Chapter 10
Exponential of a Matrix, Polar Decomposition, and Classical Groups

Polar decomposition and exponentiation are fundamental tools in the theory of finite-dimensional Lie groups and Lie algebras. We do not consider these notions here in their full generality, but restrict attention to their matricial aspects.

10.1 The Polar Decomposition

The polar decomposition of matrices is defined by analogy with that in the complex plane: if $z \in \mathbb{C}^*$, there exists a unique pair $(r,q) \in (0,+\infty) \times S^1$ (S^1 denotes the unit circle, the set of complex numbers of modulus 1) such that $z = rq$. If z acts on $\mathbb{C}$ (or on $\mathbb{C}^*$) by multiplication, this action can be decomposed as the product of a rotation of angle θ (where $q = \exp(i\theta)$) with a homothety of ratio $r > 0$. The fact that these two actions commute is a consequence of the commutativity of the multiplicative group $\mathbb{C}^*$; this commutation is false for the polar decomposition in $\mathbf{GL}_n(k)$, $k = \mathbb{R}$ or $\mathbb{C}$, because the general linear group is not commutative.

The factors $(0,+\infty)$ and S^1 are replaced by $\mathbf{HPD}_n$, the open cone of matrices of $\mathbf{M}_n(\mathbb{C})$ that are Hermitian positive-definite, and the unitary group $\mathbf{U}_n$. In $\mathbf{M}_n(\mathbb{R})$, we play instead with $\mathbf{SPD}_n$, the set of symmetric positive-definite matrices, and the orthogonal group $\mathbf{O}_n$. The groups $\mathbf{U}_n$ and $\mathbf{O}_n$ are compact, because they are closed and bounded in $\mathbf{M}_n(K = \mathbb{R},\mathbb{C})$. The columns of unitary matrices are unit vectors, so that $\mathbf{U}_n$ is bounded. On the other hand, $\mathbf{U}_n$ is defined by an equation $U^*U = I_n$, where the map $U \mapsto U^*U$ is continuous; hence $\mathbf{U}_n$ is closed. By the same arguments, $\mathbf{O}_n$ is compact.

Theorem 10.1 *For every $M \in \mathbf{GL}_n(\mathbb{C})$, there exists a unique pair*

$$(H,Q) \in \mathbf{HPD}_n \times \mathbf{U}_n$$

such that $M = HQ$. If $M \in \mathbf{GL}_n(\mathbb{R})$, then $(H,Q) \in \mathbf{SPD}_n \times \mathbf{O}_n$.

D. Serre, *Matrices,* Graduate Texts in Mathematics 216,
DOI 10.1007/978-1-4419-7683-3_10,

The map $M \mapsto (H,Q)$*, called the* polar decomposition *of* M*, is a homeomorphism (i.e., a bicontinuous bijection) between* $\mathbf{GL}_n(\mathbb{C})$ *and* $\mathbf{HPD}_n \times \mathbf{U}_n$ *(respectively, between* $\mathbf{GL}_n(\mathbb{R})$ *and* $\mathbf{SPD}_n \times \mathbf{O}_n$*).*

Proof. Existence. Because $MM^* \in \mathbf{HPD}_n$, we can set $H := \sqrt{MM^*}$ (the square root was defined in Section 6.1). Then $Q := H^{-1}M$ satisfies $Q^*Q = M^*H^{-2}M = M^*(MM^*)^{-1}M = I_n$, hence $Q \in \mathbf{U}_n$.

In the real case ($M \in \mathbf{GL}_n(\mathbb{R})$), MM^* is real symmetric. In fact, H is real symmetric. Hence Q is real orthogonal.

Uniqueness. Let $M = HQ$ be a polar decomposition, then $MM^* = HQQ^*H = H^2$. Because of the uniqueness of the positive-definite square root, we have $H = \sqrt{MM^*}$ and thus $Q = H^{-1}M$.

Smoothness. The map $(H,Q) \mapsto HQ$ is polynomial, hence continuous. Conversely, it is enough to prove that $M \mapsto (H,Q)$ is sequentially continuous, because $\mathbf{GL}_n(\mathbb{C})$ is a metric space. Let $(M_k)_{k\in\mathbb{N}}$ be a convergent sequence in $\mathbf{GL}_n(\mathbb{C})$, with limit M. Let us denote by $M_k = H_kQ_k$ and $M = HQ$ their respective polar decompositions. Because $\mathbf{U}_n$ is compact, the sequence $(Q_k)_{k\in\mathbb{N}}$ admits a cluster point R, that is, a limit of some subsequence $(Q_{k_\ell})_{l\in\mathbb{N}}$, with $k_\ell \to +\infty$. Then $H_{k_\ell} = M_{k_\ell}Q^*_{k_\ell}$ converges to $S := MR^*$. The matrix S is Hermitian positive-semidefinite (because it is the limit of the H_{k_ℓ}s) and invertible (because it is the product of M and R^*). It is thus positive-definite. Hence, SR is a polar decomposition of M. The uniqueness part ensures that $R = Q$ and $S = H$. The sequence $(Q_k)_{k\in\mathbb{N}}$, which is relatively compact and has at most one cluster point (namely Q), converges to Q. Finally, $H_k = M_kQ^*_k$ converges to $MQ^* = H$.

□

Remark

There is as well a polar decomposition $M = QH$ with the same properties. We may speak of left- and right-polar decomposition. We use one or the other depending on the context. We warn the reader that for a given matrix, the H-factors in both decompositions do not coincide. For example, in $M = HQ$, H is the square root of MM^*, although in $M = QH$, it is the square root of M^*M. However, the Q-factors coincide.

10.2 Exponential of a Matrix

The ground field is here $k = \mathbb{C}$. By restriction, we can also treat the case $k = \mathbb{R}$. Because $z \mapsto \exp z$ is holomorphic over $\mathbb{C}$, we may define $\exp(A)$ for every matrix through the functional calculus developed in Section 5.5. However, it is more efficient to give an explicit definition of $\exp A$ by using the exponential series. Both ways are, of course, equivalent.

For A in $\mathbf{M}_n(\mathbb{C})$, the series

$$\sum_{k=0}^{\infty} \frac{1}{k!} A^k$$

converges normally (which means that the series of norms is convergent), because for any matrix norm, we have

$$\sum_{k=0}^{\infty} \left\| \frac{1}{k!} A^k \right\| \leq \sum_{k=0}^{\infty} \frac{1}{k!} \|A\|^k = \exp \|A\|.$$

Because $\mathbf{M}_n(\mathbb{C})$ is complete, the series is convergent, and the estimation above shows that it converges uniformly on every bounded set. Its sum, denoted by $\exp A$, thus defines a continuous map $\exp : \mathbf{M}_n(\mathbb{C}) \to \mathbf{M}_n(\mathbb{C})$, called the *exponential*. When $A \in \mathbf{M}_n(\mathbb{R})$, we have $\exp A \in \mathbf{M}_n(\mathbb{R})$.

Given two matrices A and B in the general position, the binomial formula is not valid: $(A+B)^k$ does not necessarily coincide with

$$\sum_{j=0}^{j=k} \binom{k}{j} A^j B^{k-j}.$$

It thus follows that $\exp(A+B)$ differs in general from $\exp A \cdot \exp B$. A correct statement is the following.

Proposition 10.1 *Let $A, B \in \mathbf{M}_n(\mathbb{C})$ be commuting matrices; that is, $AB = BA$. Then* $\exp(A+B) = (\exp A)(\exp B)$.

Proof. The proof is exactly the same as for the exponential of complex numbers. We observe that because the series defining the exponential of a matrix is normally convergent, we may compute the product $(\exp A)(\exp B)$ by multiplying term by term the series

$$(\exp A)(\exp B) = \sum_{j,k=0}^{\infty} \frac{1}{j!k!} A^j B^k.$$

In other words,

$$(\exp A)(\exp B) = \sum_{\ell=0}^{\infty} \frac{1}{\ell!} C_\ell,$$

where

$$C_\ell := \sum_{j+k=\ell} \frac{\ell!}{j!k!} A^j B^k.$$

From the assumption $AB = BA$, we know that the binomial formula holds. Therefore, $C_\ell = (A+B)^\ell$, which proves the proposition. □

Noting that $\exp 0_n = I_n$ and that A and $-A$ commute, we derive the following consequence.

Corollary 10.1 *For every $A \in \mathbf{M}_n(\mathbb{C})$,* $\exp A$ *is invertible, and its inverse is* $\exp(-A)$.

