Lie Groups and Matrix Groups

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 one variable calculus 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 (Zwischenwertsatz) or the Maximal Value Theorem that each 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 translation 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), which we could call the level of *continuous groups*. If we want to apply calculus techniques to study a group, we need something like a *differentiable group*, and this means a *Lie group*[1]. We shall see that for Lie groups we shall also need a translation mechanism telling us how to pass from group theoretic properties of G to properties of its "derivative" $\mathbf{L}(G)$, the *Lie algebra of* G. We think of $\mathbf{L}(G)$ as a "linear" object attached to the "non-linear" object G, because $\mathbf{L}(G)$ is a vector space endowed with an additional algebraic structure $[\cdot,\cdot]$, the *Lie bracket*. This is a bilinear map $\mathbf{L}(G)\times\mathbf{L}(G)\to\mathbf{L}(G)$ satisfying the axioms

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

which can be considered as linearized versions of the group axioms of G.

Historically abstract groups arose first in the setting of Galois theory in the early 19^{th} century, where they arose as automorphism groups of fields. The groups considered in Galois theory at these times were all finite. Later in the

[1] The norwegian mathematician Marius Sophus Lie (1842–1899) was the first to study differentiability properties of groups in a systematic way.

19^{th} century "continuous" groups became important in geometric contexts to the extent that a *geometry on a space* S was considered as the same structure as a *group acting on this space*. Then the *geometric properties* are those invariant under the action of the group. This is the content of Felix Klein's "Erlanger Programm" from 1872[2]. A typical example is the group $\mathrm{Mot}(E_2)$ of motions (orientation preserving isometries) of the euclidean plane. The length of an interval or the area of a triangle are properties preserved by this group, hence a geometric quantity. It was an important conceptual step to observe that changing the group means changing the notion of a geometric quantity. For example the automorphism group $\mathrm{Aff}(A_2)$ of the two-dimensional affine plane A_2 does not preserve the area of a triangle (it is larger than the euclidean group), hence the area of a triangle cannot be considered as an affine geometric quantity. 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 clarified) to study symmetries of differential equations.

In this course we shall approach the general concept of a Lie group by first discussing certain groups of matrices and groups arising in geometric contexts (Chapter I). All these groups will later turn out to be Lie groups. In Chapter II we study the central tool in the theory of Lie groups that permits us to reverse the differentiation process from a Lie group G to its Lie algebra $\mathbf{L}(G)$: the exponential function $\exp_G: \mathbf{L}(G) \to G$, which is a generalization of the matrix exponential function used in the theory of linear differential equations with constant coefficients.

In Chapter III we then turn to the question which kind of structure on a group G turns it into a Lie group. Here we shall see that there is an intermediate level given by a group topology on G. The essential feature of a Lie group G is that there is an open identity neighborhood $U \subseteq G$ which is homeomorphic to an open subset of $\mathbb{R}^n$ such that the group operations are smooth mappings in a neighborhood of the identity element.

After Chapters I-III the concept of a Lie group is essentially clarified, so that we can move on and have a closer look at constructions producing new Lie groups from given ones. In particular, we turn to covering groups of Lie groups, which already leads us beyond the class of linear Lie groups and discuss the general concept of a Lie group.

---

[2] Christian Felix Klein (1849–1925) held the chair of geometry in Erlangen for a few years and the "Erlanger Programm" was his "Programmschrift," where he formulated his research plans when he came to Erlangen.

I. Concrete matrix groups

In this section $\mathbb{K}$ denotes the field $\mathbb{R}$ of real numbers or the field $\mathbb{C}$ of complex numbers. First we mainly study the group $\mathrm{GL}_n(\mathbb{K})$ of invertible $n \times n$-matrices and introduce some of its subgroups.

I.1. The general linear group

First we introduce some notation. We write $M_n(\mathbb{K})$ for the ring of $(n \times n)$-matrices with entries in $\mathbb{K}$ and $\mathbf{1}$ for the identity matrix. Then

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

is the set of invertible elements in $M_n(\mathbb{K})$, and we know from Linear Algebra that invertibility of a matrix can be tested with its determinant. Therefore we have

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

This group is called the *general linear group.*

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

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

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

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

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 subgroups of $\mathrm{GL}_n(\mathbb{K})$ carry a natural topology.

Lemma I.1.1. *The group* $\mathrm{GL}_n(\mathbb{K})$ *has the following properties:*

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

(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 $b_{ij}(g) := \det(g_{mk})_{m \neq j, k \neq i}$. Then Cramer's Rule says that the inverse of g is given by

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

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

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 polynomials in the entries of a and b, the product is a smooth map. ∎

Put in more abstract terms, Lemma I.1.1(ii) says that $\mathrm{GL}_n(\mathbb{K})$ is a topological group, a concept that we shall study more thoroughly in a later chapter. 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 easily define the most important classes of matrix groups.

Definition I.1.2. One easily verifies that the following sets are indeed groups. 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.

(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 that $\mathrm{O}_n(\mathbb{C}) \neq \mathrm{U}_n(\mathbb{C})$.
(5) The *special unitary group:* $\mathrm{SU}_n(\mathbb{K}) := \mathrm{SL}_n(\mathbb{K}) \cap \mathrm{U}_n(\mathbb{K})$.

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 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 $\langle A.z, z\rangle > 0$, 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 I.1.3. *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 $M_n(\mathbb{C}) \cong \mathbb{C}^{n^2}$, we have to show that they are closed and bounded.

Boundedness: 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 of $\mathbb{C}^n$ 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, and therefore $\mathrm{U}_n(\mathbb{C})$ is bounded.

Closedness: The functions

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

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 $\mathrm{SL}_n(\mathbb{K})$ 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. ■

Proposition I.1.4. (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 $g = up$ with u unitary and p positive definite.

Proof. We know from Linear Algebra that for each hermitian matrix A there exists an orthonormal basis $v_1, \ldots, v_n$ of $\mathbb{K}^n$ consisting of eigenvectors of A, and that all the corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$ are real. 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} is given with respect to the basis $(v_1, \ldots, v_n)$ by $A^{-1}.v_j = \lambda_j^{-1} v_j$.

