n$ we then have

$$0 < \langle gv, gv\rangle = \langle g^*gv, v\rangle,$$

showing that g^*g is positive definite. Let $p := \sqrt{g^*g}$ and define $u := gp^{-1}$. Then

$$uu^* = gp^{-1}p^{-1}g^* = gp^{-2}g^* = g(g^*g)^{-1}g^* = gg^{-1}(g^*)^{-1}g^* = \mathbf{1}$$

implies that $u \in \mathrm{U}_n(\mathbb{K})$, and it is clear that $m(u,p) = g$.

m **is injective:** If $m(u,p) = m(w,q) = g$, then $g = up = wq$ implies that

$$p^2 = p^*p = (up)^*up = g^*g = (wq)^*wq = q^2,$$

so that p and q are positive definite square roots of the same positive definite matrix g^*g, hence coincide by (2) above. Now $p = q$, and therefore $u = gp^{-1} = gq^{-1} = w$.

It remains to show that m is a homeomorphism. Its continuity is obvious, so that it remains to prove the continuity of the inverse map m^{-1}. Let $g_j = u_j p_j \to g = up$. We have to show that $u_j \to u$ and $p_j \to p$. Since $\mathrm{U}_n(\mathbb{K})$ is compact, the sequence (u_j) has a subsequence (u_{j_k}) converging to some $w \in \mathrm{U}_n(\mathbb{K})$ by the Bolzano–Weierstraß Theorem. Then $p_{j_k} = u_{j_k}^{-1} g_{j_k} \to w^{-1}g =: q \in \mathrm{Herm}_n(\mathbb{K})$ and $g = wq$. For each $v \in \mathbb{K}^n$, we then have

$$0 \leq \langle p_{j_k} v, v\rangle \to \langle qv, v\rangle,$$

showing that all eigenvalues of q are ≥ 0. Moreover, $q = w^{-1}g$ is invertible, and therefore q is positive definite. Now $m(u,p) = m(w,q)$ yields $u = w$ and $p = q$. Since each convergent subsequence of (u_j) converges to u, the sequence itself converges to u (Exercise 2.1.9), and therefore $p_j = u_j^{-1} g_j \to u^{-1} g = p$. ☐

We shall see later that the set $\mathrm{Pd}_n(\mathbb{K})$ is homeomorphic to a vector space (Proposition 3.3.5), so that, topologically, the group $\mathrm{GL}_n(\mathbb{K})$ is a product of the compact group $\mathrm{U}_n(\mathbb{K})$ and a vector space. Therefore, the "interesting" part of the topology of $\mathrm{GL}_n(\mathbb{K})$ is contained in the compact group $\mathrm{U}_n(\mathbb{K})$.

Remark 2.1.6 (Normal forms of unitary and orthogonal matrices). We recall some facts from linear algebra:

(a) For each $u \in \mathrm{U}_n(\mathbb{C})$, there exists an orthonormal basis $v_1, \ldots, v_n$ consisting of eigenvectors of g (see [La93, Thm. XV.6.7]). This means that the unitary matrix s whose columns are the vectors $v_1, \ldots, v_n$ satisfies

$$s^{-1}us = \mathrm{diag}(\lambda_1, \ldots, \lambda_n),$$

where $uv_j = \lambda_j v$ and $|\lambda_j| = 1$.

The proof of this normal form is based on the existence of an eigenvector v_1 of u, which in turn follows from the existence of a zero of the characteristic polynomial. Since u is unitary, it preserves the hyperplane $v_1^\perp$ of dimension $n-1$. Now one uses induction to obtain an orthonormal basis $v_2, \ldots, v_n$ consisting of eigenvectors.

(b) For elements of $\mathrm{O}_n(\mathbb{R})$, the situation is more complicated because real matrices do not always have real eigenvectors.

Let $A \in M_n(\mathbb{R})$ and consider it as an element of $M_n(\mathbb{C})$. We assume that A does not have a real eigenvector. Then there exists an eigenvector $z \in \mathbb{C}^n$ corresponding to some eigenvalue $\lambda \in \mathbb{C}$. We write $z = x + iy$ and $\lambda = a + ib$. Then

$$Az = Ax + iAy = \lambda z = (ax - by) + i(ay + bx).$$

Comparing real and imaginary part yields

$$Ax = ax - by \quad \text{and} \quad Ay = ay + bx.$$

Therefore, the two-dimensional subspace generated by x and y in $\mathbb{R}^n$ is invariant under A.