Given two conjugate matrices $B = P^{-1}AP$, we have $B^k = P^{-1}A^kP$ for each integer k and thus

$$\exp(P^{-1}AP) = P^{-1}(\exp A)P. \tag{10.1}$$

If $D = \mathrm{diag}(d_1, \dots, d_n)$ is diagonal, we have $\exp D = \mathrm{diag}(\exp d_1, \dots, \exp d_n)$. Of course, this formula, or more generally (10.1), can be combined with Jordan reduction in order to compute the exponential of a given matrix. Let us keep in mind, however, that Jordan reduction cannot be carried out explicitly.

Let us introduce a real parameter t and define a function g by $g(t) = \exp tA$. From Proposition 10.1, we see that g satisfies the functional equation

$$g(s+t) = g(s)g(t). \tag{10.2}$$

We have $g(0) = I_n$ and

$$\frac{g(t) - g(0)}{t} - A = \sum_{k=2}^{\infty} \frac{t^{k-1}}{k!} A^k.$$

Using any matrix norm, we deduce that

$$\left\| \frac{g(t) - g(0)}{t} - A \right\| \le \frac{\mathrm{e}^{\|tA\|} - 1 - \|tA\|}{|t|},$$

from which we obtain

$$\lim_{t \to 0} \frac{g(t) - g(0)}{t} = A.$$

We conclude that g has a derivative at $t = 0$, with $g'(0) = A$. Using the functional equation (10.2), we then obtain that g is differentiable everywhere, with

$$g'(t) = \lim_{s \to 0} \frac{g(t)g(s) - g(t)}{s} = g(t)A.$$

We observe that we also have

$$g'(t) = \lim_{s \to 0} \frac{g(s)g(t) - g(t)}{s} = Ag(t).$$

From either of these differential equations we see that g is actually infinitely differentiable. We retain the formula

$$\frac{d}{dt} \exp tA = A \exp tA = (\exp tA)A. \tag{10.3}$$

This differential equation is sometimes the most practical way to compute the exponential of a matrix. This is of particular relevance when A has real entries and has at least one nonreal eigenvalue, if one wishes to avoid the use of complex numbers.

Proposition 10.2 *For every* $A \in \mathbf{M}_n(\mathbb{C})$,

$$\det \exp A = \exp \operatorname{Tr} A. \tag{10.4}$$

Proof. We begin with a reduction of A of the form $A = P^{-1}TP$, where T is upper-triangular. Because T^k is still triangular, with diagonal entries equal to t_{jj}^k, $\exp T$ is triangular too, with diagonal entries equal to $\exp t_{jj}$. Hence

$$\det \exp T = \prod_j \exp t_{jj} = \exp \sum_j t_{jj} = \exp \operatorname{Tr} T.$$

This is the expected formula, inasmuch as $\exp A = P^{-1}(\exp T)P$ and $\operatorname{Tr} A = \operatorname{Tr} T$. □

Because $(M^*)^k = (M^k)^*$, we have $(\exp M)^* = \exp(M^*)$. In particular, the exponential of a skew-Hermitian matrix is unitary, for then

$$(\exp M)^* \exp M = \exp(M^*) \exp M = \exp(-M) \exp M = I_n.$$

Likewise, the exponential of an Hermitian matrix is Hermitian positive-definite, because

$$\exp M = \left(\exp \frac{1}{2}M\right)^2 = \left(\exp \frac{1}{2}M\right)^* \exp \frac{1}{2}M$$

and the fact that $\exp(M/2)$ is nonsingular. This calculation also shows that if M is Hermitian, then

$$\sqrt{\exp M} = \exp \frac{1}{2}M.$$

We have the following more accurate statement.

Proposition 10.3 *The map* $\exp : \mathbf{H}_n \to \mathbf{HPD}_n$ *is a homeomorphism.*

Proof. Injectivity: Let $A, B \in \mathbf{H}_n$ with $\exp A = \exp B =: H$. Then

$$\exp \frac{1}{2}A = \sqrt{H} = \exp \frac{1}{2}B.$$

By induction, we have

$$\exp 2^{-m}A = \exp 2^{-m}B, \quad m \in \mathbb{Z}.$$

Subtracting I_n, multiplying by 2^m, and passing to the limit as $m \to +\infty$, we obtain

$$\left.\frac{d}{dt}\right|_{t=0} \exp tA = \left.\frac{d}{dt}\right|_{t=0} \exp tB;$$

that is, $A = B$.

Surjectivity: Let $H \in \mathbf{HPD}_n$ be given. Then $H = U^* \operatorname{diag}(d_1, \dots, d_n)U$, where U is unitary and $d_j \in (0, +\infty)$. From above, we know that $H = \exp M$ for

$$M := U^* \operatorname{diag}(\log d_1, \dots, \log d_n)U,$$

which is Hermitian.

Continuity: The continuity of exp has already been proved. Let us investigate the continuity of the reciprocal map. Let $(H^\ell)_{\ell\in\mathbb{N}}$ be a sequence in $\mathbf{HPD}_n$ that converges to $H \in \mathbf{HPD}_n$. We denote by $M^\ell, M \in \mathbf{H}_n$, the Hermitian matrices whose exponentials are H^ℓ and H. The continuity of the spectral radius gives

$$\lim_{\ell\to+\infty} \rho(H^\ell) = \rho(H), \quad \lim_{\ell\to+\infty} \rho\left((H^\ell)^{-1}\right) = \rho\left(H^{-1}\right). \tag{10.5}$$

Because $\mathrm{Sp}(M^\ell) = \log \mathrm{Sp}(H^\ell)$, we have

$$\rho(M^\ell) = \log\max\left\{\rho(H^\ell), \rho\left((H^\ell)^{-1}\right)\right\}. \tag{10.6}$$

Keeping in mind that the restriction to $\mathbf{H}_n$ of the induced norm $\|\cdot\|_2$ coincides with that of the spectral radius ρ, we deduce from (10.5) and (10.6) that the sequence $(M^\ell)_{l\in\mathbb{N}}$ is bounded. If N is a cluster point of the sequence, the continuity of the exponential implies $\exp N = H$. But the injectivity shown above implies $N = M$. The sequence $(M^\ell)_{l\in\mathbb{N}}$, bounded with a unique cluster point, is convergent.

□

10.3 Structure of Classical Groups

Proposition 10.4 *Let G be a subgroup of $\mathbf{GL}_n(\mathbb{C})$. We assume that G is stable under the map $M \mapsto M^*$ and that for every $M \in G \cap \mathbf{HPD}_n$, the square root $\sqrt{M}$ is an element of G. Then G is stable under polar decomposition. Furthermore, polar decomposition is a homeomorphism between G and*

$$(G \cap \mathbf{U}_n) \times (G \cap \mathbf{HPD}_n).$$

This proposition applies in particular to subgroups of $\mathbf{GL}_n(\mathbb{R})$ that are stable under transposition and under extraction of square roots in $\mathbf{SPD}_n$. One has then

$$G \overset{\text{homeo}}{\sim} (G \cap \mathbf{O}_n) \times (G \cap \mathbf{SPD}_n).$$

Proof. Let $M \in G$ be given and let HQ be its polar decomposition. Because $MM^* \in G \cdot G = G$, we have $H^2 \in G$, hence $H \in G$ by assumption. Finally, we have $Q = H^{-1}M \in G^{-1} \cdot G = G$. An application of Theorem 10.1 finishes the proof. □

We apply this general result to the classical groups $\mathbf{U}(p,q)$, $\mathbf{O}(p,q)$ (where $n = p+q$) and $\mathbf{Sp}_m$ (where $n = 2m$). These are, respectively, the *unitary* group of the Hermitian form $|z_1|^2 + \cdots + |z_p|^2 - |z_{p+1}|^2 - \cdots - |z_n|^2$, the orthogonal group of the quadratic form $x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_n^2$, and the symplectic group. They are defined by $G = \{M \in \mathbf{M}_n(k) \mid M^*JM = J\}$, with $k = \mathbb{C}$ for $\mathbf{U}(p,q)$, $k = \mathbb{R}$ otherwise.

The matrix J equals

$$\begin{pmatrix} I_p & 0_{p\times q} \\ 0_{q\times p} & -I_q \end{pmatrix},$$

for $\mathbf{U}(p,q)$ and $\mathbf{O}(p,q)$, and

$$\begin{pmatrix} 0_m & I_m \\ -I_m & 0_m \end{pmatrix},$$

for $\mathbf{Sp}_m$. In each case, $J^2 = \pm I_n$.