(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 $B.v_j = \sqrt{\lambda_j} v_j$. Then $B^2 = A$ is obvious and since all λ_j are real and the v_j are orthonormal, B is positive definite because

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

for $\sum_j \mu_j v_j \neq 0$. 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 I.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

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

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 $m^{-1}: \mathrm{GL}_n(\mathbb{K}) \to \mathrm{U}_n(\mathbb{K}) \times \mathrm{Pd}_n(\mathbb{K})$. Let $g_m = u_m p_m \to g = up$. We have to show that $u_m \to u$ and $p_m \to p$. Since $\mathrm{U}_n(\mathbb{K})$ is compact, the sequence (u_m) has a subsequence (u_{m_k}) converging to some $w \in \mathrm{U}_n(\mathbb{K})$. Then $p_{m_k} = u_{m_k}^{-1} g_{m_k} \to w^{-1}g =: q \in \mathrm{Herm}_n(\mathbb{K})$ and $g = wq$. For each $v \in \mathbb{K}^n$ we then have

$$0 \leq \langle p_{m_k}.v, v\rangle \to \langle q.v, v\rangle,$$

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_m) converges to u, the sequence itself converges to u (Exercise I.1.8), and therefore $p_m = u_m^{-1} g_m \to u^{-1}g = p$. ■

We shall see later that the set $\mathrm{Pd}_n(\mathbb{K})$ is homeomorphic to a vector space, 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 I.1.5. (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. 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 $u.v_j = \lambda_j v$ 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 eigenvectors.

Let $A \in M_n(\mathbb{R})$ and consider it as an element on $M_n(\mathbb{C})$. 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 $W_1 \subseteq \mathbb{R}^n$ of $\dim W_1 \in \{1, 2\}$. Then

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

is a subspace of dimension $n - \dim W_1$ which is also invariant (Exercise I.1.13), 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

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

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

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

If $\det r = -1$, then 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 I.1.12), 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:

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

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 I.1.6. (a) *The group* $\mathrm{U}_n(\mathbb{C})$ *is arcwise connected.*

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

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

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 I.5(a)). Let $\lambda_1, \ldots, \lambda_n$ denote the corresponding eigenvectors. 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 open because det is continuous. Therefore $\mathrm{O}_n(\mathbb{R})$ is not connected and hence not arcwise connected. If we show 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 connected.

Let $g \in \mathrm{SO}_n(\mathbb{R})$. In the normal form of g discussed in Remark I.1.5, 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 of $\mathbb{R}^n$ the linear map defined by g has a matrix of the form

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

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

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

■

Corollary I.1.7. *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 $\mathrm{GL}_n(\mathbb{K})$ shows that

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

as topological spaces (this means that $\cong$ stands for a topological isomorphism, i.e., a homeomorphism). Since $\mathrm{Pd}_n(\mathbb{K})$ is an open convex set, it is arcwise connected (Exercise I.1.5). 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 I.1.6. ■

Normal subgroups of $\mathrm{GL}_n(\mathbb{K})$

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

Definition I.1.8. 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 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 $G := G_1 \times \ldots \times G_n$ 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 we have $G = G_1 \cdots G_n$. ■

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

Proposition I.1.9. (a) *For any field* $\mathbb{K}$ *we have* $Z(\mathrm{GL}_n(\mathbb{K})) = \mathbb{K}^\times \mathbf{1}$.
(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 non-zero entry 1 in position (i,j). For $i \neq j$ we then have $E_{ij}^2 = 0$, so that $(\mathbf{1} + E_{ij})(\mathbf{1} - E_{ij}) = \mathbf{1}$ 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 $k, l \in \{1, \ldots, n\}$ we get

$$g_{ki}\delta_{jl} = (gE_{ij})_{kl} = (E_{ij}g)_{kl} = \delta_{ik}g_{jl}.$$

For $k = i$ and $l = j$ we obtain $g_{ii} = g_{jj}$ and for $k = j = l$, 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

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

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

Remark I.1.10. The subgroups

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

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 $\mathrm{GL}_n(\mathbb{K})$. ■

Exercises for Section I.1.

Exercise I.1.1. (a) 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.

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

$$B.V^\lambda(A) \subseteq V^\lambda(A)$$

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

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

(c) 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.

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

(e)* Find a commuting set $\mathcal{A}$ of diagonalizable endomorphisms of a vector space V which cannot be simultaneously diagonalized. ∎

Exercise I.1.2. $\mathrm{SO}_n(\mathbb{K})$ is a closed normal subgroup of $\mathrm{O}_n(\mathbb{K})$ of index 2. For every $g \in \mathrm{O}_n(\mathbb{K})$ with $\det(g) = -1$, we have

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

Exercise I.1.3. For each subset $M \subseteq M_n(\mathbb{K})$ the *centralizer*

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

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

Exercise I.1.4. We identify $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ by the map $z = x + iy \mapsto (x, y)$ and write I for the real linear map $x \mapsto ix, \mathbb{R}^{2n} \to \mathbb{R}^{2n}$. Then

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

is a closed subgroup of $\mathrm{GL}_{2n}(\mathbb{R})$.

Exercise I.1.5. A subset C of a real vector space V is called a *convex cone* if C is convex and for each $\lambda > 0$ we have $\lambda C \subseteq C$.

Show that $\mathrm{Pd}_n(\mathbb{K})$ is an open convex cone in $\mathrm{Herm}_n(\mathbb{K})$. Hint: The verification of the convexity is easy. To see that $\mathrm{Pd}_n(\mathbb{K})$ is open, show first that for each $r > 0$ we have $B_r(r\mathbf{1}) = rB_1(\mathbf{1}) \subseteq \mathrm{Pd}_n(\mathbb{K})$ (here $B_r(x)$ denotes the open ball of radius r around x) by considering the eigenvalues and using that for $A \in \mathrm{Herm}_n(\mathbb{K})$ we have

$$\|A\| = \max\{|\lambda| \colon \det(A - \lambda\mathbf{1}) = 0\}.$$

Now observe that $\mathrm{Pd}_n(\mathbb{K}) = \bigcup_{r>0} B_r(r\mathbf{1})$ because for $A \in \mathrm{Pd}_n(\mathbb{K})$ with maximal eigenvalue r we have $A \in B_r(r\mathbf{1})$.

