Linear Algebra II for Physics

Alberto S. Cattaneo

Institute of Mathematics, University of Zurich

Last update: June 3, 2021

Contents

CHAPTER 1

Recap of Linear Algebra I

This chapter is a summary of the basic concepts introduced in Linear Algebra I which will be used in this course.

1.1. Groups, rings, and fields

DEFINITION 1.1 (Groups). A group is a set G with a distinguished element e, called the neutral element, and an operation, called "multiplication,"

$$\begin{array}{ccc} G \times G & \to & G \\ (a,b) & \mapsto & ab \end{array},$$

which is associative—i.e., $a(bc) = (ab)c$ for all $a, b, c \in G$—and which satisfies $ae = ea = a$. Moreover, for all $a \in G$, there is an element denoted by a^{-1}, and called the inverse of a, satisfying $aa^{-1} = a^{-1}a = e$ for all $a \in G$.

One can show that the inverse is unique, that $(a^{-1})^{-1} = a$ and that $(ab)^{-1} = b^{-1}a^{-1}$ for all $a, b \in G$.

DEFINITION 1.2 (Abelian groups). A group G is called abelian if, in addition, $ab = ba$ for all $a, b \in G$. If G is abelian, one often uses the additive notation in which the neutral element is denoted by 0, the multiplication is called "addition" and denoted by

$$(a,b) \mapsto a + b,$$

and the inverse of an element a is denoted by $-a$ (and called the "additive inverse" or the "opposite" of a).

EXAMPLES 1.3. Here are some examples of groups:

(1) $\mathbb{Z}$ with the usual 0 and the usual addition is an abelian group (in additive notation).
(2) $\mathbb{Z}_{>1}$ with $e = 1$ and the usual multiplication is an abelian group (not in additive notation).

(3) Invertible $n \times n$ matrices form the group GL_n with e the identity matrix and the usual multiplication of matrices; this group is nonabelian for $n > 1$.[1]
(4) The set $\mathrm{Aut}(S)$ of bijective maps of a set S to itself form a group with multiplication given by the composition and neutral element given by the identity map. If $S = \{1, \ldots, n\}$ this group is called the symmetric group on n elements and is denoted by S_n (or $\mathrm{Sym}(n)$); its elements are called permutations.

DEFINITION 1.4 (Rings). A ring is an abelian group $(R, 0, +)$ together with a second associative operation, called "multiplication,"

$$\begin{array}{ccc} R \times R & \to & R \\ (a,b) & \mapsto & ab \end{array},$$

which is also distributive; i.e.,

$$a(b+c) = ab + ac \quad \text{and} \quad (a+b)c = ac + bc$$

for all $a, b, c \in R$. A ring R is called:

(1) a ring with one if it possesses a special element, denoted by 1, such that $a1 = 1a = a$ for all $a \in R$;
(2) a commutative ring if $ab = ba$ for all $a, b \in R$.

EXAMPLES 1.5. Here are some examples of rings:

(1) $\mathbb{Z}$—with the usual addition, multiplication, zero, and one—is a commutative ring with one.
(2) The set $2\mathbb{Z}$ of even numbers—with the usual addition, multiplication, and zero—is a commutative ring without one.
(3) The set $\mathrm{Mat}_{n\times n}$ of $n \times n$ matrices, with the usual addition and multiplication, with 0 the matrix whose all entries are 0, and with 1 the identity matrix, is a ring with one. It is noncommutative for $n > 1$.
(4) Polynomials form a commutative ring with one.
(5) If I is an open interval, the set $C^0(I)$ of continuous functions on I, the set $C^k(I)$ of k times continuously differentiable functions on I, and the set $C^\infty(I)$ of functions on I that are continuously differentiable any number of times are commutative rings with one. Recall that the operations are defined as

$$(f+g)(x) := f(x) + g(x), \quad (fg)(x) := f(x)\, g(x), \qquad x \in I.$$

[1] We assume, without explicitly recapping, the knowledge of matrices, including the notion of sum, product, and transposition. We will recap trace and determinant in Section 1.6.

The zero element is the function $0(x) = 0$ for all x and the one element is the function $1(x) = 1$ for all x.

DEFINITION 1.6 (Fields). A field $\mathbb{K}$ is a commutative ring with one in which every element different from zero is invertible. This is equivalent to saying that $\mathbb{K}^\times := \mathbb{K} \setminus \{0\}$ is a commutative group (not in additive notation).

The only fields we are going to consider in this course are the field $\mathbb{R}$ of real numbers and the field $\mathbb{C}$ of complex numbers.

Many of the results we present actually hold for any field, and most of the results hold for any field of characteristic zero, like $\mathbb{R}$ or $\mathbb{C}$, i.e., a field $\mathbb{K}$ such that there is no nonzero integer n satisfying $na = 0$ for a nonzero element a of $\mathbb{K}$.[2]

DEFINITION 1.7 (Subobjects). A subset of a group/ring/field which retains all the structures is called a subgroup/subring/subfield.

1.2. Vector spaces

A vector space over a field $\mathbb{K}$, whose elements are called scalars, is an abelian group $(V, +, 0)$—in additive notation—whose elements are called vectors, together with an operation

$$\begin{array}{ccc} \mathbb{K} \times V & \to & V \\ (\lambda, v) & \mapsto & \lambda v \end{array},$$

called multiplication by a scalar or scalar multiplication,[3] satisfying

$$\lambda(\mu v) = (\lambda\mu)v, \quad (\lambda + \mu)v = \lambda v + \mu v, \quad \lambda(v + w) = \lambda v + \lambda w,$$

for all $\lambda, \mu \in \mathbb{K}$ and all $v, w \in V$.

EXAMPLE 1.8 (Column vectors). The set $\mathbb{K}^n$ of n-tuples of scalars, conventionally arranged in a column and called column vectors, is a vector space with

$$\begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} + \begin{pmatrix} w^1 \\ \vdots \\ w^n \end{pmatrix} := \begin{pmatrix} v^1 + w^1 \\ \vdots \\ v^n + w^n \end{pmatrix}, \quad 0 := \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

and

$$\lambda \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} := \begin{pmatrix} \lambda v^1 \\ \vdots \\ \lambda v^n \end{pmatrix}.$$

[2]By na one means the sum $a + \cdots + a$ with n summands.

[3]not to be confused with the scalar product a.k.a. the dot product or the inner product

The scalars v^i are called the components of the column vector, which is usually denoted by the corresponding boldface latter $\boldsymbol{v}$.[4]

EXAMPLE 1.9 (Row vectors). The set $(\mathbb{K}^n)^*$ of n-tuples of scalars, conventionally arranged in a row and called row vectors, is a vector space with

$$\big(v_1, \dots, v_n\big) + \big(w_1, \dots, w_n\big) := \big(v_1 + w_1, \dots, v_n + w_n\big)\,, \quad 0 := \big(0, \dots, 0\big)\,,$$

and

$$\lambda\,\big(v_1, \dots, v_n\big) = \big(\lambda v_1, \dots, \lambda v_n\big)\,.$$

The scalars v^i are called the components of the row vector, which is also usually denoted by the corresponding boldface latter $\boldsymbol{v}$.[4]

REMARK 1.10 (Components and indices). The scalars forming a row or column vectors are called its components. We will consistently denote the components of a column vectors with an upper index and the components of a row vectors with a lower index. This is a nowadays standard convention (especially in physics) that comes in handy with Einstein's convention for sums (which we will introduce in Definition 1.33).

EXAMPLE 1.11 (The trivial vector space). The vector space $V = \{0\}$ consisting only of the neutral element 0 is called the trivial vector space. It is denoted by 0, but also by $\mathbb{K}^0$ (and, if you wish, $(\mathbb{K}^0)^*$).