This can be applied to $g \in \mathrm{O}_n(\mathbb{R})$ as follows. The argument above implies that there exists an invariant subspace $W_1 \subseteq \mathbb{R}^n$ with $\dim W_1 \in \{1, 2\}$. Then

$$W_1^\perp := \big\{v \in \mathbb{R}^n \colon \langle v, W_1 \rangle = \{0\}\big\}$$

is a subspace of dimension $n - \dim W_1$ which is also invariant (Exercise 2.1.14), and we apply induction to see that $\mathbb{R}^n$ is a direct sum of g-invariant subspaces $W_1, \ldots, W_k$ of dimension ≤ 2. Therefore, the matrix g is conjugate by an orthogonal matrix s to a block matrix of the form

$$d = \mathrm{diag}(d_1, \ldots, d_k),$$

where d_j is the matrix of the restriction of the linear map corresponding to g to W_j.

To understand the structure of the d_j, we have to take a closer look at the case $n \leq 2$. For $n = 1$ the group $\mathrm{O}_1(\mathbb{R}) = \{\pm 1\}$ consists of two elements, and for $n = 2$ an element $r \in \mathrm{O}_2(\mathbb{R})$ can be written as

$$r = \begin{pmatrix} a & \mp b \\ b & \pm a \end{pmatrix} \quad \text{with} \quad \det r = \pm\big(a^2 + b^2\big) = \pm 1,$$

because the second column contains a unit vector orthogonal to the first one. With $a = \cos\alpha$ and $b = \sin\alpha$ we get

$$r = \begin{pmatrix} \cos\alpha & \mp\sin\alpha \\ \sin\alpha & \pm\cos\alpha \end{pmatrix}.$$

For $\det r = -1$, we obtain

$$r^2 = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}\begin{pmatrix} a & b \\ b & -a \end{pmatrix} = \mathbf{1},$$

but this implies that r is an orthogonal reflection with the two eigenvalues ± 1 (Exercise 2.1.13), hence has two orthogonal eigenvectors.

In view of the preceding discussion, we may therefore assume that the first m of the matrices d_j are of the rotation form

$$d_j = \begin{pmatrix} \cos\alpha_j & -\sin\alpha_j \\ \sin\alpha_j & \cos\alpha_j \end{pmatrix},$$

that $d_{m+1}, \ldots, d_\ell$ are -1, and that $d_{\ell+1}, \ldots, d_n$ are 1:

$$\begin{pmatrix}
\cos\alpha_1 & -\sin\alpha_1 & & & & & & & & & \\
\sin\alpha_1 & \cos\alpha_1 & & & & & & & & & \\
& & \ddots & & & & & & & & \\
& & & \cos\alpha_m & -\sin\alpha_m & & & & & & \\
& & & \sin\alpha_m & \cos\alpha_m & & & & & & \\
& & & & & -1 & & & & & \\
& & & & & & \ddots & & & & \\
& & & & & & & -1 & & & \\
& & & & & & & & 1 & & \\
& & & & & & & & & \ddots & \\
& & & & & & & & & & 1
\end{pmatrix}.$$

For $n = 3$, we obtain in particular the normal form

$$d = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & \pm 1 \end{pmatrix}.$$

From this normal form we immediately read off that $\det d = 1$ is equivalent to d describing a rotation around an axis consisting of fixed points (the axis is $\mathbb{R}e_3$ for the normal form matrix).

Proposition 2.1.7. (a) *The group* $\mathrm{U}_n(\mathbb{C})$ *is arcwise connected.*

(b) *The group* $\mathrm{O}_n(\mathbb{R})$ *has the two arc components*

$$\mathrm{SO}_n(\mathbb{R}) \quad \text{and} \quad \mathrm{O}_n(\mathbb{R})_- := \big\{g \in \mathrm{O}_n(\mathbb{R}) \colon \det g = -1\big\}.$$