Exercise I.1.6. 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 $\gamma(\pi) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ and $\mathrm{im}\,\gamma = \mathrm{SO}_2(\mathbb{R})$. ∎

Exercise I.1.7. Show that the group $\mathrm{O}_n(\mathbb{C})$ is homeomorphic to the topological product of the subgroup

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

Hint: Show that for $g \in \mathrm{O}_n(\mathbb{C})$ with polar decomposition $g = up$ both components u and p are contained in $\mathrm{O}_n(\mathbb{C})$. Compare the polar decomposition of $(g^\top)^{-1}$ and g. Why is $(p^\top)^{-1} \in \mathrm{Pd}_n(\mathbb{C})$? ■

Exercise I.1.8. Let (X, d) be a compact metric space and $(x_n)_{n\in\mathbb{N}}$ a sequence in X. Show that $\lim_{n\to\infty} x_n$ is equivalent to the condition that each convergent subsequence $(x_{n_k})_{k\in\mathbb{N}}$ converges to x. Hint: For metric spaces compactness is equivalent to sequential compactness, which means that every sequence has a convergent subsequence. ■

Exercise I.1.9. If $A \in \mathrm{Herm}_n(\mathbb{K})$ satisfies $\langle A.v, v\rangle = 0$ for each $v \in \mathbb{K}^n$, then $A = 0$. Hint: The hermitian form $b(x, y) := \langle A.x, y\rangle$ satisfies the *polarization identity*

$$b(x,y) = \tfrac{1}{4}\big(b(x+y,x+y) - b(x-y,x-y) + ib(x+iy,x+iy) - ib(x-iy,x-iy)\big)$$

for $\mathbb{K} = \mathbb{C}$, and for $\mathbb{K} = \mathbb{R}$ we have

$$b(x,y) = \tfrac{1}{4}\big(b(x+y,x+y) - b(x-y,x-y)\big).$$

Exercise I.1.10. Show that for a complex matrix $A \in M_n(\mathbb{C})$ the following are equivalent:
(1) $A^* = A$.
(2) $\langle A.v, v\rangle \in \mathbb{R}$ for each $v \in \mathbb{C}^n$.
Hint: Write $A = B + iC$ with B, C hermitian and use Exercise I.1.9 to show that $C = 0$ if (2) holds. ■

Exercise I.1.11. (a) Show that a matrix $A \in \mathrm{Herm}_n(\mathbb{K})$ is hermitian if and only if there exists an orthonormal basis $v_1, \ldots, v_n$ of $\mathbb{K}^n$ 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 $v_1, \ldots, v_n$ of $\mathbb{K}^n$ 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 $v_1, \ldots, v_n$ of $\mathbb{K}^n$ and $\lambda_j \in \mathbb{C}$ with $Av_j = \lambda_j v_j$. ■

Exercise I.1.12. (a) Let V be a vector space and $E \in \mathrm{End}(V)$ with $A^2 = \mathbf{1}$. 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 $\lambda_1, \ldots, \lambda_n$. Hint: The subalgebra $\mathbb{K}[A] := \mathrm{span}\{A^k : k \in \mathbb{N}_0\} \subseteq \mathrm{End}(V)$ is isomorphic to $\mathbb{K}^n$ with pointwise multiplication and the basis vectors $e_j \in \mathbb{K}^n$ correspond to the projections onto the eigenspace of A. ■

Exercise I.1.13. Let $\beta: V \times V \to \mathbb{K}$ be a bilinear map and $g \in \mathrm{O}(V, \beta)$ an isometry. For a subspace $E \subseteq V$ we write

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

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

I.2. Groups and geometry

In Definition I.1.2 we have defined certain matrix groups by concrete conditions on the matrices. If we think of matrices as linear maps, described with respect to a basis, 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.

Definition I.2.1. (Groups and bilinear forms)

(a) (The abstract general linear group) Let V be a $\mathbb{K}$-vector space, where $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. 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 of V, then the map

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

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

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

(b) Let V be an n-dimensional vector space with basis $v_1, \ldots, v_n$ and $\beta: V \times V \to \mathbb{K}$ 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

$$\mathrm{O}(V,\beta) := \{g \in \mathrm{GL}(V) \colon (\forall v, w \in V)\, \beta(g.v, g.w) = \beta(v,w)\}$$

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

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

If $v_1, \ldots, v_n$ is an orthonormal basis for β, i.e., $B = \mathbf{1}$, then

$$\Phi^{-1}(\mathrm{O}(V,\beta)) = \mathrm{O}_n(\mathbb{K})$$

is the orthogonal group defined in Section I.1. Note that orthonormal bases can only exist for symmetric bilinear forms (Why?).

For $V = \mathbb{K}^{2n}$ and the block (2×2)-matrix

$$B := \begin{pmatrix} \mathbf{0} & \mathbf{1}_n \\ -\mathbf{1}_n & \mathbf{0} \end{pmatrix}$$

we see that $B^\top = -B$, and the group

$$\mathrm{Sp}_{2n}(\mathbb{K}) := \{g \in \mathrm{GL}_{2n}(\mathbb{K}) \colon g^\top B g = B\}$$

is called the *symplectic group*. The corresponding skew-symmetric bilinear form on $\mathbb{K}^{2n}$ is given by

$$\beta(x,y) = x^\top B y = \sum_{i=1}^n x_i y_{n+i} - x_{n+i} y_i.$$

(c) A symmetric bilinear form β on V is called *non-degenerate* if $\beta(v, V) = \{0\}$ implies $v = 0$. For $\mathbb{K} = \mathbb{C}$ every non-degenerate symmetric bilinear form β possesses an orthonormal basis (this builds on the existence of square roots of non-zero complex numbers; see Exercise I.2.1), so that for every such form β we get

$$\mathrm{O}(V,\beta) \cong \mathrm{O}_n(\mathbb{C}).$$

For $\mathbb{K} = \mathbb{R}$ the situation is more complicated, since negative real numbers do not have a square root in $\mathbb{R}$, there might not be an orthonormal basis, but if β is non-degenerate, there always exists an orthogonal basis $v_1, \ldots, v_n$ and $p \in \{1, \ldots, n\}$ such that $\beta(v_j, v_j) = 1$ for $j = 1, \ldots, p$ and $\beta(v_j, v_j) = -1$ for $j = p+1, \ldots, n$. Let $q := n - p$ and $I_{p,q}$ denote the corresponding matrix

$$I_{p,q} = \begin{pmatrix} \mathbf{1}_p & 0 \\ 0 & -\mathbf{1}_q \end{pmatrix} \in M_{p+q}(\mathbb{R}).$$

Then $\mathrm{O}(V, \beta)$ is isomorphic to the group

$$\mathrm{O}_{p,q}(\mathbb{R}) := \{g \in \mathrm{GL}_n(\mathbb{R})\colon g^\top I_{p,q} g = I_{p,q}\},$$

where $\mathrm{O}_{n,0}(\mathbb{R}) = \mathrm{O}_n(\mathbb{R})$.