Proposition 10.5 *Let J be a complex $n\times n$ matrix satisfying $J^2 = \pm I_n$. The subgroup G of $\mathbf{M}_n(\mathbb{C})$ defined by the equation $M^*JM = J$ is invariant under polar decomposition. If $M \in G$, then $|\det M| = 1$.*

Proof. Let $M \in G$. Then $\det J = \det M^* \det M \det J$; that is, $|\det M|^2 = 1$. In particular M is nonsingular. Then we have

$$M^{-*}JM^{-1} = M^{-*}(M^*JM)M^{-1} = J,$$

thus $M^{-1} \in G$. If $M,N \in G$, we also have

$$(MN)^*J(MN) = N^*(M^*JM)N = N^*JN = J$$

and again $MN \in G$. Thus G is a group.

For stability under adjunction, let us write, for $M \in G$,

$$M^*JM(JM^*) = J^2M^* = \pm M^* = M^*J^2.$$

Simplifying by M^*J on the left, there remains $MJM^* = J$; that is, $M^* \in G$.

Because G is a group, $M \in G$ implies $M^k \in G$; that is, $(M^*)^kJ = JM^{-k}$ for every $k \in \mathbb{N}$. By linearity, it follows that $p(M^*)J = Jp(M^{-1})$ holds for every polynomial $p \in \mathbb{R}[X]$. Let us assume in addition that $M \in \mathbf{HPD}_n$. We then have $M = U^* \operatorname{diag}(d_1,\dots,d_n)U$, where U is unitary and the d_js are positive real numbers. Let A be the set formed by the numbers d_j and $1/d_j$. There exists a polynomial p with real entries such that $p(a) = \sqrt{a}$ for every $a \in A$. Then we have $p(M) = \sqrt{M}$ and $p(M^{-1}) = \sqrt{M}^{-1}$. Because $M^* = M$, we also have $p(M)J = Jp(M^{-1})$; that is, $\sqrt{M}J = J\sqrt{M}^{-1}$. Hence $\sqrt{M} \in G$. From Proposition 10.4, G is stable under polar decomposition. □

This leads us to our main result.

Theorem 10.2 *Under the hypotheses of Proposition 10.5, the group G is homeomorphic to $(G\cap \mathbf{U}_n)\times \mathbb{R}^d$, for a suitable integer d.*

Comments

- Of course, if $G = \mathbf{O}(p,q)$ or $\mathbf{Sp}_m$, the subgroup $G\cap\mathbf{U}_n$ can also be written as $G\cap\mathbf{O}_n$.

- We call $G\cap\mathbf{U}_n$ a *maximal compact subgroup* of G, because one can prove that it is not a proper subgroup of a compact subgroup of G. Another deep result, which is beyond the scope of this book, is that every maximal compact subgroup of G is a conjugate of $G\cap\mathbf{U}_n$. In the sequel, when speaking about the maximal compact subgroup of G, we always have in mind $G\cap\mathbf{U}_n$.
- In practice, the maximal compact subgroup is both homeo- and iso-morphic to some simpler classical group.

10.3.1 Calculation Trick

According to Theorem 10.2, one important step in the study of a classical group G is the calculation of $G\cap\mathbf{U}_n$. Both the definitions of G and of $\mathbf{U}_n$ are quadratic, thus nonlinear. However, the calculation can be done linearly, in part: If $M\in G\cap\mathbf{U}_n$, then we have $M^*JM=J$ and $M^*=M^{-1}$, whence the linear equation

$$JM=MJ. \tag{10.7}$$

The elements of $G\cap\mathbf{U}_n$ are precisely the unitary matrices satisfying (10.7). This equation gives easy information about the blocks of M. There remains to describe the unitary matrices that have a rather simple prescribed block form.

Proof. (of Theorem 10.2.)

According to Proposition 10.4, the proof amounts to showing that the factor $G\cap\mathbf{HPD}_n$ is homeomorphic to some $\mathbb{R}^d$. To do this, we define

$$\mathscr{G}:=\{N\in\mathbf{M}_n(k)\,|\,\exp tN\in G,\forall t\in\mathbb{R}\}.$$

Lemma 14. *The set $\mathscr{G}$ defined above satisfies*

$$\mathscr{G}=\{N\in\mathbf{M}_n(k)\,|\,N^*J+JN=0_n\}.$$

Proof. If $N^*J+JN=0_n$, let us set $M(t)=\exp tN$. Then $M(0)=I_n$ and, thanks to (10.3)

$$\frac{d}{dt}M(t)^*JM(t)=M^*(t)(N^*J+JN)M(t)=0_n,$$

so that $M(t)^*JM(t)\equiv J$. We thus have $N\in\mathscr{G}$. Conversely, if $M(t):=\exp tN\in G$ for every t, then the derivative at $t=0$ of $M^*(t)JM(t)=J$ gives $N^*J+JN=0_n$. □

Lemma 15. *The map* $\exp:\mathscr{G}\cap\mathbf{H}_n\to G\cap\mathbf{HPD}_n$ *is a homeomorphism.*

Proof. We must show that $\exp:\mathscr{G}\cap\mathbf{H}_n\to G\cap\mathbf{HPD}_n$ is onto. Let $M\in G\cap\mathbf{HPD}_n$ and let N be the Hermitian matrix such that $\exp N=M$. Let $p\in\mathbb{R}[X]$ be a polynomial with real entries such that for every $\lambda\in\operatorname{Sp}M\cup\operatorname{Sp}M^{-1}$, we have $p(\lambda)=\log\lambda$. Such a polynomial exists, because the numbers λ are real and positive.

Let $N=U^*DU$ be a unitary diagonalization of N. We have

$$M = \exp N = U^*(\exp D)U \qquad \text{and} \qquad M^{-1} = \exp(-N) = U^* \exp(-D)U.$$

Hence, $p(M) = N$ and $p(M^{-1}) = -N$. However, $M \in G$ implies $MJ = JM^{-1}$, and therefore $q(M)J = Jq(M^{-1})$ for every $q \in \mathbb{R}[X]$. With $q = p$, we obtain $NJ = -JN$. $\square$

These two lemmas complete the proof of the theorem, because $\mathscr{G} \cap \mathbf{H}_n$ is an $\mathbb{R}$-vector space. The integer d mentioned in the theorem is its dimension. $\square$

We warn the reader that neither $\mathscr{G}$ nor $\mathbf{H}_n$ is a $\mathbb{C}$-vector space. We present examples in the next section which show that $\mathscr{G} \cap \mathbf{H}_n$ can be naturally $\mathbb{R}$-isomorphic to a $\mathbb{C}$-vector space, which is a source of confusion. One therefore must be cautious when computing d.

The reader eager to learn more about the theory of classical groups is advised to have a look at the book of Mneimné and Testard [30] or the one by Knapp [26].

10.4 The Groups $\mathbf{U}(\mathbf{p},\mathbf{q})$

Let us begin with the study of the maximal compact subgroup of $\mathbf{U}(p,q)$. If $M \in \mathbf{U}(p,q) \cap \mathbf{U}_n$, let us write M blockwise:

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

where $A \in \mathbf{M}_p(\mathbb{C})$, and so on. As discussed above, we have (10.7); here

$$\begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix},$$

which yields $B = 0$ and $C = 0$: M is block diagonal. Then M is unitary if and only if A and D are so. This shows that the maximal compact subgroup of $\mathbf{U}(p,q)$ is isomorphic (not only homeomorphic) to $\mathbf{U}_p \times \mathbf{U}_q$.

Next, $\mathscr{G} \cap \mathbf{H}_n$ is the set of matrices

$$N = \begin{pmatrix} A & B \\ B^* & D \end{pmatrix},$$

where $A \in \mathbf{H}_p$, $D \in \mathbf{H}_q$, which satisfy $NJ + JN = 0_n$; that is, $A = 0_p$, $D = 0_q$. Hence $\mathscr{G} \cap \mathbf{H}_n$ is $\mathbb{R}$-isomorphic to $\mathbf{M}_{p\times q}(\mathbb{C})$. One therefore has $d = 2pq$.

Proposition 10.6 *The unitary group $\mathbf{U}(p,q)$ is homeomorphic to $\mathbf{U}_p \times \mathbf{U}_q \times \mathbb{R}^{2pq}$. In particular, $\mathbf{U}(p,q)$ is connected.*

There remains to show connectedness. It is a straightforward consequence of the following lemma.