Proof. (a) First we consider $\mathrm{U}_n(\mathbb{C})$. To see that this group is arcwise connected, let $u \in \mathrm{U}_n(\mathbb{C})$. Then there exists an orthonormal basis $v_1, \ldots, v_n$ of eigenvectors of u (Remark 2.1.6(a)). Let $\lambda_1, \ldots, \lambda_n$ denote the corresponding eigenvalues. Then the unitarity of u implies that $|\lambda_j| = 1$, and we therefore find $\theta_j \in \mathbb{R}$ with $\lambda_j = e^{\theta_j i}$. Now we define a continuous curve

$$\gamma \colon [0,1] \to \mathrm{U}_n(\mathbb{C}), \quad \gamma(t)v_j := e^{t\theta_j i} v_j, \ j = 1, \ldots, n.$$

We then have $\gamma(0) = \mathbf{1}$ and $\gamma(1) = u$. Moreover, each $\gamma(t)$ is unitary because the basis $(v_1, \ldots, v_n)$ is orthonormal.

(b) For $g \in \mathrm{O}_n(\mathbb{R})$, we have $gg^\top = \mathbf{1}$, and therefore $1 = \det(gg^\top) = (\det g)^2$. This shows that

$$\mathrm{O}_n(\mathbb{R}) = \mathrm{SO}_n(\mathbb{R}) \,\dot\cup\, \mathrm{O}_n(\mathbb{R})_-$$

and both sets are closed in $\mathrm{O}_n(\mathbb{R})$ because det is continuous. Therefore, $\mathrm{O}_n(\mathbb{R})$ is not connected, and hence not arcwise connected. Suppose we knew that $\mathrm{SO}_n(\mathbb{R})$ is arcwise connected and $x, y \in \mathrm{O}_n(\mathbb{R})_-$. Then $\mathbf{1}, x^{-1}y \in \mathrm{SO}_n(\mathbb{R})$ can be connected by an arc $\gamma \colon [0,1] \to \mathrm{SO}_n(\mathbb{R})$, and then $t \mapsto x\gamma(t)$ defines an arc $[0,1] \to \mathrm{O}_n(\mathbb{R})_-$ connecting x to y. So it remains to show that $\mathrm{SO}_n(\mathbb{R})$ is arcwise connected.

Let $g \in \mathrm{SO}_n(\mathbb{R})$. In the normal form of g discussed in Remark 2.1.6, the determinant of each two-dimensional block is 1, so that the determinant is the product of all -1-eigenvalues. Hence their number is even, and we can write each consecutive pair as a block

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

This shows that with respect to some orthonormal basis for $\mathbb{R}^n$ the linear map defined by g has a matrix of the form

$$g = \begin{pmatrix} \cos\alpha_1 & -\sin\alpha_1 & & & & & & \\ \sin\alpha_1 & \cos\alpha_1 & & & & & & \\ & & \ddots & & & & & \\ & & & \cos\alpha_m & -\sin\alpha_m & & & \\ & & & \sin\alpha_m & \cos\alpha_m & & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix}.$$

Now we obtain an arc $\gamma\colon [0,1] \to \mathrm{SO}_n(\mathbb{R})$ with $\gamma(0) = \mathbf{1}$ and $\gamma(1) = g$ by

$$\gamma(t) := \begin{pmatrix} \cos t\alpha_1 & -\sin t\alpha_1 & & & & & & \\ \sin t\alpha_1 & \cos t\alpha_1 & & & & & & \\ & & \ddots & & & & & \\ & & & \cos t\alpha_m & -\sin t\alpha_m & & & \\ & & & \sin t\alpha_m & \cos t\alpha_m & & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix}.$$

$\square$

Corollary 2.1.8. *The group* $\mathrm{GL}_n(\mathbb{C})$ *is arcwise connected and the group* $\mathrm{GL}_n(\mathbb{R})$ *has two arc-components given by*

$$\mathrm{GL}_n(\mathbb{R})_\pm := \{g \in \mathrm{GL}_n(\mathbb{R})\colon\ \pm\det g > 0\}.$$

Proof. If $X = A \times B$ is a product space, then the arc-components of X are the sets of the form $C \times D$, where $C \subseteq A$ and $D \subseteq B$ are arc-components (easy Exercise!). The polar decomposition of $\mathrm{GL}_n(\mathbb{K})$ yields a homeomorphism

$$\mathrm{GL}_n(\mathbb{K}) \cong \mathrm{U}_n(\mathbb{K}) \times \mathrm{Pd}_n(\mathbb{K}).$$

Since $\mathrm{Pd}_n(\mathbb{K})$ is an open convex set, it is arcwise connected (Exercise 2.1.6). Therefore, the arc-components of $\mathrm{GL}_n(\mathbb{K})$ are in one-to-one correspondence with those of $\mathrm{U}_n(\mathbb{K})$ which have been determined in Proposition 2.1.7. $\square$

2.1.2 Normal Subgroups of $\mathbf{GL}_n(\mathbb{K})$

We shall frequently need some basic concepts from group theory which we recall in the following definition.