(d) Let V be an n-dimensional complex vector space and $\beta\colon V \times V \to \mathbb{C}$ a sesquilinear form, i.e., β is linear in the first and antilinear in the second argument. Then we also choose a basis $v_1, \ldots, v_n$ in V and define $B = (b_{jk}) := (\beta(v_j, v_k))_{j,k=1,\ldots,n}$, but now we obtain

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

We write

$$\mathrm{U}(V, \beta) := \{g \in \mathrm{GL}(V)\colon (\forall v, w \in V)\, \beta(g.v, g.w) = \beta(v, w)\}$$

for the corresponding *unitary group* and find

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

If $v_1, \ldots, v_n$ is an orthonormal basis for β, i.e., $B = \mathbf{1}$, then

$$\Phi^{-1}(\mathrm{U}(V, \beta)) = \mathrm{U}_n(\mathbb{C}) = \{g \in \mathrm{GL}_n(\mathbb{C})\colon g^* = g^{-1}\}$$

is the unitary group over $\mathbb{C}$. We call β *hermitian* if it is sesquilinear and satisfies $\beta(y, x) = \overline{\beta(x, y)}$. In this case one has to face the same problems as for symmetric forms on real vector spaces, but there always exists an orthogonal basis $v_1, \ldots, v_n$ and $p \in \{1, \ldots, n\}$ with $\beta(v_j, v_j) = 1$ for $j = 1, \ldots, p$ and $\beta(v_j, v_j) = -1$ for $j = p+1, \ldots, n$. With $q := n - p$ and

$$I_{p,q} = \begin{pmatrix} \mathbf{1}_p & 0 \\ 0 & -\mathbf{1}_q \end{pmatrix} \in M_n(\mathbb{C})$$

we then define the *indefinite unitary groups* by

$$\mathrm{U}_{p,q}(\mathbb{C}) := \{g \in \mathrm{GL}_n(\mathbb{C})\colon g^\top I_{p,q} \overline{g} = I_{p,q}\}.$$

Since $I_{p,q}$ has real entries, we have

$$\mathrm{U}_{p,q}(\mathbb{C}) = \{g \in \mathrm{GL}_n(\mathbb{C})\colon g^* I_{p,q} g = I_{p,q}\},$$

where $\mathrm{U}_{n,0}(\mathbb{C}) = \mathrm{U}_n(\mathbb{C})$. ∎

Definition I.2.2. (a) Let V be a vector space. We consider the *affine group* $\mathrm{Aff}(V)$ of all maps $V \to V$ of the form

$$\varphi_{v,g}(x) = gx + v, \quad g \in \mathrm{GL}(V), v \in V.$$

We write elements $\varphi_{(v,g)}$ of $\mathrm{Aff}(V)$ simply as pairs (v,g). Then the composition in $\mathrm{Aff}(V)$ is given by

$$(v,g)(w,h) = (v + g.w, gh),$$

$(0,\mathbf{1})$ is the identity, and inversion is given by

$$(v,g)^{-1} = (-g^{-1}.v, g^{-1}).$$

For $V = \mathbb{K}^n$ we put $\mathrm{Aff}_n(\mathbb{K}) := \mathrm{Aff}(\mathbb{K}^n)$. Then the map

$$\Phi\colon \mathrm{Aff}_n(\mathbb{K}) \to \mathrm{GL}_{n+1}(\mathbb{K}), \quad \Phi(v,g) = \begin{pmatrix} [g] & v \\ 0 & 1 \end{pmatrix}$$

is an injective group homomorphism, where $[g]$ denotes the matrix of the linear map with respect to the canonical basis of $\mathbb{K}^n$.

(b) (The euclidean isometry group) Let $V = \mathbb{R}^n$ and consider the euclidean metric $d(x,y) := \|x - y\|_2$ on $\mathbb{R}^n$. We define

$$\mathrm{Iso}_n(\mathbb{R}) := \{g \in \mathrm{Aff}(\mathbb{R}^n)\colon (\forall x, y \in V)\ d(g.v, g.w) = d(v,w)\}.$$

This is the group of *affine isometries* of the euclidean n-space. Actually one can show that every isometry of a normed space $(V, \|\cdot\|)$ is an affine map (Exercise I.2.5). This implies that

$$\mathrm{Iso}(\mathbb{R}^n) = \{g\colon \mathbb{R}^n \to \mathbb{R}^n\colon (\forall x, y \in \mathbb{R}^n)\ d(g.v, g.w) = d(v,w)\}. \qquad \blacksquare$$

We have seen in Definition I.1.8 how to form direct products of groups. If $G = G_1 \times G_2$ is a direct product of the groups G_1 and G_2, then we identify G_1 and G_2 with the corresponding subgroups of $G_1 \times G_2$, i.e., we identify $g_1 \in G_1$ with (g_1, e) and $g_2 \in G_2$ with (e, g_2). Then G_1 and G_2 are normal subgroups of G and the product map

$$m\colon G_1 \times G_2 \to G, \quad (g_1, g_2) \mapsto g_1 g_2 = (g_1, g_2)$$

is a group isomorphism, i.e., each element $g \in G$ has a unique decomposition $g = g_1 g_2$ with $g_1 \in G_1$ and $g_2 \in G_2$.

The affine group $\mathrm{Aff}(V)$ has a structure which is similar. The translation group $V \cong \{(v,\mathbf{1})\colon v \in V\}$ and the linear group $\mathrm{GL}(V) \cong \{(0,g)\colon g \in \mathrm{GL}(V)\}$ are subgroups, and each element (v,g) has a unique representation as a product $(v,\mathbf{1})(0,g)$, but in this case $\mathrm{GL}(V)$ is not a normal subgroup, whereas V is normal. The following lemma introduces a concept that is important to understand the structure of groups which have similar decompositions.