Lemma 16. *The unitary group $\mathbf{U}_n$ is connected.*

Inasmuch as $\mathbf{GL}_n(\mathbb{C})$ is homeomorphic to $\mathbf{U}_n \times \mathbf{HPD}_n$ (via polar decomposition), hence to $\mathbf{U}_n \times \mathbf{H}_n$ via the exponential), it is equivalent to the following statement.

Lemma 17. *The linear group* $\mathbf{GL}_n(\mathbb{C})$ *is connected.*

Proof. Let $M \in \mathbf{GL}_n(\mathbb{C})$ be given. Define $A := \mathbb{C} \setminus \{(1-\lambda)^{-1} | \lambda \in \mathrm{Sp}(M)\}$, which is arcwise connected because its complement is finite. The set A contains the origin and the point $z = 1$, because $0 \notin \mathrm{Sp}(M)$. There exists a path γ joining 0 to 1 in A : $\gamma \in \mathscr{C}([0,1];A)$, $\gamma(0) = 0$, and $\gamma(1) = 1$. Let us define $M(t) := \gamma(t)M + (1-\gamma(t))I_n$. By construction, $M(t)$ is invertible for every t, and $M(0) = I_n$, $M(1) = M$. The connected component of I_n is thus all of $\mathbf{GL}_n(\mathbb{C})$. □

10.5 The Orthogonal Groups $\mathbf{O(p,q)}$

The analysis of the maximal compact subgroup and of $\mathscr{G} \cap \mathbf{H}_n$ for the group $\mathbf{O}(p,q)$ is identical to that for $\mathbf{U}(p,q)$. On the one hand, $\mathbf{O}(p,q) \cap \mathbf{O}_n$ is isomorphic to $\mathbf{O}_p \times \mathbf{O}_q$. However, $\mathscr{G} \cap \mathbf{H}_n$ is isomorphic to $\mathbf{M}_{p\times q}(\mathbb{R})$, which is of dimension $d = pq$.

Proposition 10.7 *Let* $n \geq 1$. *The group* $\mathbf{O}(p,q)$ *is homeomorphic to* $\mathbf{O}_p \times \mathbf{O}_q \times \mathbb{R}^{pq}$. *The number of its connected components is two if p or q is zero, four otherwise.*

Proof. We must show that $\mathbf{O}_n$ has two connected components. It is the disjoint union of $\mathbf{SO}_n$ (matrices of determinant $+1$) and of $\mathbf{O}_n^-$ (matrices of determinant -1). Because $\mathbf{O}_n^- = M \cdot \mathbf{SO}_n$ for any matrix $M \in \mathbf{O}_n^-$ (e.g., a hyperplane symmetry), there remains to show that the special orthogonal group $\mathbf{SO}_n$ is connected, in fact arcwise connected. We use the following property:

Lemma 18. *Given* $M \in \mathbf{O}_n$, *there exists* $Q \in \mathbf{O}_n$ *such that the matrix* $Q^{-1}MQ$ *has the form*

$$\begin{pmatrix} (\cdot) & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & (\cdot) \end{pmatrix}, \tag{10.8}$$

where the diagonal blocks are of size 1×1 *or* 2×2 *and are orthogonal. The* 1×1 *blocks are* (± 1), *whereas those of size* 2×2 *are rotation matrices:*

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}. \tag{10.9}$$

Let us apply Lemma 18 to $M \in \mathbf{SO}_n$. The determinant of M, which is the product of the determinants of the diagonal blocks, equals $(-1)^m$, m being the multiplicity of the eigenvalue -1. Because $\det M = 1$, m is even, and we can assemble the diagonal -1s pairwise in order to form matrices of the form (10.9), with $\theta = \pi$. Finally, there exists $Q \in \mathbf{O}_n$ such that

$$M = Q^T \begin{pmatrix} \begin{pmatrix} \cos\theta_1 & \sin\theta_1 \\ -\sin\theta_1 & \cos\theta_1 \end{pmatrix} & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & \begin{pmatrix} \cos\theta_r & \sin\theta_r \\ -\sin\theta_r & \cos\theta_r \end{pmatrix} & \ddots & & 0 \\ \vdots & & \ddots & 1 & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix} Q.$$

Let us now define a matrix $M(t)$ by the same formula, in which we replace the angles θ_j by $t\theta_j$. We thus obtain a path in $\mathbf{SO}_n$, from $M(0) = I_n$ to $M(1) = M$. The connected component of I_n is thus the whole of $\mathbf{SO}_n$. □

We now prove Lemma 18. As an orthogonal matrix, M is normal. From Theorem 5.5, it decomposes into a matrix of the form (10.8), the 1×1 diagonal blocks being the real eigenvalues. These eigenvalues are ± 1, inasmuch as M is orthogonal. The diagonal blocks 2×2 are direct similitude matrices. However, they are isometries, because $Q^{-1}MQ$ is orthogonal. Hence they are rotation matrices.

10.5.1 Notable Subgroups of $\mathbf{O}(\mathbf{p},\mathbf{q})$

We describe here the four connected components of $\mathbf{O}(p,q)$ when $p,q \geq 1$.

Let us write $M \in \mathbf{O}(p,q)$ blockwise

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

where $A \in \mathbf{M}_p(\mathbb{R})$, and so on. When writing $M^T J M = J$, we find in particular $A^T A = C^T C + I_p$, as well as $D^T D = B^T B + I_q$. From the former identity, $A^T A$ is larger than I_p as a symmetric matrix, hence $\det A$ cannot vanish. More precisely, $|\det A| \geq 1$. Likewise, the latter shows that $\det D$ does not vanish. The continuous map $M \mapsto (\det A, \det D)$ thus sends $\mathbf{O}(p,q)$ to $((-\infty,-1]\cup[1,+\infty))^2$. The sign map from $(-\infty,-1]\cup[1,+\infty)$ to $\{-,+\}$ is continuous, therefore we define a continuous function

$$\begin{aligned} \mathbf{O}(p,q) &\xrightarrow{\sigma} \{-,+\}^2 \sim (\mathbb{Z}/2\mathbb{Z})^2, \\ M &\mapsto (\operatorname{sgn}\det A, \operatorname{sgn}\det D). \end{aligned}$$

The diagonal matrices whose diagonal entries are ± 1 belong to $\mathbf{O}(p,q)$. It follows that σ is onto. Because σ is continuous, the preimage G_α of an element α of $\{-,+\}^2$ is the union of some connected components of $\mathbf{O}(p,q)$. Let $n(\alpha)$ be the number of these components. Then $n(\alpha) \geq 1$ (σ being onto), and $\sum_\alpha n(\alpha)$ equals 4, the number of connected components of $\mathbf{O}(p,q)$. There are four terms in this sum, therefore we obtain $n(\alpha) = 1$ for every α. Finally, the connected components of $\mathbf{O}(p,q)$ are the G_αs, where $\alpha \in \{-,+\}^2$.

The left multiplication by an element M of $\mathbf{O}(p,q)$ is continuous, bijective, and its inverse (another multiplication) is continuous. It thus induces a permutation of the set π_0 of connected components of $\mathbf{O}(p,q)$. Because σ induces a bijection between π_0 and $\{-,+\}^2$, there exists a permutation q_M of $\{-,+\}^2$ such that $\sigma(MM') = q_M(\sigma(M'))$. Likewise, the multiplication at the right by M' is a homeomorphism, whence another permutation $p_{M'}$ of $\{-,+\}^2$ such that $\sigma(MM') = p_{M'}(\sigma(M))$. The equality

$$p_{M'}(\sigma(M)) = q_M(\sigma(M'))$$

shows that p_M and q_M actually depend only on $\sigma(M)$. In other words, $\sigma(MM')$ depends only on $\sigma(M)$ and $\sigma(M')$. A direct evaluation in the special case of matrices in $\mathbf{O}(p,q) \cap \mathbf{O}_n(\mathbb{R})$ leads to the following conclusion.

Proposition 10.8 ($p,q \geq 1$) *The connected components of $G = \mathbf{O}(p,q)$ are the sets $G_\alpha := \sigma^{-1}(\alpha)$, defined by $\alpha_1 \det A > 0$ and $\alpha_2 \det D > 0$, when a matrix M is written blockwise as above. The map $\sigma : \mathbf{O}(p,q) \to \{-,+\}^2$ is a surjective group homomorphism; that is, $\sigma(MM') = \sigma(M)\sigma(M')$. In particular:*

1. $G_\alpha^{-1} = G_\alpha$.
2. $G_\alpha \cdot G_{\alpha'} = G_{\alpha\alpha'}$.