Definition 2.1.9. Let G be a group with identity element e.

(a) A subgroup $N \subseteq G$ is called *normal* if $gN = Ng$ holds for all $g \in G$. We write this as $N \trianglelefteq G$. The normality implies that the quotient set G/N (the set of all cosets of the subgroup N) inherits a natural group structure by

$$gN \cdot hN := ghN$$

for which eN is the identity element and the quotient map $q\colon G \to G/N$ is a surjective group homomorphism with kernel $N = \ker q = q^{-1}(eN)$.

On the other hand, all kernels of group homomorphisms are normal subgroups, so that the normal subgroups are precisely those which are kernels of group homomorphisms.

It is clear that G and $\{e\}$ are normal subgroups. We call G *simple* if $G \neq \{e\}$ and these are the only normal subgroups.

(b) The subgroup $Z(G) := \{g \in G\colon (\forall x \in G) gx = xg\}$ is called the *center of* G. It obviously is a normal subgroup of G. For $x \in G$, the subgroup

$$Z_G(x) := \{g \in G\colon gx = xg\}$$

is called its *centralizer*. Note that $Z(G) = \bigcap_{x \in G} Z_G(x)$.

(c) If $G_1, \ldots, G_n$ are groups, then the product set $G := G_1 \times \cdots \times G_n$ has a natural group structure given by

$$(g_1, \ldots, g_n)(g_1', \ldots, g_n') := (g_1 g_1', \ldots, g_n g_n').$$

The group G is called the *direct product* of the groups G_j, $j = 1, \ldots, n$. We identify G_j with a subgroup of G. Then all subgroups G_j are normal subgroups and $G = G_1 \cdots G_n$.

In the following, we write $\mathbb{R}_+^\times :=]0, \infty[$.

Proposition 2.1.10. (a) $Z(\mathrm{GL}_n(\mathbb{K})) = \mathbb{K}^\times \mathbf{1}$.

(b) *The multiplication map*

$$\varphi\colon (\mathbb{R}_+^\times, \cdot) \times \mathrm{SL}_n(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R})_+, \quad (\lambda, g) \mapsto \lambda g$$

is a homeomorphism and a group isomorphism, i.e., an isomorphism of topological groups.

Proof. (a) It is clear that $\mathbb{K}^\times \mathbf{1}$ is contained in the center of $\mathrm{GL}_n(\mathbb{K})$. To see that each matrix $g \in Z(\mathrm{GL}_n(\mathbb{K}))$ is a multiple of $\mathbf{1}$, we consider the elementary matrix $E_{ij} := (\delta_{ij})$ with the only nonzero entry 1 in position (i, j). For $i \neq j$, we then have $E_{ij}^2 = 0$, so that $(\mathbf{1} + E_{ij})(\mathbf{1} - E_{ij}) = \mathbf{1}$, which implies that $T_{ij} := \mathbf{1} + E_{ij} \in \mathrm{GL}_n(\mathbb{K})$. From the relation $gT_{ij} = T_{ij}g$ we immediately get $gE_{ij} = E_{ij}g$ for $i \neq j$, so that for $k, \ell \in \{1, \ldots, n\}$ we get

$$g_{ki}\delta_{j\ell} = (gE_{ij})_{k\ell} = (E_{ij}g)_{k\ell} = \delta_{ik}g_{j\ell}.$$

For $k = i$ and $\ell = j$, we obtain $g_{ii} = g_{jj}$; and for $k = j = \ell$, we get $g_{ji} = 0$. Therefore, $g = \lambda\mathbf{1}$ for some $\lambda \in \mathbb{K}$.