In the following we write $\mathrm{Aut}(G)$ for the set of automorphisms of the group G and note that this set is a group under composition of maps. In particular the inverse of a group automorphism is an automorphism.

Lemma I.2.3. (a) *Let N and H be groups, write* $\mathrm{Aut}(N)$ *for the group of all automorphisms of N, and suppose that $\delta\colon H \to \mathrm{Aut}(N)$ is a group homomorphism. Then we define a multiplication on $N \times H$ by*

$$(n,h)(n',h') := (n\delta(h)(n'), hh'). \tag{2.1}$$

This multiplication turns $N \times H$ into a group denoted by $N \rtimes_\delta H$, where $N \cong N \times \{e\}$ is a normal subgroup, $H \cong \{e\} \times H$ is a subgroup, and each element $g \in N \rtimes_\delta H$ has a unique representation as $g = nh$, $n \in N$, $h \in H$.

(b) *If, conversely, G is a group, $N \trianglelefteq G$ a normal subgroup and $H \subseteq G$ a subgroup with the property that the multiplication map $m\colon N \times H \to G$ is bijective, i.e., $NH = G$ and $N \cap H = \{e\}$, then*

$$\delta\colon H \to \mathrm{Aut}(N), \quad \delta(h)(n) := hnh^{-1} \tag{2.2}$$

is a group homomorphism, and the map

$$m\colon N \rtimes_\delta H \to G, \quad (n,h) \mapsto nh$$

is a group isomorphism.

Proof. (a) We have to verify the associativity of the multiplication and the existence of an inverse. The associativity follows from

$$\begin{aligned}
&\big((n,h)(n',h')\big)(n'',h'') \\
&= (n\delta(h)(n'), hh')(n'',h'') = (n\delta(h)(n')\delta(hh')(n''), hh'h'') \\
&= (n\delta(h)(n')\delta(h)\big(\delta(h')(n'')\big), hh'h'') = (n\delta(h)\big(n'\delta(h')(n'')\big), hh'h'') \\
&= (n,h)(n'\delta(h')(n''), h'h'') = (n,h)\big((n',h')(n'',h'')\big).
\end{aligned}$$

With (2.1) we immediately get the formula for the inverse

$$(n,h)^{-1} = (\delta(h^{-1})(n^{-1}), h^{-1}). \tag{2.2}$$

(b) Since

$$\delta(h_1h_2)(n) = h_1h_2n(h_1h_2)^{-1} = h_1(h_2nh_2^{-1})h_1^{-1} = \delta(h_1)\delta(h_2)(n),$$

the map $\delta\colon H \to \mathrm{Aut}(N)$ is a group homomorphism. Moreover, the multiplication map m satisfies

$$m(n,h)m(n',h') = nhn'h' = (nhn'h^{-1})hh' = m((n,h)(n',h')),$$

hence is a group homomorphism. It is bijective by assumption. ■

Definition I.2.4. The group $N \rtimes_\delta H$ constructed in Lemma I.2.3 from the data (N, H, δ) is called the *semidirect product* of N and H with respect to δ. If it is clear from the context what δ is, then we simply write $N \rtimes H$ instead of $N \rtimes_\delta H$. ■

If δ is trivial, i.e., $\delta(h) = \mathrm{id}_N$ for each $h \in H$, then $N \rtimes_\delta H \cong N \times H$ is a direct product. In this sense semidirect products generalize direct products. Below we shall see several concrete examples of groups which can most naturally be described as semidirect products of known groups.

One major point in studying semidirect products is that for any normal subgroup $N \trianglelefteq G$, we think of the groups N and G/N as building blocks of the group G. For each semidirect product $G = N \rtimes H$ we have $G/N \cong H$, so that the two building blocks N and $G/N \cong H$ are the same, although the groups might be quite different, f.i. $\mathrm{Aff}(V)$ and $V \times \mathrm{GL}(V)$ are very different groups: In the latter group $N = V \times \{0\}$ is a central subgroup and in the first group it is not. On the other hand there are situations where G cannot be build from N and $H := G/N$ as a semidirect product. This works if and only if there exists a group homomorphism $\sigma\colon G/N \to G$ with $\sigma(gN) \in gN$ for each $g \in G$. An example where such a homomorphism does not exist is

$$G = C_4 := \{z \in \mathbb{C}^\times : z^4 = 1\} \quad \text{and} \quad N := C_2 := \{z \in \mathbb{C}^\times : z^2 = 1\} \trianglelefteq G.$$

In this case $G \not\cong N \rtimes H$ for any group H because then $H \cong G/N \cong C_2$, so that the fact that G is abelian would lead to $G \cong C_2 \times C_2$, contradicting the existence of elements of order 4 in G.

Example I.2.5. (a) We know already the following examples of semidirect products from Definition I.2.2. The affine group $\mathrm{Aff}(V)$ of a vector space is isomorphic to the semidirect product

$$\mathrm{Aff}(V) \cong V \rtimes_\delta \mathrm{GL}(V), \quad \delta(g)(v) = gv.$$

Similarly we have

$$\mathrm{Aff}_n(\mathbb{R}) \cong \mathbb{R}^n \rtimes_\delta \mathrm{GL}_n(\mathbb{R}), \quad \delta(g)(v) = gv.$$

We furthermore have the subgroup $\mathrm{Iso}_n(\mathbb{R})$, which, in view of

$$\mathrm{O}_n(\mathbb{R}) = \{g \in \mathrm{GL}_n(\mathbb{R}) \colon (\forall x \in \mathbb{R}^n)\|g.x\| = \|x\|\}$$

(cf. Exercise I.2.6) satisfies

$$\mathrm{Iso}(\mathbb{R}^n) \cong \mathbb{R}^n \rtimes \mathrm{O}_n(\mathbb{R}).$$

The group of *euclidean motions of* $\mathbb{R}^n$ is the subgroup

$$\mathrm{Mot}_n(\mathbb{R}) := \mathbb{R}^n \rtimes \mathrm{SO}_n(\mathbb{R})$$

of those isometries preserving orientation.