Remark

The map σ admits a right inverse, namely

$$\alpha \mapsto M^\alpha := \operatorname{diag}(\alpha_1, 1, \ldots, 1, \alpha_2).$$

The group $\mathbf{O}(p,q)$ is therefore the semidirect product of G_{++} with $(\mathbb{Z}/2\mathbb{Z})^2$.

We deduce immediately from the proposition that $\mathbf{O}(p,q)$ possesses five open and closed normal subgroups, the preimages of the five subgroups of $(\mathbb{Z}/2\mathbb{Z})^2$:

- $\mathbf{O}(p,q)$ itself.
- G_{++}, which we also denote by G_0 (see Exercise 15), the connected component of the unit element I_n.
- $G_{++} \cup G_\alpha$, for the three other choices of an element α.

One of these groups, namely $G_{++} \cup G_{--}$ is equal to the kernel $\mathbf{SO}(p,q)$ of the homomorphism $M \mapsto \det M$. In fact, this kernel is open and closed, thus is the union of connected components of $\mathbf{O}(p,q)$. However, the sign of $\det M$ for $M \in G_\alpha$ is that of $\alpha_1\alpha_2$, which can be seen directly from the case of diagonal matrices M^α.

10.5.2 The Lorentz Group $\mathbf{O(1,3)}$

If $p = 1$ and $q = 3$, the group $\mathbf{O}(1,3)$ is isomorphic to the orthogonal group of the Lorentz quadratic form $dt^2 - dx_1^2 - dx_2^2 - dx_3^2$, which defines the space–time distance in special relativity.[1] Each element M of $\mathbf{O}(1,3)$ corresponds to the transformation

$$\begin{pmatrix} t \\ x \end{pmatrix} \mapsto M \begin{pmatrix} t \\ x \end{pmatrix},$$

which we still denote by M, by abuse of notation. This transformation preserves the light cone of equation $t^2 - x_1^2 - x_2^2 - x_3^2 = 0$. Because it is a homeomorphism of $\mathbb{R}^4$, it permutes the connected components of the complement $\mathscr{C}$ of that cone. There are three such components (see Figure 10.1):

- The convex set $C_+ := \{(t,x) \mid \|x\| < t\}$;
- The convex set $C_- := \{(t,x) \mid t < -\|x\|\}$;
- The "ring" $\mathscr{A} := \{(t,x) \mid |t| < \|x\|\}$.

Clearly, C_+ and C_- are homeomorphic. For example, they are so via the time reversal $t \mapsto -t$. However, they are not homeomorphic to $\mathscr{A}$, because the latter is homeomorphic to $S^2 \times \mathbb{R}^2$ (here, S^2 denotes the unit sphere), which is not contractible, whereas a convex set is always contractible. Because M is a homeomorphism, one deduces that necessarily, $M\mathscr{A} = \mathscr{A}$, and $MC_+ = C_\pm$, $MC_- = C_\mp$.

The transformations that preserve C_+ (we say that they preserve the *time arrow*), and therefore every connected component of $\mathscr{C}$, form the *orthochronous Lorentz group*, denoted $\mathbf{O}^+(1,3)$. Its elements are those that send the vector $\mathbf{e}_0 := (1,0,0,0)^T$ to C_+; that is, those for which the first component of $M\mathbf{e}_0$ is positive. Because this component is A (here a scalar), this group must be $G_{++} \cup G_{+-}$. The transformations that preserve the time arrow and the orientation form the group $G_{++} =: \mathbf{SO}^+(1,3)$.

10.6 The Symplectic Group $\mathbf{Sp_n}$

To begin with, we describe the maximal compact subgroup $\mathbf{Sp}_n \cap \mathbf{O}_{2n}$. If

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathbf{Sp}_n \cap \mathbf{O}_{2n},$$

with blocks of size $n \times n$, then M satisfies (10.7); that is,

$$C = -B, \qquad D = A.$$

Hence

[1] We have selected a system of units in which the speed of light equals one.

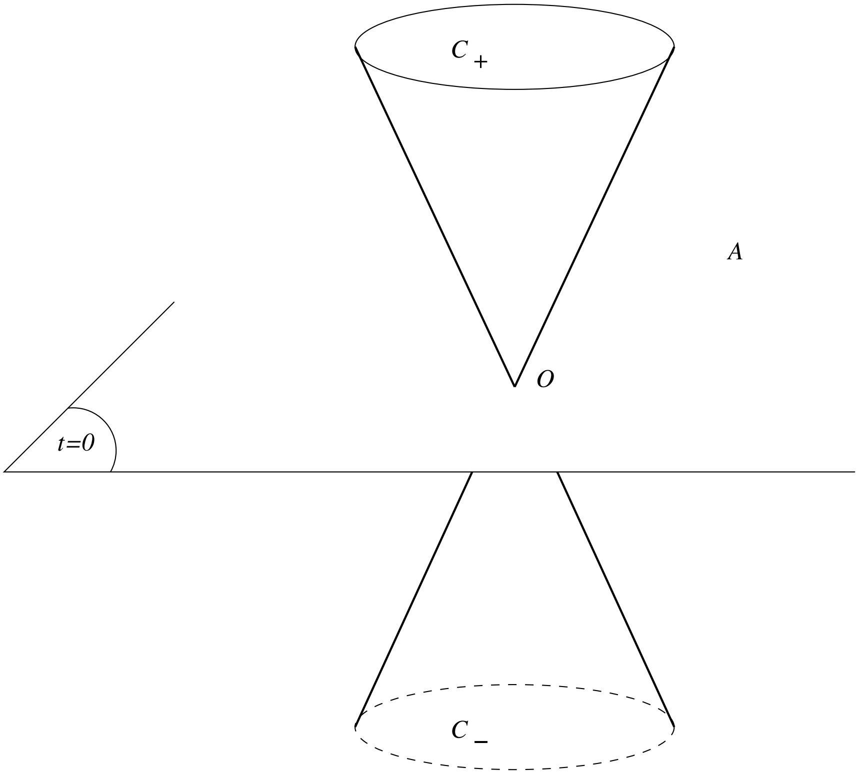

Fig. 10.1 The Lorentz cone. The spatial dimension has been reduced from 3 to 2 for the sake of clarity.

$$M = \begin{pmatrix} A & B \\ -B & A \end{pmatrix}.$$

There remains to write that M is orthogonal, which gives

$$A^T A + B^T B = I_n, \quad A^T B = B^T A.$$

This amounts to saying that $A + \mathrm{i}B$ is unitary. One immediately checks that the map $M \mapsto A + \mathrm{i}B$ is an isomorphism from $\mathbf{Sp}_n \cap \mathbf{O}_{2n}$ onto $\mathbf{U}_n$.

Next, if

$$N = \begin{pmatrix} A & B \\ B^T & D \end{pmatrix}$$

is symmetric and $NJ + JN = 0_{2n}$, we have, in fact,

$$N = \begin{pmatrix} A & B \\ B & -A \end{pmatrix},$$

where A and B are symmetric. Hence $\mathcal{G} \cap \mathbf{Sym}_{2n}$ is isomorphic to $\mathbf{Sym}_n \times \mathbf{Sym}_n$, that is, to $\mathbb{R}^{n(n+1)}$.

Proposition 10.9 *The symplectic group* $\mathbf{Sp}_n$ *is homeomorphic to* $\mathbf{U}_n \times \mathbb{R}^{n(n+1)}$.

Corollary 10.2 *In particular, every symplectic matrix has determinant* $+1$.

Indeed, Proposition 10.9 implies that $\mathbf{Sp}_n$ is connected. Because the determinant is continuous, with values in $\{-1, 1\}$, it is constant, equal to $+1$.

Remark

Corollary 10.2 follows as well, and for every scalar field, from the formula (3.21).