REMARK 1.12 (The zero notation). Observe that the symbol 0 is used for all of the following:

(1) The neutral element of an abelian group in additive notation.
(2) The zero element of a ring or a field.
(3) The zero element of a vector space.
(4) The vector space consiting only of the zero element ($0 = \{0\}$).
(5) A constant map having value 0 (e.g., the continuous real function $x \mapsto 0$ for all $x \in \mathbb{R}$, or the map $V \to W$, $v \mapsto 0$, where V and W are vector spaces).
(6) A matrix whose entries are all equal to 0 (even though we will prefer the notation $\mathbf{0}$).

EXAMPLE 1.13 (Polynomials). The ring $\mathbb{K}[x]$ of polynomials in an undetermined x with coefficients in $\mathbb{K}$ (i.e., expressions of the form $p = a_0 + a_1 x + \cdots + a_d x^d$, for some d, and $a_i \in \mathbb{K}$ for all i) is also a vector space over $\mathbb{K}$ with scalar multiplication $\lambda p := \lambda a_0 + \lambda a_1 x + \cdots + \lambda a_d x^d$ and the usual addition the addition of polynomials (i.e., addition of the coefficients).

[4] Other common notations are $\vec{v}$ and $\underline{v}$.

EXAMPLE 1.14 (Functions). The rings of functions $C^k(I)$, $k \in \mathbb{N} \cup \{\infty\}$, of Example 1.5.(5) are also vector spaces over $\mathbb{R}$ with scalar multiplication $(\lambda f)(x) := \lambda f(x)$.

DEFINITION 1.15 (Subspaces). A subset W of a vector space V that retains all the structures is called a (vector) subspace. Equivalently, $W \subseteq V$ is a subspace iff for every $w, \widetilde{w} \in W$ and for every $\lambda \in \mathbb{K}$ we have $w + \widetilde{w} \in W$ and $\lambda w \in W$.

DEFINITION 1.16 ((Direct) sums of subspaces). If W_1 and W_2 are subspaces of V, we denote by $W_1 + W_2$ the subset of elements of V consisting of sums of elements of W_1 and W_2; i.e.,

$$W_1 + W_2 = \{w_1 + w_2,\ w_1 \in W_1,\ w_2 \in W_2\}.$$

It is also a subspace of V. If $W_1 \cap W_2 = \{0\}$, the sum is called the direct sum and is denoted by $W_1 \oplus W_2$.

REMARK 1.17. A vector $v \in W_1 \oplus W_2$ uniquely decomposes as $v = w_1 + w_2$ with $w_1 \in W_1$ and $w_2 \in W_2$. The vectors w_1 and w_2 are called the components of v in the direct sum.

PROOF. If $v = \widetilde{w}_1 + \widetilde{w}_2$ were another decomposition, by taking the difference we would get $w_1 - \widetilde{w}_1 = \widetilde{w}_2 - w_2$. Since the left hand side is in W_1, the right hand side is in W_2, and $W_1 \cap W_2 = \{0\}$, we have $w_1 - \widetilde{w}_1 = 0 = \widetilde{w}_2 - w_2$. □

REMARK 1.18. Vice versa, if every vector in $W_1 + W_2$ has a unique decomposition $v = w_1 + w_2$ with $w_i \in W_i$, $i = 1, 2$, then this is a direct sum (i.e., $W_1 \cap W_2 = 0$).

PROOF. Suppose $v \in W_1 \cap W_2$. From $0 = v - v$, we see that the first summand, v, is the component of 0 in W_1 and the second summand, $-v$, is its component in W_2. Since we can decompose the zero vector also as $0 = 0 + 0$ (and 0 belongs to both W_1 and W_2), by the assumed uniqueness of the decomposition, we then have $v = 0$. □

DEFINITION 1.19. If W is a subspace of V, a susbpace W' such that $V = W \oplus W'$ is called a complement.

Every subspace admits a complement, see Lemma 1.27 and Proposition 1.55. This is elementary in the case of finite-dimensional vector spaces and requires the axiom of choice for infinite-dimensional ones (see Digression 1.56).

DEFINITION 1.20. We may generalize Definition 1.16 to a sum of several subspaces $W_1, \dots, W_k$:

$$W_1 + \cdots + W_k := \{v_1 + \cdots + v_k,\ v_i \in W_i,\ i = 1, \dots, k\}.$$

We generalize the notion of direct sum via the property in Remark 1.18:

DEFINITION 1.21. A sum $W_1 + \cdots + W_k$ is called a **direct sum**, and it is denoted by

$$W_1 \oplus \cdots \oplus W_k \quad \text{or} \quad \bigoplus_{i=1}^{k} W_i,$$

if every vector v in it has a unique decomposition $v = w_1 + \cdots + w_k$ with $w_i \in W_i$, $i = 1, \ldots, k$.

REMARK 1.22. An easier criterion is the following. A sum $W_1 + \cdots + W_k$ is a direct sum iff the zero vector has a unique decomposition.

PROOF. If the sum is direct, the zero vector has a unique decomposition by definition, like every other vector.

Vice versa, suppose that the zero vector has a unique decomposition and that a vector v can be written both as $w_1 + \cdots + w_k$ and as $w'_1 + \cdots + w'_k$ with $w_i, w'_i \in W_i$. By taking the difference of these two decompositions, we get $(w_1 - w'_1) + \cdots + (w_k - w'_k) = 0$, which then implies $w_i = w'_i$ for every i. Therefore, the decomposition of every vector in the sum is unique. □

REMARK 1.23. A consequence of this is that for any $0 < r < k$ we have

$$\bigoplus_{i=1}^{k} W_i = \bigoplus_{i=1}^{r} W_i \oplus \bigoplus_{i=r+1}^{k} W_i.$$

PROOF. We have to prove that $\bigoplus_{i=1}^{r} W_i \cap \bigoplus_{i=r}^{k} W_i = 0$. If v is in the intersection, we may uniquely decompose it as $w_1 + \cdots + w_r$ and as $w_{r+1} + \cdots + w_k$ with $w_i \in W_i$. Taking the difference, we get $0 = w_1 + \cdots + w_r - w_{r+1} - \cdots - w_k$. By uniqueness of the decomposition in $\bigoplus_{i=1}^{k} W_i$, we get $w_i = 0$ for every i. □

REMARK 1.24. Note that, by definition, $W_1 \oplus W_2 = W_2 \oplus W_1$ and that, by the last remark, $(W_1 \oplus W_2) \oplus W_3 = W_1 \oplus (W_2 \oplus W_3)$, where W_1, W_2 and W_3 are subspaces of V. Therefore, the direct sum is commutative and associative. It also have a "neutral element," namely the zero subspace $0 := \{0\}$.

REMARK 1.25 (Infinite sums). Let $(W_i)_{i \in S}$ be a, possibly infinite, collection of subspaces of a vector space V. Their sum is the subspace of V consisting of all vectors of the form $w_{i_1} + \cdots + w_{i_k}$, $w_{i_j} \in W_{i_j}$ for $j = 1, \ldots, k$ and some integer k. This sum is direct, denoted by

$$\bigoplus_{i \in S} W_i,$$

if each vector in it has a unique decomposition (or, equivalently, if $w_{i_1} + \cdots + w_{i_k} = 0$, $w_{i_j} \in W_{i_j}$, $i_j \neq i_{j'}$ for $j \neq j'$, implies $w_{i_j} = 0$ for all j).