(b) For each group G we can form the semidirect product group

$$G \rtimes_\delta \operatorname{Aut}(G), \quad \delta(\varphi)(g) = \varphi(g). \qquad \blacksquare$$

Example I.2.6. (The concrete Galilei group[1]) We consider the vector space

$$M := \mathbb{R}^4 \cong \mathbb{R}^3 \times \mathbb{R}$$

as the space of pairs (q,t) describing *events* in a four-dimensional (non-relativistic) *spacetime*. Here q stands for the spatial coordinate of the event and t for the (absolute) time of the event. The set M is called *Galilei spacetime*. There are three types of symmetries of this spacetime:

(1) The special Galilei transformations:

$$G_v \colon \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3 \times \mathbb{R}, \quad (q,t) \mapsto (q + vt, t) = \begin{pmatrix} \mathbf{1} & v \\ 0 & 1 \end{pmatrix} \begin{pmatrix} q \\ t \end{pmatrix},$$

describing movements with constant velocity v.

(2) Rotations:

$$\mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3 \times \mathbb{R}, \quad (q,t) \mapsto (Aq, t), \quad A \in \mathrm{SO}_3(\mathbb{R}),$$

(3) Space translations

$$T_v \colon \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3 \times \mathbb{R}, \quad (q,t) \mapsto (q + v, t),$$

and time translations

$$T_\beta \colon \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3 \times \mathbb{R}, \quad (q,t) \mapsto (q, t + \beta).$$

All these maps are affine maps on $\mathbb{R}^4$. The subgroup $\Gamma \subseteq \mathrm{Aff}(4,\mathbb{R})$ generated by the maps in (1), (2) and (3) is called the *proper (orthochrone) Galilei group*. The full *Galilei group* Γ_{ext} is obtained if we add the time reversion $T(q,t) := (q,-t)$ and the space reflection $S(q,t) := (-q,t)$. Both are not contained in Γ.

Roughly stated, *Galilei's relativity principle* states that *the basic physical laws of closed systems are invariant under transformations of the proper Galilei group* (see [Sch95, Sect. II.2] for more information on this perspective). It means that Γ is the natural symmetry group of non-relativistic mechanics.

[1] Galileo Galilei, 1564–1642, was an italian mathematician and philosopher. He held professorships in Pisa and Padua, later he was mathematician and -philosopher at the court in Florence. The Galilei group is the symmetry group of non-relativistic kinematics in three dimensions.

To describe the structure of the group Γ, we first observe that by (3) it contains the subgroup $\Gamma_t \cong (\mathbb{R}^4, +)$ of all spacetime translations. The maps under (1) and (2) are linear maps on $\mathbb{R}^4$. They generate the group

$$\Gamma_\ell := \{(v, A) \colon A \in \mathrm{SO}_3(\mathbb{R}), v \in \mathbb{R}^3\},$$

where we write (v, A) for the affine map given by $(q, t) \mapsto (Aq + vt, t)$. The composition of two such maps is given by

$$(v, A).\big((v', A').(q, t)\big) = (A(A'q + v't) + vt, t) = (AA'q + (Av' + v)t, t),$$

so that the product in Γ_ℓ is

$$(v, A)(v', A') = (v + Av', AA').$$

We conclude that

$$\Gamma_\ell \cong \mathbb{R}^3 \rtimes \mathrm{SO}_3(\mathbb{R})$$

is isomorphic to the group $\mathrm{Mot}_3(\mathbb{R})$ of motions of euclidean space. We thus obtain the description

$$\Gamma \cong \mathbb{R}^4 \rtimes (\mathbb{R}^3 \rtimes \mathrm{SO}_3(\mathbb{R})) \cong \mathbb{R}^4 \rtimes \mathrm{Mot}_3(\mathbb{R}),$$

where $\mathrm{Mot}_3(\mathbb{R})$ acts on $\mathbb{R}^4$ by $(v, A).(q, t) := (Aq + vt, t)$, which corresponds to the natural embedding $\mathrm{Aff}_3(\mathbb{R}) \to \mathrm{GL}_4(\mathbb{R})$ discussed in Example I.2.2.

For the extended Galilei group one easily obtains

$$\Gamma_{\mathrm{ext}} \cong \Gamma \rtimes \{S, T, ST, \mathbf{1}\} \cong \Gamma \rtimes (C_2 \times C_2),$$

because the group $\{S, T, ST, \mathbf{1}\}$ generated by S and T is a four element group intersecting the normal subgroup Γ trivially. Therefore the description as a semidirect product follows from the second part of Lemma I.2.3. ■

Example I.2.7. (The concrete Poincaré group) In the preceding example we have viewed four-dimensional spacetime as a product of space $\mathbb{R}^3$ with time $\mathbb{R}$. This picture changes if one wants to incorporate special relativity. Here the underlying spacetime is *Minkowski space*, which is $M = \mathbb{R}^4$, endowed with the *Lorentz form*

$$\beta(x, y) := x_1 y_1 + x_2 y_2 + x_3 y_3 - x_4 y_4.$$

The group

$$L := \mathrm{O}_{3,1}(\mathbb{R}) \cong \mathrm{O}(\mathbb{R}^4, \beta)$$

is called the *Lorentz group*. This is the symmetry group of relativistic (classical) mechanics.

The Lorentz group has several important subgroups:

$$L_+ := \mathrm{SO}_{3,1}(\mathbb{R}) := L \cap \mathrm{SL}_4(\mathbb{R}) \quad \text{and} \quad L^\uparrow := \{g \in L \colon g_{44} \leq -1\}.$$

The condition $g_{44} \geq 1$ comes from the observation that for $e_4 = (0,0,0,1)^\top$ we have

$$-1 = \beta(e_4, e_4) = \beta(g.e_4, g.e_4) = g_{14}^2 + g_{24}^2 + g_{34}^2 - g_{44}^2,$$

so that $g_{44}^2 \geq 1$. Therefore either $g_{44} \geq 1$ or $g_{44} \leq -1$. To understand geometrically why $L^\uparrow$ is a subgroup, we consider the quadratic form

$$q(x) := \beta(x,x) = x_1^2 + x_2^2 + x_3^2 - x_4^2$$

on $\mathbb{R}^4$. Since q is invariant under L, the action of the group L on $\mathbb{R}^4$ preserves the double cone

$$C := \{x \in \mathbb{R}^4 \colon q(x) \leq 0\} = \{x \in \mathbb{R}^4 \colon |x_4| \geq \|(x_1, x_2, x_3)\|\}.$$