Exercises

1. In the left- and right-polar decompositions of $M \in \mathbf{GL}_n(\mathbb{C})$, the unitary factors equal respectively $(MM^*)^{-1/2}M$ and $M(M^*M)^{-1/2}$. Deduce that they are equal to each other. **Hint:** More generally, $f(MM^*)M = Mf(M^*M)$ for every polynomial.
   Show that M is normal if and only if the Hermitian factor is the same in the left- and right-polar decompositions.
2. Let $M \in \mathbf{M}_n(k)$ be given, with $k = \mathbb{R}$ or $\mathbb{C}$. Show that there exists a polynomial $P \in k(X)$, of degree at most $n-1$, such that $P(M) = \exp M$. However, show that this polynomial cannot be chosen independently of the matrix.
   Compute this polynomial when M is nilpotent.
3. For $t \in \mathbb{R}$, define *Pascal's matrix* $P(t)$ by $p_{ij}(t) = 0$ if $i < j$ (the matrix is lower-triangular) and
$$p_{ij}(t) = t^{i-j}\binom{i-1}{j-1}$$
   otherwise. Let us emphasize that for just this once in this book, P is an *infinite* matrix, meaning that its indices range over the infinite set $\mathbb{N}^*$. Compute $P'(t)$ and deduce that there exists a matrix L such that $P(t) = \exp(tL)$. Calculate L explicitly.
4. We use *Schur's norm* $\|A\| = (\mathrm{Tr} A^*A)^{1/2}$.
   a. If $A \in \mathbf{M}_n(\mathbb{C})$, show that there exists $Q \in \mathbf{U}_n$ such that $\|A - Q\| \leq \|A - U\|$ for every $U \in \mathbf{U}_n$. We define $S := Q^{-1}A$. We therefore have $\|S - I_n\| \leq \|S - U\|$ for every $U \in \mathbf{U}_n$.
   b. Let $H \in \mathbf{H}_n$ be an Hermitian matrix. Show that $\exp(\mathrm{i}tH) \in \mathbf{U}_n$ for every $t \in \mathbb{R}$. Compute the derivative at $t = 0$ of
$$t \mapsto \|S - \exp(\mathrm{i}tH)\|^2$$

and deduce that $S \in \mathbf{H}_n$.

c. Let D be a diagonal matrix, unitarily similar to S. Show that $\|D - I_n\| \leq \|DU - I_n\|$ for every $U \in \mathbf{U}_n$. By selecting a suitable U, deduce that $S \geq 0_n$.

d. If $A \in \mathbf{GL}_n(\mathbb{C})$, show that QS is the polar decomposition of A.

e. Deduce that if $H \in \mathbf{HPD}_n$ and if $U \in \mathbf{U}_n$, $U \neq I_n$, then $\|H - I_n\| < \|H - U\|$.

f. Finally, show that if $H \in \mathbf{H}_n$, $H \geq 0_n$ and $U \in \mathbf{U}_n$, then $\|H - I_n\| \leq \|H - U\|$.

5. Let $A \in \mathbf{M}_n(\mathbb{C})$ and $h \in \mathbb{C}$. Show that $I_n - hA$ is invertible as soon as $|h| < 1/\rho(A)$. One then denotes its inverse by $R(h;A)$ (the *resolvant* of A).

   a. Let $r \in (0, 1/\rho(A))$. Show that there exists a $c_0 > 0$ such that for every $h \in \mathbb{C}$ with $|h| \leq r$, we have

   $$\|R(h;A) - \mathrm{e}^{hA}\| \leq c_0 |h|^2.$$

   b. Verify the formula

   $$C^m - B^m = (C-B)C^{m-1} + \cdots + B^{\ell-1}(C-B)C^{m-\ell} + \cdots + B^{m-1}(C-B),$$

   and deduce the bound

   $$\|R(h;A)^m - \mathrm{e}^{mhA}\| \leq c_0 m |h|^2 \mathrm{e}^{c_2 m|h|},$$

   when $|h| \leq r$ and $m \in \mathbb{N}$.

   c. Show that for every $t \in \mathbb{C}$,

   $$\lim_{m \to +\infty} R(t/m;A)^m = \mathrm{e}^{tA}.$$

   This is the convergence of the implicit Euler difference scheme for the differential equation

   $$\frac{dx}{dt} = Ax. \tag{10.10}$$

6. a. Let $J(a;r)$ be a Jordan block of size r, with $a \in \mathbb{C}^*$. Let $b \in \mathbb{C}$ be such that $a = \mathrm{e}^b$. Show that there exists a nilpotent $N \in \mathbf{M}_r(\mathbb{C})$ such that $J(a;r) = \exp(bI_r + N)$.

   b. Show that $\exp : \mathbf{M}_n(\mathbb{C}) \to \mathbf{GL}_n(\mathbb{C})$ is onto, but that it is not one-to-one. Deduce that $X \mapsto X^2$ is onto $\mathbf{GL}_n(\mathbb{C})$. Verify that it is not onto $\mathbf{M}_n(\mathbb{C})$.

7. a. Show that the matrix

   $$J_2 = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$$

   is not the square of any matrix of $\mathbf{M}_2(\mathbb{R})$.

   b. Show, however, that the matrix $J_4 := \mathrm{diag}(J_2, J_2)$ is the square of a matrix of $\mathbf{M}_4(\mathbb{R})$.

   Show also that the matrix

   $$J_3 = \begin{pmatrix} J_2 & I_2 \\ 0_2 & J_2 \end{pmatrix}$$

is not the square of a matrix of $\mathbf{M}_4(\mathbb{R})$.

c. Show that J_2 is not the exponential of any matrix of $\mathbf{M}_2(\mathbb{R})$. Compare with the previous exercise.

d. Show that J_4 is the exponential of a matrix of $\mathbf{M}_4(\mathbb{R})$, but that J_3 is not.

8. Let $\mathbf{A}_n(\mathbb{C})$ be the set of skew-Hermitian matrices of size n. Show that $\exp : \mathbf{A}_n(\mathbb{C}) \to \mathbf{U}_n$ is onto. **Hint:** If U is unitary, diagonalize it.

9. a. Let $\theta \in \mathbb{R}$ be given. Compute $\exp B$, where

$$B = \begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix}.$$

b. Let $\mathbf{A}_n(\mathbb{R})$ be the set of real skew-symmetric matrices of size n. Show that $\exp : \mathbf{A}_n(\mathbb{R}) \to \mathbf{SO}_n$ is onto. **Hint:** Use the reduction of direct orthogonal matrices.

10. (J. Duncan.) We denote $\langle x, y\rangle$ the usual sesquilinear product in $\mathbb{C}^n$. Let $M \in \mathbf{GL}_n(\mathbb{C})$ be given. Thanks to Exercise 1, the left- and right-polar decompositions of M write $M = UH = KU$, with $H = \sqrt{M^*M}$ and $K = \sqrt{MM^*}$.

a. Prove that $U\sqrt{H} = \sqrt{K}U$.

b. Check that

$$\langle Mx, y\rangle = \langle \sqrt{H}x, U^*\sqrt{K}y\rangle, \qquad \forall x, y \in \mathbb{C}^n.$$

Deduce that

$$|\langle Mx, y\rangle|^2 \le \langle Hx, x\rangle\, \langle Ky, y\rangle.$$

c. More generally, let a rectangular matrix $A \in \mathbf{M}_{n\times m}(\mathbb{C})$ be given. Prove the generalized Cauchy–Schwarz inequality

$$|\langle Ax, y\rangle|^2 \le \langle \sqrt{A^*A}x, x\rangle\, \langle \sqrt{AA^*}y, y\rangle, \qquad \forall x, y \in \mathbb{C}^n.$$

Hint: Use the decompositions

$$\mathbb{C}^m = \ker A \oplus^\perp R(A^*), \qquad \mathbb{C}^n = \ker A^* \oplus^\perp R(A),$$

then apply the case above to the restriction of A from $R(A^*)$ to $R(A)$.

11. Let $\phi : \mathbf{M}_n(\mathbb{R}) \to \mathbb{R}$ be a nonnull map satisfying $\phi(AB) = \phi(A)\phi(B)$ for every $A, B \in \mathbf{M}_n(\mathbb{R})$. If $\alpha \in \mathbb{R}$, we set $\delta(\alpha) = |\phi(\alpha I_n)|^{1/n}$. We have seen, in Exercise 5 of Chapter 5, that $|\phi(M)| = \delta(\det M)$ for every $M \in \mathbf{M}_n(\mathbb{R})$.

a. Show that on the range of $M \mapsto M^2$ and on that of $M \mapsto \exp M$, $\phi \equiv \delta \circ \det$.

b. Deduce that $\phi \equiv \delta \circ \det$ on $\mathbf{SO}_n$ (use Exercise 9) and on $\mathbf{SPD}_n$.

c. Show that either $\phi \equiv \delta \circ \det$ or $\phi \equiv (\operatorname{sgn}(\det))\delta \circ \det$.