Definition 1.26 (Direct sums of vector spaces). If V_1 and V_2 are vector spaces over the same field, we denote by $V_1 \oplus V_2$—and call it the direct sum of V_1 and V_2—the Cartesian product $V_1 \times V_2$ of pairs of elements of V_1 and V_2 with the following vector space structure:

$$(v_1, v_2)+(\widetilde{v}_1, \widetilde{v}_2) := (v_1+\widetilde{v}_1, v_2+\widetilde{v}_2),\ 0 := (0, 0),\ \lambda(v_1, v_2) := (\lambda v_1, \lambda v_2).$$

The spaces V_1 and V_2 are identified with the subspaces $\{(v, 0), v \in V_1\}$ and $\{(0, v), v \in V_2\}$ of $V := V_1 \oplus V_2$. Under this identification, $V_1 \cap V_2 = \{0\}$, so the notation for the direct sum of the vector spaces V_1 and V_2 fits with that of the direct sum of the subspaces V_1 and V_2 of V. This generalizes to a collection $V_1, \ldots, V_k$ of vector spaces over the same field. By $\bigoplus_{i=1}^k V_i$, we denote the Cartesian product $V_1 \times \cdots \times V_k$ with

$$\begin{aligned}(v_1, \ldots, v_k) + (\widetilde{v}_1, \ldots \widetilde{v}_k) &:= (v_1 + \widetilde{v}_1, \ldots, v_k + \widetilde{v}_k),\\ 0 &:= (0, \ldots, 0),\\ \lambda(v_1, \ldots, v_k) &:= (\lambda v_1, \ldots, \lambda v_k).\end{aligned}$$

Again, we may regard V_i as the subspace of $\bigoplus_{i=1}^k V_i$ consisting of $k$tuples with 0 in each position but the ith.

Note that $\mathbb{K}^n = \mathbb{K} \oplus \cdots \oplus \mathbb{K}$ with n summands. As a subspace of $\mathbb{K}^n$, the kth summand consists of vectors in which only the kth component may be different from zero.

Moreover, $\mathbb{K}^n = \mathbb{K}^r \oplus \mathbb{K}^l$ for all nonnegative integers r and l with $r + l = n$. In this case, as a subspace of $\mathbb{K}^n$, the first summand consists of vectors whose last l components are zero, and the second summand consists of vectors whose first r components are zero. We use this decomposition (with $l = 1$) for the following

Lemma 1.27. *Every subspace of $\mathbb{K}^n$ has a complement.*

Proof. Let W be a subspace of $\mathbb{K}^n$. We want to show that we can always find a subspace W' of $\mathbb{K}^n$ such that $W \oplus W' = \mathbb{K}^n$ (i.e., $W \cap W' = \{0\}$ and $W + W' = \mathbb{K}^n$).

If $n = 0$, there is nothing to prove, since necessarily $W = \{0\}$ and $W' = \{0\}$.

Otherwise, we prove the lemma by induction on $n > 0$. If $n = 1$, the proof is immediate: In case $W = \{0\}$, we take $W' = \mathbb{K}$. If, otherwise, W contains a nonzero vector $\boldsymbol{v}$, then $W = \mathbb{K}$, since every vector in $\mathbb{K}$ can be written as $\lambda \boldsymbol{v}$; therefore, $W' = \{0\}$.

Now assume we have proved the lemma for $\mathbb{K}^n$, and let W be a subspace of $\mathbb{K}^{n+1}$. Let W_1 be the subspace of vectors in W whose last component is zero and let W_2 be the subspace of vectors of W whose first n components are zero. We can view W_1 as a subspace of the first summand, $\mathbb{K}^n$, in the decomposition $\mathbb{K}^{n+1} = \mathbb{K}^n \oplus \mathbb{K}$ and W_2 as a subspace of the second summand, $\mathbb{K}$. By the induction assumption, there is a complement W_1' of W_1 in the first summand and a complement W_2' of W_2 in the second. Then $W_1' \oplus W_2'$ is a complement of W in $\mathbb{K}^n$. □

1.3. Linear maps

A map $F\colon V \to W$ between $\mathbb{K}$-vector spaces is called a linear map if

$$F(\lambda v + \mu \widetilde{v}) = \lambda F(v) + \mu F(\tilde{v})$$

for all $\lambda, \mu \in \mathbb{K}$ and all $v, \widetilde{v} \in V$.

EXAMPLES 1.28. Here are some examples of linear maps:

(1) The inclusion map of a subspace is linear.
(2) If V is the direct sum of vector spaces V_1 and V_2—i.e., $V = V_1 \oplus V_2$ as in Definition 1.26—then we have the linear maps, called canonical projections, $\pi_i\colon V \to V_i$, $i = 1, 2$, given by

$$\pi_i(v_1, v_2) = v_i.$$

If we regard V_1 and V_2 as subspaces of V, we may also regard the projections as linear maps $P_i\colon V \to V$:

$$P_1(v_1, v_2) = (v_1, 0) \quad \text{and} \quad P_2(v_1, v_2) = (0, v_2).$$

More precisely, $P_i = \iota_i \circ \pi_i$ where ι_i is the inclusion of V_i into V.
(3) Multiplication, from the left, by an $m \times n$ matrix defines a linear map $\mathbb{K}^n \to \mathbb{K}^m$.
(4) Multiplication, from the right, by an $m \times n$ matrix defines a linear map $(\mathbb{K}^m)^* \to (\mathbb{K}^n)^*$.
(5) The derivative defines a linear map $C^k(I) \to C^{k-1}(I)$, $f \mapsto f'$ (we assume $k \in \mathbb{N}_{>0} \cup \{\infty\}$).

REMARK 1.29. Here are some facts and notations.

(1) If F is linear, one often writes Fv instead of $F(v)$.
(2) The image of a linear map $F\colon V \to W$[5] is denoted by $\operatorname{im} F$ or $F(V)$ and is a subspace of W.

[5] i.e., the set of vectors $w \in W$ for which there is a $v \in V$ with $w = F(v)$

(3) The subset of elements of V mapped to 0 by a linear map $F: V \to W$ is denoted by $\ker F$ and is called its kernel. It is a subspace of V. A linear map F turns out to be injective iff $\ker F = \{0\}$.
(4) The composition of linear maps, say, $F: V \to W$ and $G: W \to Z$, is automatically linear. Instead of $G \circ F$ one often writes GF.
(5) If a linear map F is linear and invertible, its inverse map F^{-1} is automatically linear.
(6) A linear map $F: V \to W$ is also called a homomorphism from V to W.
(7) An invertible linear map $F: V \to W$ is also called an isomorphism from V to W.
(8) If an isomorphism from V to W exists, then V and W are called isomorphic and one writes $V \cong W$.
(9) A linear map $F: V \to V$ is also called an endomorphism of V or a linear operator (or just an operator) on V. If it is invertible, it is also called an automorphism. The identity map, denoted by Id or Id_V or 1, is an automorphism.
(10) If F is an endomorphism of V, a subspace W of V is called F-invariant if $F(W) \subseteq W$ (i.e., $F(w) \in W$ for every $w \in W$). The restriction of F to an invariant subspace W then yields an endomorphism of W.

We introduce the following sets of linear maps and their additional structures:

(1) $\mathrm{Hom}(V, W)$ is the set of all homomorphisms from V to W. If $F, G \in \mathrm{Hom}(V, W)$, we define $F + G \in \mathrm{Hom}(V, W)$ by

$$(F + G)(v) := F(v) + G(v).$$

We denote by 0 the zero homomorphism $0(v) = 0$ for all $v \in V$. With the scalar multiplication $(\lambda F)(v) := \lambda F(v)$, the set $\mathrm{Hom}(V, W)$ is a vector space over $\mathbb{K}$.
(2) $\mathrm{End}(V) = \mathrm{Hom}(V, V)$ is the set of all endomorphisms of V. As a particular case of the above, it is a vector space over $\mathbb{K}$. It is also a ring with one, where the multiplication is given by the composition and the one element is the identity map.
(3) $\mathrm{Aut}(V) \subset \mathrm{End}(V)$ is the set of all automorphisms of V. It is a group with multiplication given by composition.