Let

$$C_\pm := \{x \in C \colon \pm x_4 \geq 0\} = \{x \in \mathbb{R}^4 \colon \pm x_4 \geq \|(x_1, x_2, x_3)\|\}.$$

Then $C = C_+ \cup C_-$ with $C_+ \cap C_- = \{0\}$ and the sets $C_\pm$ are both convex cones, as follows easily from the convexity of the norm function on $\mathbb{R}^3$ (Exercise). Each element $g \in L$ preserves the set $C \setminus \{0\}$ which has the two arc-components $C_\pm \setminus \{0\}$. The continuity of the map $g \colon C \setminus \{0\} \to C \setminus \{0\}$ now implies that we have two possibilities. Either $g.C_+ = C_+$ or $g.C_+ = C_-$. In the first case we have $g_{44} \geq 1$ and in the latter case we have $g_{44} \leq -1$.

In the physical literature one sometimes finds $\mathrm{SO}_{3,1}(\mathbb{R})$ as the notation for $L_+^\uparrow := L_+ \cap L^\uparrow$, which is inconsistent with the standard notation for matrix groups.

The *(proper) Poincaré group* is the corresponding affine group

$$P := \mathbb{R}^4 \rtimes L_+^\uparrow.$$

This group is the identity component of the *inhomogeneous Lorentz group* $\mathbb{R}^4 \rtimes L$. Some people use the name Poincaré group only for the universal covering group $\widetilde{P}$ of P which is isomorphic to $\mathbb{R}^4 \rtimes \mathrm{SL}_2(\mathbb{C})$, as we shall see below in Example V.4.6(3).

The topological structure of the Poincaré- and Lorentz group will be discussed after Chapter II, when we have refined information on the polar decomposition obtained from the exponential function. Then we shall see that the Lorentz group L has four arc-components

$$L_+^\uparrow, \quad L_+^\downarrow, \quad L_-^\uparrow \quad \text{and} \quad L_-^\downarrow,$$

where

$$L_\pm := \{g \in L \colon \det g = \pm 1\}, \quad L^\downarrow := \{g \in L \colon g_{44} \leq -1\}$$

and

$$L_\pm^\uparrow := L_\pm \cap L^\uparrow, \quad L_\pm^\downarrow := L_\pm \cap L^\downarrow. \qquad \blacksquare$$

Exercises for Section I.2.

Exercise I.2.1. (a) Let β be a symmetric bilinear form on a finite-dimensional complex vector space V. Show that there exists an orthogonal basis $v_1, \ldots, v_n$ with $\beta(v_j, v_j) = 1$ for $j = 1, \ldots, p$ and $\beta(v_j, v_j) = 0$ for $j > p$. Hint: Use induction on $\dim V$. If $\beta \neq 0$, then there exists $v_1 \in V$ with $\beta(v_1, v_1) = 1$ (polarization identity). Now proceed with the space $v_1^\perp := \{v \in V : \beta(v_1, v) = 0\}$.

(b) Show that each invertible symmetric matrix $B \in \mathrm{GL}_n(\mathbb{C})$ can be written as $B = AA^\top$ for some $A \in \mathrm{GL}_n(\mathbb{C})$. Hint: Consider the symmetric bilinear form $\beta(x, y) = x^\top By$. ■

Exercise I.2.2. Let β be a symmetric bilinear form on a finite-dimensional real vector space V. Show that there exists an orthogonal basis $v_1, \ldots, v_n$ with $\beta(v_j, v_j) = 1$ for $j = 1, \ldots, p$, $\beta(v_j, v_j) = -1$ for $j = p+1, \ldots, q$, and $\beta(v_j, v_j) = 0$ for $j > q$. ■

Exercise I.2.3. Let β be a skew-symmetric bilinear form on a finite-dimensional vector space V which is non-degenerate in the sense that $\beta(v, V) = \{0\}$ implies $v = 0$. Show that there exists a basis $v_1, \ldots, v_n$, $w_1, \ldots, w_n$ of V with

$$\beta(v_i, w_j) = \delta_{ij} \quad \text{and} \quad \beta(v_i, v_j) = \beta(w_i, w_j) = 0.$$

Hint: Pick $v_1 \in V \setminus \{0\}$ and find $w_1 \in V$ with $\beta(v_1, w_1) = 1$. Then consider the restriction β_1 of β to the subspace

$$V_1 := \{v_1, w_1\}^\perp = \{x \in V : \beta(x, v_1) = \beta(x, v_2) = 0\}$$

and argue by induction. Why is β_1 non-degenerate? ■

Exercise I.2.4. (Metric characterization of midpoints) Let $(X, \|\cdot\|)$ be a normed space and $x, y \in X$ distinct points. Let

$$M_0 := \{z \in X : \|z - x\| = \|z - y\| = \tfrac{1}{2}\|x - y\|\} \quad \text{and} \quad m := \frac{x + y}{2}.$$

For a subset $A \subseteq X$ we define its *diameter*

$$\delta(A) := \sup\{\|a - b\| : a, b \in A\}.$$

Show that:

(1) If X is a pre-Hilbert space (i.e., a vector space with a hermitian scalar product), then $M_0 = \{m\}$ is a one-element set.

(2) For $z \in M_0$ we have $\|z - m\| \leq \frac{1}{2}\delta(M_0) \leq \frac{1}{2}\|x - y\|$.

(3) For $n \in \mathbb{N}$ we define inductively:

$$M_n := \{p \in M_{n-1} : (\forall z \in M_{n-1})\ \|z - p\| \leq \tfrac{1}{2}\delta(M_{n-1})\}.$$

Then we have for each $n \in \mathbb{N}$:
(a) M_n is a convex set.
(b) M_n is invariant under the point reflection $s_m(a) := 2m - a$ in m.
(c) $m \in M_n$.
(d) $\delta(M_n) \leq \frac{1}{2}\delta(M_{n-1})$.