12. Let A be a Banach algebra ($K = \mathbb{R}$ or $\mathbb{C}$) with a unit denoted by e. If $x \in A$, define $x^0 := e$.

a. Given $x \in A$, show that the series

$$\sum_{m\in\mathbb{N}} \frac{1}{m!}x^m$$

converges normally, hence converges in A. Its sum is denoted by $\exp x$.

b. If $x,y\in A$, $[x,y]=xy-yx$ is called the "commutator" of x and y. Verify that $[x,y]=0$ implies

$$\exp(x+y)=(\exp x)(\exp y), \quad [x,\exp y]=0.$$

c. Show that the map $t\mapsto \exp tx$ is differentiable on $\mathbb{R}$, with

$$\frac{d}{dt}\exp tx = x\exp tx = (\exp tx)x.$$

d. Let $x,y\in A$ be given. Assume all along this part that $[x,y]$ commutes with x and y.
   i. Show that $(\exp -tx)xy(\exp tx)=xy+t[y,x]x$.
   ii. Deduce that $[\exp -tx,y]=t[y,x]\exp -tx$.
   iii. Compute the derivative of $t\mapsto(\exp -ty)(\exp -tx)\exp t(x+y)$. Finally, prove the Campbell–Hausdorff formula

$$\exp(x+y)=(\exp x)(\exp y)\left(\exp\frac{1}{2}[y,x]\right).$$

e. In $A=\mathbf{M}_3(\mathbb{R})$, construct an example that satisfies the above hypothesis ($[x,y]$ commutes with x and y), where $[x,y]$ is nonzero.

13. Show that the map

$$H\mapsto f(H):=(\mathrm{i}I_n+H)(\mathrm{i}I_n-H)^{-1}$$

induces a homeomorphism from $\mathbf{H}_n$ onto the set of matrices of $\mathbf{U}_n$ whose spectrum does not contain -1. Find an equivalent of $f(tH)-\exp(-2itH)$ as $t\to 0$.

14. Let G be a group satisfying the hypotheses of Proposition 10.5.

a. Show that $\mathscr{G}$ is a *Lie algebra*, meaning that it is stable under the bilinear map $(A,B)\mapsto[A,B]:=AB-BA$.

b. Show that for $t\to 0+$,

$$\exp(tA)\exp(tB)\exp(-tA)\exp(-tB)=I_n+t^2[A,B]+O(t^3).$$

Deduce another proof of the stability of $\mathscr{G}$ by $[\cdot,\cdot]$.

c. Show that the map $M\mapsto[A,M]$ is a derivation, meaning that the Jacobi identity

$$[A,[B,C]]=[[A,B],C]+[B,[A,C]]$$

holds true.

15. A *topological group* is a group G endowed with a topology for which the maps $(g,h) \mapsto gh$ and $g \mapsto g^{-1}$ are continuous. Show that in a topological group, the connected component of the unit element is a normal subgroup. Show also that the open subgroups are closed. Illustrate this result with $G = \mathbf{O}(p,q)$. Give an example of a closed subgroup that is not open.
16. One identifies $\mathbb{R}^{2n}$ with $\mathbb{C}^n$ by the map

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto x + \mathrm{i}y.$$

Therefore, every matrix $M \in \mathbf{M}_{2n}(\mathbb{R})$ defines an $\mathbb{R}$-linear map $\tilde{M}$ from $\mathbb{C}^n$ into itself.

   a. Let

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathbf{M}_{2n}(\mathbb{R})$$

   be given. Under what condition on the blocks A,B,C,D is the map $\tilde{M}$ $\mathbb{C}$-linear?

   b. Show that $M \mapsto \tilde{M}$ is an isomorphism from $\mathbf{Sp}_n \cap \mathbf{O}_{2n}$ onto $\mathbf{U}_n$.

17. Let k be a field and

$$P = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

be an orthogonal matrix, with A and D square.
Prove that

$$\det D = \pm \det A.$$

Hint: multiply P by

$$\begin{pmatrix} A^T & C^T \\ 0 & I \end{pmatrix}.$$

Extend this result to elements P of a group $\mathbf{O}(p,q)$.

18. Let $A \in \mathbf{M}_n(\mathbb{C})$ be given, and $U(t) := \exp(tA)$.

   a. Show that $\|U(t)\| \le \exp(t\|A\|)$ for $t \ge 0$ and any matrix norm. Deduce that the integral

$$\int_0^{+\infty} \mathrm{e}^{-2\gamma t} U(t)^* U(t)\, dt$$

   converges for every $\gamma > \|A\|$.

   b. Denote H_γ the value of this integral, when it is defined. Computing the derivative at $h = 0$ of $h \mapsto U(h)^* H_\gamma U(h)$, by two different methods, deduce that H_γ is a solution of

$$A^* X + XA = 2\gamma X - I_n, \quad X \in \mathbf{HPD}_n. \tag{10.11}$$

c. Let γ be larger than the supremum of the real parts of eigenvalues of A. Show that Equation (10.11) admits a unique solution in $\mathbf{HPD}_n$, and that the above integral converges.
d. In particular, if the spectrum of M has positive real part, and if $K \in \mathbf{HPD}_n$ is given, then the Lyapunov equation

$$M^*H + HM = K, \qquad H \in \mathbf{HPD}_n$$

admits a unique solution.
Let $x(t)$ be a solution of the differential equation $\dot{x} + Mx = 0$, show that $t \mapsto x^*Hx$ decays, and strictly if $x \neq 0$.

19. This exercise shows that a matrix $M \in \mathbf{GL}_n(\mathbb{R})$ is the exponential of a real matrix if and only if it is the square of another real matrix.

a. Show that, in $\mathbf{M}_n(\mathbb{R})$, every exponential is a square.
b. Given a matrix $A \in \mathbf{M}_n(\mathbb{C})$, we denote $\mathscr{A}$ the $\mathbb{C}$-algebra spanned by A, that is, the set of matrices $P(A)$ as P runs over $\mathbb{C}[X]$.
    i. Check that $\mathscr{A}$ is commutative, and that the exponential map is a homomorphism from $(\mathscr{A},+)$ to $(\mathscr{A}^*,\times)$, where $\mathscr{A}^*$ denotes the subset of invertible matrices (a multiplicative group.)
    ii. Show that $\mathscr{A}^*$ is an open and connected subset of $\mathscr{A}$.
    iii. Let E denote $\exp(\mathscr{A})$, so that E is a subgroup of $\mathscr{A}^*$. Show that E is a neighbourhood of the identity. **Hint:** Use the implicit function theorem.
    iv. Deduce that E is closed in $\mathscr{A}^*$. **Hint:** The complement F of E in $\mathscr{A}$ satisfies $F = F \cdot E$ and thus is open. Conclude that $E = \mathscr{A}^*$.
    v. Finally, show that every matrix $B \in \mathbf{GL}_n(\mathbb{C})$ reads $B = \exp(P(B))$ for some polynomial $P \in \mathbb{C}[X]$.
c. Let $B \in \mathbf{GL}_n(\mathbb{R})$ and $P \in \mathbb{C}[X]$ be as above. Show that

$$B^2 = \exp(P(B) + \bar{P}(B)).$$

Conclusion?

20. (See also [41].)
Let M belong to $\mathbf{Sp}_n(\mathbb{R})$. We recall that $M^TJM = J$, $J^2 = -I_{2n}$, and $J^T = -J$.

a. Show that the characteristic polynomial is reciprocal:

$$P_M(X) = X^{2n} P_M\left(\frac{1}{X}\right).$$

Deduce a classification of the eigenvalues of M.
b. Define the quadratic form

$$q(x) := 2x^TJMx.$$

Verify that M is a q-isometry.

c. Let $(e^{-i\theta}, e^{i\theta})$ be a pair of *simple* eigenvalues of M on the unit circle. Let Π be the corresponding invariant subspace:

$$\Pi := \ker(M^2 - 2(\cos\theta)M + I_{2n}).$$

i. Show that $J\Pi^\perp$ is invariant under M.
ii. Using the formula (5.3), show that $e^{\pm i\theta}$ are not eigenvalues of $M|_{J\Pi^\perp}$.
iii. Deduce that $\mathbb{R}^{2n} = \Pi \oplus J\Pi^\perp$.

d. (Continued.)
i. Show that q does not vanish on $\Pi \setminus \{0\}$. Hence q defines a Euclidean structure on Π.
ii. Check that $M|_\Pi$ is direct (its determinant is positive).
iii. Show that $M|_\Pi$ is a rotation with respect to the Euclidean structure defined by q, whose angle is either θ or $-\theta$.

e. More generally, assume that a plane Π is invariant under a symplectic matrix M, with corresponding eigenvalues $e^{\pm i\theta}$, and that Π is not Lagrangian: $(x,y) \mapsto y^T Jx$ is not identically zero on Π. Show that $M|_\Pi$ acts as rotation of angle $\pm\theta$. In particular, if $M = J$, show that $\theta = +\pi/2$.

f. Let H be an invariant subspace of M, on which the form q is either positive or negative-definite. Prove that the spectrum of $M|_H$ lies in the unit circle and that $M|_H$ is semisimple (the Jordan form is diagonal).

g. Equivalently, let λ be an eigenvalue of M (say a simple one) with $\lambda \notin \mathbb{R}$ and $|\lambda| \neq 1$. Let H be the invariant subspace associated with the eigenvalues $(\lambda, \bar{\lambda}, 1/\lambda, 1/\bar{\lambda})$. Show that the restriction of the form q to H is neither positive nor negative-definite. Show that the invariant subspace K associated with the eigenvalues λ and $\bar{\lambda}$ is q-isotropic. Thus, if $q|_H$ is nondegenerate, its signature is $(2,2)$.