Remark 1.30 (Injective linear maps). Note that a linear map F is injective iff $\ker F = 0$. In fact, if F is injective, then $Fv = 0 = F0$

implies $v = 0$. On the other hand, the equality $Fv = Fv'$ implies, by linearity, that $v - v' \in \ker F$, so if $\ker F = 0$ then we have $v = v'$.

DEFINITION 1.31 (Dual space). The vector space $\mathrm{Hom}(V, \mathbb{K})$ is usually denoted by V^* and called the dual space of V. Note that V^*, like every Hom space, is itself a vector space. An element α of V^* is a linear map $V \to \mathbb{K}$ and is usually called a linear functional. In addition to the notation $\alpha(v)$ to indicate $\alpha \in V^*$ evaluated on $v \in V$, one often writes (α, v).[6]

EXAMPLE 1.32. A row vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in (\mathbb{K}^n)^*$ defines a linear functional on $\mathbb{K}^n$ via

$$\begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \mapsto (\boldsymbol{\alpha}, \boldsymbol{v}) := \sum_{i=1}^{n} \alpha_i v^i. \tag{1.1}$$

One can show that these are all possible linear functionals on $\mathbb{K}^n$, so $(\mathbb{K}^n)^*$ is the dual of $\mathbb{K}^n$, which justifies the notation.

DEFINITION 1.33 (Einstein's convention). In these notes we will follow a very handy convention introduced by A. EINSTEIN according to which, whenever an index appears twice in an expression, once as a lower and once as an upper index, then a sum over that index is understood.

EXAMPLE 1.34. According to Einstein's convention, the evaluation of a row vector $\boldsymbol{\alpha}$ on a column vector $\boldsymbol{v}$, as in equation (1.1), is simply written as

$$(\boldsymbol{\alpha}, \boldsymbol{v}) = \alpha_i v^i.$$

DEFINITION 1.35 (Bidual space). The dual space of the dual space V^* of a vector space V is denoted by V^{**} and is called the bidual space of V.

REMARK 1.36. Note that every $v \in V$ defines a linear functional on V^* by $V^* \ni \alpha \mapsto v(\alpha) := \alpha(v)$. Therefore, we may regard V as a subspace of V^{**}. We will see (Proposition 1.62) that, if V is finite-dimensional, one actually has $V = V^{**}$.

REMARK 1.37 (Direct sum). Suppose $V = V_1 \oplus V_2$ as in Example 1.28.(2). Then we have the following relations among the projections:

$$P_1 + P_2 = 1, \quad P_1^2 = P_2, \quad P_2^2 = P_2, \quad P_1 P_2 = P_2 P_1 = 0.$$

[6]This notation actually indicates the induced bilinear map $V^* \times V \to \mathbb{K}$.

REMARK 1.38. More generally, an endomorphism P of V is called a projection if

$$P^2 = P.$$

Note that $Q := 1-P$ is also a projection and that we have $PQ = QP = 0$. This is related to the previous remark as follows: define $V_1 := \operatorname{im} P$ and $V_2 := \operatorname{im} Q$. Then we have $V = V_1 \oplus V_2$ and $P_1 = P$ and $P_2 = Q$.

REMARK 1.39 (The dual map). A linear map $F\colon V \to W$ induces a dual map $F^*\colon W^* \to V^*$ as follows: an element α of W^*—i.e., a linear functional on W—is mapped to $F^*\alpha \in V^*$ defined by

$$(F^*\alpha)(v) := \alpha(Fv).$$

Note that F^* is also linear. Moreover, F^{**} restricted to V is the map F again. We will see in Section 1.5.2 that the dual of a map is related to the transposition of matrices.

1.4. Bases

A basis of a $\mathbb{K}$-vector space V is a collection $(e_i)_{i\in S}$ of elements of V such that for every vector $v \in V$ there are are uniquely determined scalars $v^i \in \mathbb{K}$, only finitely many of which are different from zero, such that

$$v = \sum_i v^i e_i.$$

Note that, by omitting the zero summands, this is a sum of finitely many terms. The scalars v^i are called the components of v in the given basis. Using Einstein's convention (see Definition 1.33), we write the expansion of v in the basis $(e_i)_{i\in S}$ as

$$v = v^i e_i.$$

REMARK 1.40. In order to use Einstein's convention, one has to be consistent with the positioning of the indices. Typically we will use lower indices for basis elements and, consequently, upper indices for components of vectors. In some cases, see below, we use upper indices for basis elements and, consequently, lower indices for components of vectors.[7]

EXAMPLE 1.41 (The standard bases). The space $\mathbb{K}^n$ has the standard basis $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ where $\boldsymbol{e}_i$ denotes the column vector that has a 1

[7]As will be explained later, we use upper indices for a basis of a dual space.

in the ith position and a 0 otherwise. The space $(\mathbb{K}^n)^*$ also has a standard basis, now with upper indices $(\boldsymbol{e}^1, \dots, \boldsymbol{e}^n)$, where $\boldsymbol{e}^i$ denotes the row vector that has a 1 in the ith position and a 0 otherwise. Namely:

$$\boldsymbol{v} = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \qquad \Longrightarrow \qquad \boldsymbol{v} = v^i \boldsymbol{e}_i$$

$$\boldsymbol{v} = \begin{pmatrix} v_1, \dots, v_n \end{pmatrix} \qquad \Longrightarrow \qquad \boldsymbol{v} = v_i \boldsymbol{e}^i$$

REMARK 1.42 (Related concepts). Some important concepts are related to the notion of basis. All the vectors mentioned in the following list belong to a fixed $\mathbb{K}$-vector space V.

(1) A linear combination of a finite collection $v_1, \dots, v_k$ of vectors is a vector of the form $\lambda^i v_i$ with $\lambda^i \in \mathbb{K}$. The set

$$\mathrm{Span}\{v_1, \dots, v_k\} := \{\lambda^i v_i,\ \lambda_i \in \mathbb{K}\}$$

of all linear combinations of $v_1, \dots, v_k$ is called their span and is a subspace of V. If the set contains a single vector v, instead of $\mathrm{Span}\{v\}$ we also use the notation $\mathbb{K}v$, so

$$\mathbb{K}v = \{\lambda v,\ \lambda \in \mathbb{K}\}.$$

(2) A linear combination of a collection $(e_i)_{i \in S}$ of vectors is by definition a linear combination of a finite subcollection. When writing $\lambda^i v_i$, it is then assumed that only finitely many λ_is are different from zero. The set $\mathrm{Span}_{i \in S}\, e_i$ of linear combinations of the vectors in the collection is also a subspace of V.

(3) A collection $(e_i)_{i \in S}$ of vectors is called a system of generators for V if every vector of V can be expressed as a linear combination of the e_is (but we do not require uniqueness of this expression). In other words, $\mathrm{Span}_{i \in S}\, e_i = V$. Each e_i is called a generator.

(4) A collection $(e_i)_{i \in S}$ of vectors is called linearly independent if a linear combination can be zero only if all the coefficients are zero. That is, if

$$\lambda^i e_i = 0 \implies \lambda_i = 0\ \forall i.$$

The collection is called linearly dependent otherwise.

(5) A basis is then the same as a linearly independent system of generators.

One can esily see that the following hold:

(1) If $F\colon V \to W$ is injective and $(e_i)_{i \in S}$ is a linearly independent family of vectors in V, then $(Fe_i)_{i \in S}$ is a linearly independent family of vectors in W.