(b) It is obvious that φ is a group homomorphism and that φ is continuous. Moreover, the map

$$\psi\colon\ \mathrm{GL}_n(\mathbb{R})_+ \to \mathbb{R}_+^\times \times \mathrm{SL}_n(\mathbb{R}), \quad g \mapsto \big((\det g)^{\frac{1}{n}}, (\det g)^{-\frac{1}{n}}g\big)$$

is continuous and satisfies $\varphi \circ \psi = \mathrm{id}$ and $\psi \circ \varphi = \mathrm{id}$. Hence φ is a homeomorphism. $\square$

Remark 2.1.11. The subgroups

$$Z\big(\mathrm{GL}_n(\mathbb{K})\big) \quad \text{and} \quad \mathrm{SL}_n(\mathbb{K})$$

are normal subgroups of $\mathrm{GL}_n(\mathbb{K})$. Moreover, for $\mathrm{GL}_n(\mathbb{R})$ the subgroup $\mathrm{GL}_n(\mathbb{R})_+$ is a proper normal subgroup and the same holds for $\mathbb{R}_+^\times\mathbf{1}$. One can show that these examples exhaust all normal arcwise connected subgroups of $\mathrm{GL}_n(\mathbb{K})$.

2.1.3 Exercises for Section 2.1

Exercise 2.1.1. Let V be a $\mathbb{K}$-vector space and $A \in \mathrm{End}(V)$. We write $V_\lambda(A) := \ker(A - \lambda\mathbf{1})$ for the *eigenspace of* A corresponding to the eigenvalue λ and $V^\lambda(A) := \bigcup_{n\in\mathbb{N}} \ker(A - \lambda\mathbf{1})^n$ for the *generalized eigenspace of* A corresponding to λ.

(a) If $A, B \in \mathrm{End}(V)$ commute, then

$$BV^\lambda(A) \subseteq V^\lambda(A) \quad \text{and} \quad BV_\lambda(A) \subseteq V_\lambda(A)$$

holds for each $\lambda \in \mathbb{K}$.

(b) If $A \in \mathrm{End}(V)$ is diagonalizable and $W \subseteq V$ is an A-invariant subspace, then $A|_W \in \mathrm{End}(W)$ is diagonalizable.

(c) If $A, B \in \mathrm{End}(V)$ commute and both are diagonalizable, then they are simultaneously diagonalizable, i.e., there exists a basis for V which consists of eigenvectors of A and B.

(d) If $\dim V < \infty$ and $\mathcal{A} \subseteq \mathrm{End}(V)$ is a commuting set of diagonalizable endomorphisms, then $\mathcal{A}$ can be simultaneously diagonalized, i.e., V is a direct sum of simultaneous eigenspaces of $\mathcal{A}$.

(e) For any function $\lambda\colon \mathcal{A} \to V$, we write $V_\lambda(\mathcal{A}) = \bigcap_{a\in\mathcal{A}} V_{\lambda(a)}(a)$ for the corresponding simultaneous eigenspace. Show that the sum $\sum_\lambda V_\lambda(\mathcal{A})$ is direct.

(f) If $\mathcal{A} \subseteq \mathrm{End}(V)$ is a finite commuting set of diagonalizable endomorphisms, then $\mathcal{A}$ can be simultaneously diagonalized.

(g) Find a commuting set of diagonalizable endomorphisms of a vector space V which cannot be diagonalized simultaneously.

Exercise 2.1.2. Let G be a topological group. Let G_0 be the *identity component*, i.e., the connected component of the identity in G. Show that G_0 is a closed normal subgroup of G.

Exercise 2.1.3. $\mathrm{SO}_n(\mathbb{K})$ is a closed normal subgroup of $\mathrm{O}_n(\mathbb{K})$ of index 2 and, for every $g \in \mathrm{O}_n(\mathbb{K})$ with $\det(g) = -1$,

$$\mathrm{O}_n(\mathbb{K}) = \mathrm{SO}_n(\mathbb{K}) \cup g\,\mathrm{SO}_n(\mathbb{K})$$

is a disjoint decomposition.

Exercise 2.1.4. For each subset $M \subseteq M_n(\mathbb{K})$, the *centralizer*

$$Z_{\mathrm{GL}_n(\mathbb{K})}(M) := \big\{g \in \mathrm{GL}_n(\mathbb{K}) : (\forall m \in M) gm = mg\big\}$$

is a closed subgroup of $\mathrm{GL}_n(\mathbb{K})$.