21. This is a sequel of Exercise 23, Chapter 7. Let Σ denote the unit sphere of $\mathbf{M}_2(\mathbb{R})$ for the induced norm $\|\cdot\|_2$. Recall that Σ is the union of the segments $[r,s]$ where $r \in \mathscr{R} := \mathbf{SO}_2(\mathbb{R})$ and $s \in \mathscr{S}$, the set of orthogonal symmetries. Both $\mathscr{R}$ and $\mathscr{S}$ are circles. Finally, two distinct segments may intersect only at an extremity.

a. Show that there is a well-defined map $\rho : \Sigma \setminus \mathscr{S} \to \mathscr{R}$, such that M belongs to some segment $[\rho(M), s)$ with $s \in \mathscr{S}$. For which M is the other extremity s unique?

b. Show that the map ρ above is continuous, and that ρ coincides with the identity over $\mathscr{R}$. We say that ρ is a *retraction* from $\Sigma \setminus \mathscr{S}$ onto $\mathscr{R}$.

c. Let $f : D \to \Sigma$ be a continuous function, where D is the unit disk of the complex plane, such that $f(\exp(i\theta))$ is the rotation of angle θ. Show that $f(D)$ contains an element of $\mathscr{S}$.
Hint: Otherwise, there would be a retraction of D onto the unit circle, which is impossible (an equivalent statement to Brouwer fixed point theorem).

Meaning. Likewise, one finds that if a disk D' is immersed in Σ, with boundary $\mathscr{S}$, then it contains an element of $\mathscr{R}$. We say that the circles $\mathscr{R}$ and $\mathscr{S}$ of Σ are *linked.*
This result tells us that $\mathscr{R}$ and $\mathscr{S}$ are linked within Σ.

22. In $\mathbb{R}^{1+m}$ we denote the generic point by $(t,x)^T$, with $t \in \mathbb{R}$ and $x \in \mathbb{R}^m$. Let $\mathscr{C}^+$ be the cone defined by $t > \|x\|$. Recall that those matrices of $\mathbf{O}(1,m)$ that preserve $\mathscr{C}^+$ form the subgroup $\mathbf{O}^+(1,m)$. The Hermitian form $(t,x) \mapsto \|x\|^2 - |t|^2$ is denoted by q.
Let M belong to $\mathbf{O}^+(1,m)$.

    a. Given a point x in the unit closed ball B of $\mathbb{R}^m$, let $(t,y)^T$ be the image of $(1,x)^T$ under M. Define $f(x) := y/t$. Prove that f is a continous map from B into itself. Deduce that it has a fixed point. Deduce that M has at least one real positive eigenvalue, associated with an eigenvector in the closure of $\mathscr{C}^+$. **Note:** If m is odd, one can prove that this eigenvector can be taken in the light cone $t = \|x\|$.
    b. If $Mv = \lambda v$ with $v \in \mathbb{C}^{1+m}$ and $q(v)) \neq 0$, show that $|\lambda| = 1$.
    c. Let $v = (t,x)$ and $w = (s,y)$ be *light* vectors (i.e., $q(v) = q(w) = 0$), linearly independent. Show that $v^*Jw \neq 0$.
    d. Assume that M admits an eigenvalue λ of modulus different from 1, v being an eigenvector. Show that $1/\lambda$ is also an eigenvector. Denote by w a corresponding eigenvector. Let $< v,w >^\circ$ be the orthogonal of v and w with respect to q. Using the previous question, show that the restriction q_1 of q to $< v,w >^\circ$ is positive-definite. Show that $< v,w >^\circ$ is invariant under M and deduce that the remaining eigenvalues have unit modulus. In particular, λ is real.
    e. Show that, for every $M \in \mathbf{O}^+(1,m)$, $\rho(M)$ is an eigenvalue of M.

23. We endow $\mathbf{M}_n(\mathbb{C})$ with the induced norm $\|\cdot\|_2$. Let G be a subgroup of $\mathbf{GL}_n(\mathbb{C})$ that is contained in the open ball $B(I_n;r)$ for some $r < 2$.

    a. Show that for every $M \in G$, there exists an integer $p \geq 1$ such that $M^p = I_n$.
    b. Let $A,B \in G$ be such that $\mathrm{Tr}(AM) = \mathrm{Tr}(BM)$ for every $M \in G$. Prove that $A = B$.
    c. Deduce that G is a finite group.
    d. Conversely, let R be a rotation in the plane ($n = 2$) of angle $\theta \notin \pi\mathbb{Q}$. Prove that the subgroup spanned by R is infinite and is contained in $B(I_2;2)$.

24. Let $m \in \mathbb{N}^*$ be given. We denote $P_m : A \mapsto A^m$ the mth power in $\mathbf{M}_n(\mathbb{C})$. Show that the differential of P_m at A is given by

$$\mathrm{d}P_m(A) \cdot B = \sum_{j=0}^{m-1} A^j B A^{m-1-j}.$$

Deduce the formula

$$\mathrm{D}\exp(A)\cdot B = \int_0^1 \mathrm{e}^{(1-t)A} B \mathrm{e}^{tA}\,dt.$$

25. Let $A \in \mathbf{M}_n(\mathbb{R})$ be a matrix satisfying $a_{ij} \geq 0$ for every pair (i,j) of distinct indices.
   a. Using the Exercise 3 of Chapter 8, show that
   $$R(h;A) := (I_n - hA)^{-1} \geq 0,$$
   for $h > 0$ small enough.
   b. Deduce that $\exp(tA) \geq 0$ for every $t > 0$. **Hint:** Use Exercise 5.
   c. Deduce that if $x(0) \geq 0$, then the solution of (10.10) is nonnegative for every nonnegative t.
   d. Deduce also that
   $$\sigma := \sup\{\Re\lambda \mid \lambda \in \mathrm{Sp}\, A\}$$
   is an eigenvalue of A.

26. We use the scalar product over $\mathbf{M}_n(\mathbb{C})$, given by $\langle M, N\rangle = \mathrm{Tr}(M^*N)$. We recall that the corresponding norm is the Schur–Frobenius norm $\|\cdot\|_F$. If $T \in \mathbf{GL}_n(\mathbb{C})$, we denote $T = U|T|$ the polar decomposition, with $|T| := \sqrt{T^*T}$ and $U \in \mathbf{U}_n$. The Aluthge transform $\Delta(T)$ is defined by
   $$\Delta(T) := |T|^{1/2} U |T|^{1/2}.$$
   a. Check that $\Delta(T)$ is similar to T.
   b. If T is normal, show that $\Delta(T) = T$.
   c. Show that $\|\Delta(T)\|_F \leq \|T\|_F$, with equality if and only if T is normal.
   d. We define Δ^n by induction, with $\Delta^n(T) := \Delta(\Delta^{n-1}(T))$.
      i. Given $T \in \mathbf{GL}_n(\mathbb{C})$, show that the sequence $\left(\Delta^k(T)\right)_{k\in\mathbb{N}}$ is bounded.
      ii. Show that its limit points are normal matrices with the same characteristic polynomial as T (Jung, Ko and Pearcy, or Ando).
      iii. Deduce that when T has only one eigenvalue μ, then the sequence converges towards μI_n.
   e. If T is not diagonalizable, show that these limit points are not similar to T.