(2) If $F: V \to W$ is surjective and $(e_i)_{i\in S}$ is a system of generators in V, then $(Fe_i)_{i\in S}$ is a system of generators in W.

Therefore,

PROPOSITION 1.43. *If $F: V \to W$ is an isomorphism and $(e_i)_{i\in S}$ is a basis of V, then $(Fe_i)_{i\in S}$ is a basis of W.*

Moreover, one has the

THEOREM 1.44. *Any two bases of the same vector space have the same cardinality.*

If V admits a finite basis (i.e., a basis $(e_i)_{i\in S}$ with S a finite set), then it is called a finite-dimensional vector space; otherwise, it is called an infinite-dimensional vector space.

DIGRESSION 1.45 (Existence of bases). By definition a finite-dimensional vector space has a basis (actually, a finite one). By the axiom of choice one can prove that every vector space has a basis (actually, the existence of bases for all vector spaces is equivalent to the axiom of choice).

DEFINITION 1.46 (Dimension). If V is finite-dimensional with a basis of cardinality n (i.e., $|S| = n$), then we set

$$\dim V = n$$

and call it the dimension of V. Usually we then choose $S = \{1, \dots, n\}$ and denote the basis by $(e_1, \dots, e_n)$. Note that the trivial vector space $V = \{0\}$ is zero-dimensional. In particular, we have

$$\dim \mathbb{K}^n = \dim(\mathbb{K}^n)^* = n \quad \forall n \in \mathbb{N}.$$

If V is infinite-dimensional, we set

$$\dim V = \infty.$$

REMARK 1.47 (Dimension over a field). The same abelian group V may sometimes be regarded as a vector space over different fields $\mathbb{K}$. In this case, it is convenient to to remember which field we are considering when computing the dimension: we will write $\dim_{\mathbb{K}} V$ for the dimension of V as a $\mathbb{K}$-vector space.

REMARK 1.48 (Complex spaces as real spaces). In particular, a case we will often encounter is that of a complex vector space V (i.e., a vector space over $\mathbb{C}$). For every real λ, we still have the scalar multiplication $v \mapsto \lambda v$, so V may also be regarded as a real vector space (i.e., a vector space over $\mathbb{R}$). If $\mathcal{B}_{\mathbb{C}} = (e_1, \dots, e_n)$ is a basis of V as a complex vector space, then[8]

$$\mathcal{B}_{\mathbb{R}} = (e_1, \dots, e_n, \mathrm{i}e_1, \dots, \mathrm{i}e_n)$$

is a basis of V as a real one. In fact, every $v \in V$ may uniquely be expanded as $\lambda^i e_i$ with $\lambda^i \in \mathbb{C}$. Writing $\lambda^i = a^i + \mathrm{i}b^i$, with a^i and b^i real, we get the expansion $v = a^i e_i + b^i \mathrm{i}e_i$, so $\mathcal{B}_{\mathbb{R}}$ is a system of generators over $\mathbb{R}$.[9] Moreover, if $a^i e_i + b^i \mathrm{i}e_i = 0$ for $a^i, b^i \in \mathbb{R}$, then $\lambda^i e_i = 0$; linear independence of $\mathcal{B}_{\mathbb{C}}$ over $\mathbb{C}$ implies, for all i, $\lambda^i = 0$ and, therefore, $a^i = b^i = 0$, which is linear independence of $\mathcal{B}_{\mathbb{R}}$ over $\mathbb{R}$. Therefore, $\mathcal{B}_{\mathbb{R}}$ is basis. We conclude that $\dim_{\mathbb{C}} V = n$ implies that $\dim_{\mathbb{R}} V = 2n$; i.e.,

$$\dim_{\mathbb{R}} V = 2 \dim_{\mathbb{C}} V.$$

One can prove that a linearly independent collection $(e_1, \dots, e_n)$ in an n-dimensional space is automatically a basis. This in particular implies the following

PROPOSITION 1.49. *If V is a finite-dimensional vector space and W is a subspace of V of the same dimension, then $W = V$.*

REMARK 1.50 (Direct sums and bases). There is a strong relationship between the notions of direct sums and bases. Namely, by Definition 1.21, a collection $(v_1, \dots, v_n)$ is a basis of V iff $V = \bigoplus_{i=1}^n \mathbb{K} v_i$.

REMARK 1.51 (Union of bases). Another relation is the following. If $V = \bigoplus_{i=1}^k W_i$ and $\mathcal{B}_i = (v_{i,j})_{j=1,\dots,d_i}$ is a basis of W_i, then

$$\mathcal{B} := \cup_{i=1}^k \mathcal{B}_i = (v_{i,j})_{\substack{i=1,\dots,k \\ j=1,\dots,d_i}}$$

is a basis of V. As a consequence,

$$\dim \bigoplus_{i=1}^k W_i = \sum_{i=1}^k \dim W_i. \tag{1.2}$$

PROOF. By Definition 1.21, every $v \in V$ uniquely decomposes as $v = w_1 + \dots + w_k$ with $w_i \in W_i$. By definition of basis, every w_i uniquely decomposes as $w_i = \sum_{j=1}^{d_i} \alpha_{i,j} v_{i,j}$. Therefore, v uniquely decomposes as $v = \sum_{i=1}^k \sum_{j=1}^{d_i} \alpha_{i,j} v_{i,j}$. □

[8]Here i denotes the imaginary unit, and $\mathrm{i}e_j$ is the scalar multiplication of the scalar i with the vector e_j.

[9]A standard convention in mathematics is to use italic characters for variables and roman characters for constants. The imaginary unit, being a constant, is then denoted by i, whereas i is a variable, like, e.g., the index in e_i. By handwriting it is however better to avoid using a variable i when the imaginary unit also appers.

REMARK 1.52 (The basis isomorphism). A basis $\mathcal{B} = (e_1, \dots, e_n)$ of V determines a linear map

$$\phi_{\mathcal{B}} \colon \mathbb{K}^n \to V$$

by

$$\phi_{\mathcal{B}} \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} := v^i e_i.$$

In particular, we have

$$\phi_{\mathcal{B}} \boldsymbol{e}_i = e_i$$

for all i. Note that $\phi_{\mathcal{B}}$ is an isomorphism with inverse $\phi_{\mathcal{B}}^{-1} \colon V \to \mathbb{K}^n$ the map that sends $v \in V$ to the column vector with components the components of v in the basis $\mathcal{B}$. The vector $\phi_{\mathcal{B}}^{-1} v$ is called the coordinate vector of v. Note that we then have

$$\dim V = n \iff V \cong \mathbb{K}^n.$$

REMARK 1.53. If V and W have the same finite dimension, then, by composition of the above, we get an isomorphism $V \to W$. Note that such an isomorphism depends on the choice of bases.

REMARK 1.54 (Bases and frames). When we define the isomorphism of Remark 1.52, the order in which we take the basis vectors $(e_1, \dots, e_n)$ matters. This is an additional choice to just the notion of a basis (which is by definition a collection, i.e., a set, of linearly independent generators). A basis with a choice of ordering is more precisely called a frame. According to a general habit, we will be sloppy about it and speak of a basis also when we actually mean a frame as, e.g., in Remark 1.52.

An immediate consequence of the basis isomorphism is the following

PROPOSITION 1.55. *Every subspace of a finite-dimensional vector space has a complement.*

PROOF. Let V be n-dimensional, and let Z be a subspace of V. By choosing a basis $\mathcal{B}$, we have the isomorphism $\phi_{\mathcal{B}} \colon \mathbb{K}^n \to V$. Then $W := \phi_{\mathcal{B}}^{-1}(Z)$ is a subspace of $\mathbb{K}^n$. By Lemma 1.27, it has a complement W'. Finally, $\phi_{\mathcal{B}}(W')$ is a complement of Z. □

DIGRESSION 1.56. By the axiom of choice, one can show that a subspace of any vector space has a complement.