Exercise 2.1.5. We identify $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ by the map $z = x + iy \mapsto (x, y)$ and write $I(x, y) := (-y, x)$ for the real linear endomorphism of $\mathbb{R}^{2n}$ corresponding to multiplication with i. Then

$$\mathrm{GL}_n(\mathbb{C}) \cong Z_{\mathrm{GL}_{2n}(\mathbb{R})}\big(\{I\}\big)$$

yields a realization of $\mathrm{GL}_n(\mathbb{C})$ as a closed subgroup of $\mathrm{GL}_{2n}(\mathbb{R})$.

Exercise 2.1.6. A subset C of a real vector space V is called a *convex cone* if C is convex and $\lambda C \subseteq C$ for each $\lambda > 0$.

Show that $\mathrm{Pd}_n(\mathbb{K})$ is an open convex cone in $\mathrm{Herm}_n(\mathbb{K})$.

Exercise 2.1.7. Show that

$$\gamma : (\mathbb{R}, +) \to \mathrm{GL}_2(\mathbb{R}), \quad t \mapsto \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}$$

is a continuous group homomorphism with $\gamma(\pi) = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and $\mathrm{im}\,\gamma = \mathrm{SO}_2(\mathbb{R})$.

Exercise 2.1.8. Show that the group $\mathrm{O}_n(\mathbb{C})$ is homeomorphic to the topological product of the subgroup

$$\mathrm{O}_n(\mathbb{R}) \cong \mathrm{U}_n(\mathbb{C}) \cap \mathrm{O}_n(\mathbb{C}) \quad \text{and the set} \quad \mathrm{Pd}_n(\mathbb{C}) \cap \mathrm{O}_n(\mathbb{C}).$$

Exercise 2.1.9. Let (X, d) be a compact metric space and $(x_n)_{n\in\mathbb{N}}$ a sequence in X. Show that $\lim_{n\to\infty} x_n = x$ is equivalent to the condition that each convergent subsequence $(x_{n_k})_{k\in\mathbb{N}}$ converges to x.

Exercise 2.1.10. If $A \in \mathrm{Herm}_n(\mathbb{K})$ satisfies $\langle Av, v\rangle = 0$ for each $v \in \mathbb{K}^n$, then $A = 0$.

Exercise 2.1.11. Show that for a complex matrix $A \in M_n(\mathbb{C})$ the following are equivalent:

(1) $A^* = A$.
(2) $\langle Av, v\rangle \in \mathbb{R}$ for each $v \in \mathbb{C}^n$.

Exercise 2.1.12. (a) Show that a matrix $A \in M_n(\mathbb{K})$ is hermitian if and only if there exists an orthonormal basis $v_1, \ldots, v_n$ for $\mathbb{K}^n$ and real numbers $\lambda_1, \ldots, \lambda_n$ with $Av_j = \lambda_j v_j$.

(b) Show that a complex matrix $A \in M_n(\mathbb{C})$ is unitary if and only if there exists an orthonormal basis $v_1, \ldots, v_n$ for $\mathbb{K}^n$ and $\lambda_j \in \mathbb{C}$ with $|\lambda_j| = 1$ and $Av_j = \lambda_j v_j$.

(c) Show that a complex matrix $A \in M_n(\mathbb{C})$ is *normal*, i.e., satisfies $A^*A = AA^*$, if and only if there exists an orthonormal basis $v_1, \ldots, v_n$ for $\mathbb{K}^n$ and $\lambda_j \in \mathbb{C}$ with $Av_j = \lambda_j v_j$.

Exercise 2.1.13. (a) Let V be a vector space and $\mathbf{1} \neq A \in \mathrm{End}(V)$ with $A^2 = \mathbf{1}$ (A is called an *involution*). Show that

$$V = \ker(A - \mathbf{1}) \oplus \ker(A + \mathbf{1}).$$

(b) Let V be a vector space and $A \in \mathrm{End}(V)$ with $A^3 = A$. Show that

$$V = \ker(A - \mathbf{1}) \oplus \ker(A + \mathbf{1}) \oplus \ker A.$$

(c) Let V be a vector space and $A \in \mathrm{End}(V)$ an endomorphism for which there exists a polynomial p of degree n with n different zeros $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ and $p(A) = 0$. Show that A is diagonalizable with eigenvalues $\lambda_1, \ldots, \lambda_n$.