(4) $\bigcap_{n\in\mathbb{N}} M_n = \{m\}$. ■

Exercise I.2.5. (Isometries of normed spaces are affine maps) Let $(X, \|\cdot\|)$ be a normed space endowed with the metric $d(x,y) := \|x-y\|$. Show that each isometry $\varphi: (X,d) \to (X,d)$ is an affine map by using the following steps:

(1) It suffices to assume that $\varphi(0) = 0$ and to show that this implies that φ is a linear map.
(2) $\varphi(\frac{x+y}{2}) = \frac{1}{2}(\varphi(x) + \varphi(y))$ for $x, y \in X$. Hint: Exercise I.2.4.
(3) φ is continuous.
(4) $\varphi(\lambda x) = \lambda\varphi(x)$ for $\lambda \in 2^{\mathbb{Z}} \subseteq \mathbb{R}$.
(5) $\varphi(x+y) = \varphi(x) + \varphi(y)$ for $x, y \in X$.
(6) $\varphi(\lambda x) = \lambda\varphi(x)$ for $\lambda \in \mathbb{R}$. ■

Exercise I.2.6. Let $\beta: V \times V \to V$ be a symmetric bilinear form on the vector space V and

$$q: V \to V, \quad v \mapsto \beta(v,v)$$

the corresponding quadratic form. Then for $\varphi \in \mathrm{End}(V)$ the following are equivalent:

(1) $(\forall v \in V)\ q(\varphi(v)) = q(v)$.
(2) $(\forall v, w \in V)\ \beta(\varphi(v), \varphi(w)) = \beta(v,w)$.

Hint: Use the polarization identity $\beta(v,w) = \frac{1}{4}\big(q(v+w) - q(v-w)\big)$. ■

Exercise I.2.7. We consider $\mathbb{R}^4 = \mathbb{R}^3 \times \mathbb{R}$, where the elements of $\mathbb{R}^4$ are considered as space time events (q,t), $q \in \mathbb{R}^3$, $t \in \mathbb{R}$. On $\mathbb{R}^4$ we have the linear (time) functional

$$\Delta: \mathbb{R}^4 \to \mathbb{R}, (x,t) \mapsto t$$

and we endow $\ker\Delta \cong \mathbb{R}^3$ with the euclidean scalar product

$$\beta(x,y) := x_1y_1 + x_2y_2 + x_3y_3.$$

Show that

$$H := \{g \in \mathrm{GL}_4(\mathbb{R}): g.\ker\Delta \subseteq \ker\Delta, g|_{\ker\Delta} \in \mathrm{O}_3(\mathbb{R})\} \cong \mathbb{R}^3 \rtimes (\mathrm{O}_3(\mathbb{R}) \times \mathbb{R}^\times)$$

and

$$G := \{g \in H: \Delta \circ g = \Delta\} \cong \mathbb{R}^3 \rtimes \mathrm{O}_3(\mathbb{R}).$$

In this sense the linear part of the Galilei group (extended by the space reflection S) is isomorphic to the symmetry group of the triple $(\mathbb{R}^4, \beta, \Delta)$, where Δ represents a universal time function and β is the scalar product on $\ker\Delta$. In the relativistic picture (Example I.2.7), the time function is combined with the scalar product in the Lorentz form. ■

Exercise I.2.8. On the four-dimensional real vector space $V := \mathrm{Herm}_2(\mathbb{C})$ we consider the symmetric bilinear form β given by

$$\beta(A,B) := \mathrm{tr}(AB) - \mathrm{tr}\,A\,\mathrm{tr}\,B.$$

Show that:
(1) The corresponding quadratic form is given by $q(A) := \beta(A,A) = -2\det A$.
(2) Show that $(V,\beta) \cong \mathbb{R}^{3,1}$ by finding a basis $E_1,\ldots,E_4$ of $\mathrm{Herm}_2(\mathbb{C})$ with

$$q(a_1E_1 + \ldots + a_4E_4) = a_1^2 + a_2^2 + a_3^2 - a_4^2.$$

(3) For $g \in \mathrm{GL}_2(\mathbb{C})$ and $A \in \mathrm{Herm}_2(\mathbb{C})$ the matrix gAg^* is hermitian and satisfies

$$q(gAg^*) = |\det(g)|^2 q(A).$$

(4) For $g \in \mathrm{SL}_2(\mathbb{C})$ we define a linear map $\rho(g) \in \mathrm{GL}(\mathrm{Herm}_2(\mathbb{C}))$ by $\rho(g)(A) := gAg^*$. Then we obtain a homomorphism

$$\rho\colon \mathrm{SL}_2(\mathbb{C}) \to \mathrm{O}(V,\beta) \cong \mathrm{O}_{3,1}(\mathbb{R}).$$

(5) Show that $\ker\rho = \{\pm\mathbf{1}\}$. ■

Exercise I.2.9. Let $\beta\colon V \times V \to \mathbb{K}$ be a bilinear form.
(1) Show that there exists a symmetric bilinear form β_+ and a skew-symmetric bilinear form β_- with $\beta = \beta_+ + \beta_-$.
(2) $\mathrm{O}(V,\beta) = \mathrm{O}(V,\beta_+) \cap \mathrm{O}(V,\beta_-)$. ■

Exercise I.2.10. (a) Let G be a group, $N \subseteq G$ a normal subgroup and $q\colon G \to G/N, g \mapsto gN$ the quotient homomorphism. Show that:
(1) If $G \cong N \rtimes_\delta H$ for a subgroup H, then $H \cong G/N$.
(2) There exists a subgroup $H \subseteq G$ with $G \cong N \rtimes_\delta H$ if and only if there exists a group homomorphism $\sigma\colon G/N \to G$ with $q \circ \sigma = \mathrm{id}_{G/N}$.
(b) Show that

$$\mathrm{GL}_n(\mathbb{K}) \cong \mathrm{SL}_n(\mathbb{K}) \rtimes_\delta \mathbb{K}^\times$$

for a suitable homomorphism $\delta\colon \mathbb{K}^\times \to \mathrm{Aut}(\mathrm{SL}_n(\mathbb{K}))$. Hint: Consider the homomorphism

$$\varphi\colon \mathbb{K}^\times \to \mathrm{GL}_n(\mathbb{K}), \quad \lambda \mapsto \mathrm{diag}(\lambda,1,\ldots,1).$$ ■

Exercise I.2.11. Show that $\mathrm{O}_{p,q}(\mathbb{C}) \cong \mathrm{O}_{p+q}(\mathbb{C})$ for $p,q \in \mathbb{N}_0$, $p+q > 0$. ■