DEFINITION 1.57 (Change of basis). If $\mathcal{B}$ and $\mathcal{B}'$ are bases of an n-dimensional space V, the composition

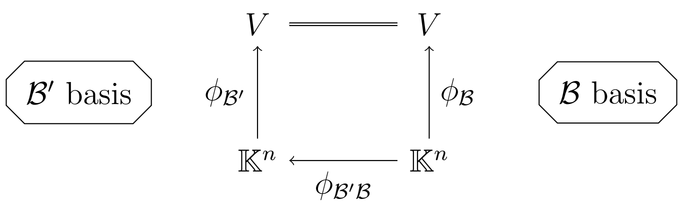
$$\phi_{\mathcal{B}'\mathcal{B}} := \phi_{\mathcal{B}'}^{-1} \phi_{\mathcal{B}} \in \operatorname{Aut}(\mathbb{K}^n)$$

is called the corresponding change of basis.

REMARK 1.58. If you have a vector $v \in V$, then $\phi_{\mathcal{B}}^{-1}v$ is the column vector of its components in the basis $\mathcal{B}$. The column vector $\phi_{\mathcal{B}'}^{-1}v$ of its components in the basis $\mathcal{B}'$ is then related to $\phi_{\mathcal{B}}^{-1}v$ by

$$\phi_{\mathcal{B}'}^{-1}v = \phi_{\mathcal{B}'\mathcal{B}}\phi_{\mathcal{B}}^{-1}v.$$

Therefore, $\phi_{\mathcal{B}'\mathcal{B}}$ maps the coordinate vector in the $\mathcal{B}$ basis to the coordinate vector in the $\mathcal{B}'$ basis. (A more descriptive, but also more cumbersome, notation would be $\phi_{\mathcal{B}'\leftarrow\mathcal{B}}$ instead of $\phi_{\mathcal{B}'\mathcal{B}}$.)



REMARK 1.59 (The dual basis). A basis $\mathcal{B} = (e_1, \dots, e_n)$ of V allows defining uniquely the components v^i of any vector v. The map

$$\begin{array}{rccc} e^i\colon & V & \to & \mathbb{K} \\ & v & \mapsto & v^i \end{array}$$

is linear for every i. The collection $\mathcal{B}^* := (e^1, \dots, e^n)$ of linear functionals is called the dual basis of V^*—more precisely, the basis of V^* dual to $\mathcal{B}$—and satisfies, by definition,

$$e^i(e_j) = \delta^i_j := \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$$

where δ^i_j is called the Kronecker delta. It is indeed a basis of V^*. In fact, if $\alpha := \alpha_i e^i = 0$, then $0 = \alpha(e_j) = \alpha_j$ for every j, so $(e^1, \dots, e^n)$ is linearly independent. Moreover, given any $\alpha \in V^*$, we get

$$\alpha(v) = \alpha(v^i e_i) = v^i\alpha(e_i) = e^i(v)\alpha(e_i) = (\alpha(e_i)e^i)(v),$$

which shows that $(e^1, \dots, e^n)$ is a system of generators and that, in particular, we can compute the component α_i of α as $\alpha(e_i)$.

EXAMPLE 1.60. The basis $(\boldsymbol{e}^1, \dots, \boldsymbol{e}^n)$ of $(\mathbb{K}^n)^*$ is the dual basis to $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$.

More generally, we have proved the

PROPOSITION 1.61. *If V is a finite-dimensional vector space, then*

$$\dim V^* = \dim V.$$

This also implies that $\dim V^{**} = \dim V$. As a consequence of Remark 1.36 and of Proposition 1.49, we get the

PROPOSITION 1.62. *If V is a finite-dimensional vector space, then*

$$V^{**} = V.$$

REMARK 1.63 (Canonical and noncanonical maps). A linear map is called canonical if it does not depend on any additional structure. For example, if W is a subspace of V, the inclusion map is canonical. If $V = W_1 \oplus W_2$, the projections from V to W_1 and W_2 are also canonical. If W is a subspace of V, we can always find a complement W', so we can write $V = W \oplus W'$ and, therefore, get a projection $V \to W$. This projection is not canonical because it depends on the choice of a complement. Similarly, we saw that that every element V defines a linear functional on V^*, so we have a canonical inclusion map of V into V^{**}. If V is finite dimensional, we then have a canonical isomorphism between V and V^{**}: we therefore write $V = V^{**}$. On the other hand, by Proposition 1.61 and Remark 1.53, we also have an isomorphism between V and V^*, but this is not canonical because it depends on the choice of a basis. Explicitly, the map $V \to V^*$ sends $v^i e_i$ to $\sum_{i=1}^n v^i e^i$ (not that we cannot use Einstein's convention in this case).

EXAMPLE 1.64. With respect to the standard basis, the isomorphism $\mathbb{K}^n \xrightarrow{\sim} (\mathbb{K}^n)^*$, as at the end of the previous remark, is the transposition map $\boldsymbol{v} \mapsto \boldsymbol{v}^\mathsf{T}$:

$$\begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \mapsto (v^1, \dots, v^n).$$

1.5. Representing matrices

If $F\colon V \to W$ is a linear map and $(e_i)_{i \in S}$ is a basis of V, then the values of F on the e_is are enough to reconstruct F; in fact, every $v \in V$ is uniquely expanded as $v^i e_i$, so by linearity we get

$$F(v) = v^i F(e_i).$$

Vice versa, we can define a linear map $F\colon V \to W$ by specifying $w_i := F(e_i) \in W$ for all $i \in S$.

If $(\bar{e}_{\bar{\imath}})_{\bar{\imath} \in \bar{S}}$ is a basis of W, we may expand $F(e_i)$ as

$$F(e_i) = A_i^{\bar{\imath}}\, \bar{e}_{\bar{\imath}}, \tag{1.3}$$

where, by Einstein's convention, a sum over $\bar{\imath}$ is understood. The scalars $A_i^{\bar{\imath}}$ are the entries of the representing matrix $\boldsymbol{A}$ of F. On a generic vector

$v = v^i e_i \in V$ we then get

$$F(v) = v^i A_i^{\bar{\imath}}\, \bar{e}_{\bar{\imath}}.$$

We now assume that V and W are finite-dimensional: say, $\dim V = n$ and $\dim W = m$. Then

$$\boldsymbol{A} = (A_i^{\bar{\imath}})_{\substack{i=1,\dots,n\\ \bar{\imath}=1,\dots,m}} = \begin{pmatrix} A_1^1 & A_2^1 & \cdots & A_n^1 \\ A_1^2 & A_2^2 & \cdots & A_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ A_1^m & A_2^m & \cdots & A_n^m \end{pmatrix}$$

is called an $m \times n$ matrix.

Remark 1.65. Note that we consistently put the indices in upper and lower positions. We will also encounter matrices with only lower indices representing bilinear forms.

Remark 1.66 (Normal form). Given a linear map $F\colon V \to W$ between finite-dimensional vector spaces, one can always find bases $e_1, \dots, e_n$ of V and $\bar{e}_1, \dots, \bar{e}_m$ of W such that the representing matrix of F reads

$$\boldsymbol{A} = \begin{pmatrix} \mathbf{1}_r & \mathbf{0}_{r,n-r} \\ \mathbf{0}_{m-r,r} & \mathbf{0}_{m-r,n-r} \end{pmatrix},$$

where $\mathbf{1}_r$ denotes the $r \times r$ identity matrix and $\mathbf{0}_{i,j}$ denotes the $i \times j$ zero matrix. Here $r = \dim \operatorname{im} F$ is called the rank of F. From this presentation it also follows that the kernel of F corresponds, under the isomorphism to $\mathbb{K}^n$ induced by the basis, to vectors of the form

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ v^{r+1} \\ \vdots \\ v^n \end{pmatrix},$$

which show that $\dim \ker F = n - r$. We thus get the dimension formula

$$\dim \ker F + \dim \operatorname{im} F = \dim V \tag{1.4}$$

for any linear map $F\colon V \to W$.