Exercise 2.1.14. Let $\beta\colon V \times V \to \mathbb{K}$ be a bilinear map and $g : V \to V$ with $\beta(gv, gw) = \beta(v, w)$ be a β-isometry. For a subspace $E \subseteq V$, we write

$$E^\perp := \big\{v \in V\colon (\forall w \in E)\ \beta(v, w) = 0\big\}$$

for its *orthogonal space*. Show that $g(E) = E$ implies that $g(E^\top) = E^\top$.

Exercise 2.1.15 (Iwasawa decomposition of $\mathrm{GL}_n(\mathbb{R})$). Let

$$T_n^+(\mathbb{R}) \subseteq \mathrm{GL}_n(\mathbb{R})$$

denote the subgroup of upper-triangular matrices with positive diagonal entries. Show that the multiplication map

$$\mu\colon\ \mathrm{O}_n(\mathbb{R}) \times T_n^+(\mathbb{R}) \to \mathrm{GL}_n(\mathbb{R}), \quad (a, b) \mapsto ab$$

is a homeomorphism.

Exercise 2.1.16. Let $\mathbb{K}$ be a field and $n \in \mathbb{N}$. Show that

$$Z\big(M_n(\mathbb{K})\big) := \big\{z \in M_n(\mathbb{K})\colon \big(\forall x \in M_n(\mathbb{K})\big)\, zx = xz\big\} = \mathbb{K}\mathbf{1}.$$

2.2 Groups and Geometry

In Definition 2.1.3, we have defined certain matrix groups by concrete conditions on the matrices. Often it is better to think of matrices as linear maps described with respect to a basis. To do that, we have to adopt a more abstract point of view. Similarly, one can study symmetry groups of bilinear forms on a vector space V without fixing a certain basis a priori. Actually, it is much more convenient to choose a basis for which the structure of the bilinear form is as simple as possible.

2.2.1 Isometry Groups

Definition 2.2.1 (Groups and bilinear forms). (a) (The abstract general linear group) Let V be a $\mathbb{K}$-vector space. We write $\mathrm{GL}(V)$ for the group of linear automorphisms of V. This is the group of invertible elements in the ring $\mathrm{End}(V)$ of all linear endomorphisms of V.

If V is an n-dimensional $\mathbb{K}$-vector space and $v_1, \ldots, v_n$ is a basis for V, then the map

$$\Phi\colon M_n(\mathbb{K}) \to \mathrm{End}(V), \quad \Phi(A)v_k := \sum_{j=1}^{n} a_{jk}v_j$$

is a linear isomorphism which describes the passage between linear maps and matrices. In view of $\Phi(\mathbf{1}) = \mathrm{id}_V$ and $\Phi(AB) = \Phi(A)\Phi(B)$, we obtain a group isomorphism

$$\Phi|_{\mathrm{GL}_n(\mathbb{K})}\colon \mathrm{GL}_n(\mathbb{K}) \to \mathrm{GL}(V).$$

(b) Let V be an n-dimensional vector space with basis $v_1, \ldots, v_n$ and $\beta\colon V \times V \to \mathbb{K}$ a bilinear map. Then $B = (b_{jk}) := (\beta(v_j, v_k))_{j,k=1,\ldots,n}$ is an $(n \times n)$-matrix, but this matrix should NOT be interpreted as the matrix of a linear map. It is the matrix of a bilinear map to $\mathbb{K}$, which is something different. It describes β in the sense that

$$\beta\Big(\sum_j x_j v_j, \sum_k y_k v_k\Big) = \sum_{j,k=1}^{n} x_j b_{jk} y_k = x^\top B y,$$

where $x^\top B y$ with column vectors $x, y \in \mathbb{K}^n$ is viewed as a matrix product whose result is a (1×1)-matrix, i.e., an element of $\mathbb{K}$.

We write

$$\mathrm{Aut}(V, \beta) := \big\{g \in \mathrm{GL}(V)\colon (\forall v, w \in V)\, \beta(gv, gw) = \beta(v, w)\big\}$$

for the *isometry group of the bilinear form* β. Then it is easy to see that

$$\Phi^{-1}\big(\mathrm{Aut}(V, \beta)\big) = \big\{g \in \mathrm{GL}_n(\mathbb{K})\colon g^\top B g = B\big\}.$$