If $F\colon V \to W$ is an isomorphism, then $\ker F = 0$ (see Remark 1.30) and $\operatorname{im} F = W$, so we get $\dim V = \dim W$. If, on the other hand, V and W have the same finite dimension n, then each of them is isomorphic to $\mathbb{K}^n$ by the basis isomorphism of Remark 1.52. Therefore,

Proposition 1.67. *Two finite-dimensional vector spaces are isomorphic iff they have the same dimension.*

1.5.1. Operations. If $\boldsymbol{B}$ is the representing matrix of another linear map $G\colon V \to W$, with respect to the same bases, the representing matrix of $F+G$ is $\boldsymbol{A}+\boldsymbol{B}$, where addition is defined entry-wise:

$$(A+B)^{\bar{\imath}}_i = A^{\bar{\imath}}_i + B^{\bar{\imath}}_i.$$

If G is instead a linear map $W \to Z$, the collection $(\tilde{e}_{\tilde{\imath}})_{\tilde{\imath}=1,\dots,k}$ is a basis of Z, and $\boldsymbol{B}$ is the corresponding $k\times m$ representing matrix of G, then the representing matrix of GF is the $k \times n$ matrix $\boldsymbol{BA}$, where the matrix multiplication is defined by[10]

$$(BA)^{\tilde{\imath}}_i = B^{\tilde{\imath}}_{\bar{\imath}} A^{\bar{\imath}}_i.$$

1.5.2. Duals. If $F^*\colon W^* \to V^*$ is the dual map, see Remark 1.39, of a linear map $F\colon V\to W$ with representing matrix $\boldsymbol{A}$, then we get, for $\alpha = \alpha_{\bar{\jmath}}\,\bar{e}^{\bar{\jmath}} \in W^*$ and $v = v^i e_i \in V$,

$$F^*(\alpha)(v) = \alpha(Fv) = \alpha_{\bar{\jmath}}\,\bar{e}^{\bar{\jmath}}(v^i A^{\bar{\imath}}_i\,\bar{e}_{\bar{\imath}}) = \alpha_{\bar{\imath}}\,A^{\bar{\imath}}_i\,v^i,$$

where $\bar{e}^1,\dots,\bar{e}^m$ is the dual basis to $\bar{e}_1,\dots,\bar{e}_m$. To compute the ith component of $F^*\alpha$, we evaluate on $v=e_i$ getting $(F^*\alpha)_i = \alpha_{\bar{\imath}}\,A^{\bar{\imath}}_i$, so

$$F^*(\alpha) = \alpha_{\bar{\imath}} A^{\bar{\imath}}_i\,e^i.$$

To get the representing matrix of F^*, we then compute

$$F^*(\bar{e}^{\bar{\imath}}) = A^{\bar{\imath}}_i\,e^i. \tag{1.5}$$

The difference between (1.3) and (1.5) is that in the former we sum over the upper index whereas in the latter we sum over the lower index of the matrix $\boldsymbol{A}$. In the usual notation with two lower indices, this correponds to summing over the first or second index, and the two representing matrices are related by the exchange of the indices—the operation known as transposition: if $\boldsymbol{A}$ is the representing matrix of F, with respect to some bases, then $\boldsymbol{A}^\mathsf{T}$ is the representing matrix of F^* with respect to the dual bases.

Remark 1.68. When working with vector spaces of row or column vectors, the standard basis is always assumed, unless otherwise specified. A linear map between such subspaces is then always understood as the corresponding matrix. We would in particular write $\boldsymbol{A}\colon \mathbb{K}^n \to \mathbb{K}^m$ to denote the linear map with representing matrix $\boldsymbol{A}$ with respect to the standard bases, i.e., the map $\boldsymbol{v}\mapsto \boldsymbol{Av}$, where we use matrix multiplication. Note that the jth row of $\boldsymbol{A}$ is equal to $\boldsymbol{A}\boldsymbol{e}_j$. One can also easily see that the dual map, acting on row vectors, is instead given by $\boldsymbol{\alpha}\mapsto\boldsymbol{\alpha A}$.

[10]Comparing with the usual way of writing a product of matrices, we see here that the upper index is the first index and the lower index is the second index.

1.5.3. Change of basis. If we want to keep track of the chosen bases, we need a more descriptive notation. Let $\mathcal{B} = (e_1, \dots, e_n)$ be the chosen basis of V and $\bar{\mathcal{B}} = (\bar{e}_1, \dots, \bar{e}_m)$ the chosen basis of W. The representing matrix of a linear map $F\colon V \to W$ with respect to the bases $\mathcal{B}$ and $\bar{\mathcal{B}}$ is then denoted by $F_{\bar{\mathcal{B}}\mathcal{B}}$.

As we identify linear maps $\mathbb{K}^n \to \mathbb{K}^m$ with their representing matrices with respect to the standard bases, we may regard $F_{\bar{\mathcal{B}}\mathcal{B}}$ as an $m \times n$ matrix or, equivalently, as a linear map. In the latter case, we have, in terms of the basis isomorphisms of Remark 1.52,

$$F_{\bar{\mathcal{B}}\mathcal{B}} = \phi_{\bar{\mathcal{B}}}^{-1} F \phi_{\mathcal{B}}.$$

If we choose another basis $\mathcal{B}'$ of V and another basis $\bar{\mathcal{B}}'$ of W, we then have, using the notation of Definition 1.57,

$$F_{\bar{\mathcal{B}}'\mathcal{B}'} = \phi_{\bar{\mathcal{B}}'}^{-1} F \phi_{\mathcal{B}'} = \phi_{\bar{\mathcal{B}}'}^{-1} \phi_{\bar{\mathcal{B}}} F_{\bar{\mathcal{B}}\mathcal{B}} \phi_{\mathcal{B}}^{-1} \phi_{\mathcal{B}'} = \phi_{\bar{\mathcal{B}}'\bar{\mathcal{B}}} F_{\bar{\mathcal{B}}\mathcal{B}} \phi_{\mathcal{B}\mathcal{B}'},$$

i.e.,

$$F_{\bar{\mathcal{B}}'\mathcal{B}'} = \phi_{\bar{\mathcal{B}}'\bar{\mathcal{B}}} F_{\bar{\mathcal{B}}\mathcal{B}} \phi_{\mathcal{B}\mathcal{B}'}.$$

It is easy to remember this formula, for it looks similar to the formula for matrix product, with indices replaced by bases.

Remark 1.69 (Equivalence of matrices). The formula for the change of bases motivates the following definition of equivalence of matrices. Two $m \times n$ matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are called equivalent if there is an invertible $m \times m$ matrix $\boldsymbol{T}$ and an invertible $n \times n$ matrix $\boldsymbol{S}$ such that

$$\boldsymbol{A} = \boldsymbol{T}\boldsymbol{B}\boldsymbol{S}.$$

By this definition, any two representing matrices of the same linear map are equivalent, as we may see by setting $\boldsymbol{A} = F_{\bar{\mathcal{B}}'\mathcal{B}'}$, $\boldsymbol{B} = F_{\bar{\mathcal{B}}\mathcal{B}}$, $\boldsymbol{S} = \phi_{\mathcal{B}\mathcal{B}'}$, and $\boldsymbol{T} = \phi_{\bar{\mathcal{B}}'\bar{\mathcal{B}}}$. Note that, explicitly, we have

$$A^{\bar{\imath}'}_{i'} = T^{\bar{\imath}'}_{\bar{\imath}} B^{\bar{\imath}}_{i} S^{i}_{i'}.$$

Also useful are the formulae

$$e'_{i'} = S^i_{i'} e_i \quad \text{and} \quad \bar{e}_{\bar{\imath}} = T^{\bar{\imath}'}_{\bar{\imath}} \bar{e}'_{\bar{\imath}'}.$$

Let us prove the first (the second is analogous):

$$S^i_{i'} e_i = S^i_{i'} \phi_{\mathcal{B}} \boldsymbol{e}_i = \phi_{\mathcal{B}}(S^i_{i'} \boldsymbol{e}_i) = \phi_{\mathcal{B}}(\phi_{\mathcal{B}\mathcal{B}'} \boldsymbol{e}_{i'}) = \phi_{\mathcal{B}'} \boldsymbol{e}_{i'} = e'_{i'},$$

where we also used the linearity of $\phi_{\mathcal{B}}$ and the definition $\phi_{\mathcal{B}\mathcal{B}'} = \phi_{\mathcal{B}}^{-1} \phi_{\mathcal{B}'}$.

1.5.4. Endomorphisms. If F is an endomorphism of a finite-dimensional vector space V, one usually chooses the same basis (say, $\mathcal{B}$) for V as source and target space. In this case, the representing matrix of V, now a square matrix,[11] with respect to the basis $\mathcal{B}$ is written $F_\mathcal{B}$ and we have

$$F_\mathcal{B} = \phi_\mathcal{B}^{-1} F \phi_\mathcal{B}.$$

If we pass to another basis (say, $\mathcal{B}'$), we then have

$$F_{\mathcal{B}'} = \phi_{\mathcal{B}'\mathcal{B}} F_\mathcal{B} \phi_{\mathcal{B}\mathcal{B}'}.$$

Observing that the isomorphisms $\phi_{\mathcal{B}'\mathcal{B}}$ and $\phi_{\mathcal{B}\mathcal{B}'}$ are inverse to each other, we can also write

$$F_{\mathcal{B}'} = \phi_{\mathcal{B}\mathcal{B}'}^{-1} F_\mathcal{B} \phi_{\mathcal{B}\mathcal{B}'}.$$

REMARK 1.70 (Similarity of matrices). The formula for the change of basis for the representing matrix of an endomorphism motivates the following definition of similarity of matrices. Two $n \times n$ matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are called similar if there is an invertible $n \times n$ matrix $\boldsymbol{S}$ such that

$$\boldsymbol{A} = \boldsymbol{S}^{-1} \boldsymbol{B} \boldsymbol{S}.$$

By this definition, any two representing matrices of the same endomorphism are similar.

1.5.5. Bilinear forms. A bilinear form on a $\mathbb{K}$-vector space V is a map $B\colon V \times V \to \mathbb{K}$ that is linear in both arguments; viz.,

$$B(\lambda_1 v_1 + \lambda_2 v_2, w) = \lambda_1 B(v_1, w) + \lambda_2 B(v_2, w),$$
$$B(v, \lambda_1 w_1 + \lambda_2 w_2) = \lambda_1 B(v, w_1) + \lambda_2 B(v, w_2).$$

If $\mathcal{B} = \{e_1, \dots, e_n\}$ is a basis of V, the representing matrix $\boldsymbol{B}$ of a bilinear form B has the entries

$$B_{ij} := B(e_i, e_j). \tag{1.6}$$

Note that, consistently with the r.h.s., we use lower indices for the entries of the representing matrix.

Now consider a new basis $\mathcal{B}' = \{e_1', \dots, e_n'\}$ and denote by $\boldsymbol{S}$ the matrix representing, in the standard basis of $\mathbb{K}^n$, the change of basis $\phi_{\mathcal{B}\mathcal{B}'}$; i.e., as we showed above,

$$e_{i'}' = S_{i'}^i e_i.$$

Denoting by $\boldsymbol{B}'$ the representing matrix of B in the basis $\mathcal{B}'$, we then have

$$B_{i'j'}' = S_{i'}^i B_{ij} S_{j'}^j.$$

[11] i.e., a matrix with the same number of rows and columns

Remark 1.71 (Congruency of matrices). The formula for the change of basis of a bilinear form motivates the following definition of congruency of matrices. Two $n \times n$ matrices $\boldsymbol{B}$ and $\boldsymbol{B}'$ are called congruent if there is an invertible $n \times n$ matrix $\boldsymbol{S}$ such that

$$\boldsymbol{B}' = \boldsymbol{S}^\mathsf{T}\boldsymbol{B}\boldsymbol{S},$$

where $^\mathsf{T}$ denotes transposition. By this definition, any two representing matrices of the same bilinear form are congruent.[12]

1.6. Traces and determinants

Trace and determinant are particularly important functions on square matrices which we review here.

The trace of an $n \times n$ matrix $\boldsymbol{A}$ is the sum of its diagonal elements. If we write $\boldsymbol{A} = (A^i_j)$, then

$$\operatorname{tr}\boldsymbol{A} = A^i_i,$$

where we used Einstein's convention. If we we write $\boldsymbol{A} = A_{ij}$, then we have to write the sum symbol explicitly: $\operatorname{tr}\boldsymbol{A} = \sum_{i=1}^n A_{ii}$. (The fact that this second notation is incompatible with Einstein's convention is related to the fact that the trace of a bilinear form is not a natural operation.)

Immediate properties of the trace are

(1) $\operatorname{tr}\boldsymbol{A}^\mathsf{T} = \operatorname{tr}\boldsymbol{A}$ for every $n \times n$ matrix, and
(2) $\operatorname{tr}\boldsymbol{A}\boldsymbol{B} = \operatorname{tr}\boldsymbol{B}\boldsymbol{A}$ for all $n \times n$ matrices $\boldsymbol{A}, \boldsymbol{B}$.

The second property implies $\operatorname{tr}\boldsymbol{S}^{-1}\boldsymbol{B}\boldsymbol{S} = \operatorname{tr}\boldsymbol{B}$, so similar matrices—see Remark 1.70—have the same trace. Therefore, we can make the following

Definition 1.72. The trace of an endomorphism F of a finite-dimensional vector space V is the trace of any of its representing matrices: $\operatorname{tr} F := \operatorname{tr} F_\mathcal{B}$, where $\mathcal{B}$ is any basis of V.

Congruent matrices may on the other hand have different traces. For this reason the trace of a bilinear form is not a well-defined concept, as it depends on the explicit choice of a basis.[13]

We conclude with a few additional properties of the trace:

[12] Note that in the explicit formula (1.6) it is the upper index of the first $\boldsymbol{S}$ that is the same as the first index of $\boldsymbol{B}$, whereas in the usual product of matrices—see footnote 10—it should be the lower index to be involved. It is for this reason that the first $\boldsymbol{S}$ is actually transposed.

[13] It is well-defined if we may restrict to a special class of bases that are related to each other by an orthogonal transformation, i.e., if we only allow congruences $\boldsymbol{B}' = \boldsymbol{S}^\mathsf{T}\boldsymbol{B}\boldsymbol{S}$ with $\boldsymbol{S}\boldsymbol{S}^\mathsf{T} = \mathbf{1}$.

(1) The trace is linear: $\operatorname{tr}(\lambda \boldsymbol{A}+\lambda \boldsymbol{B}) = \lambda \operatorname{tr} \boldsymbol{A} + \mu \operatorname{tr} \boldsymbol{B}$ for all $n \times n$ matrices $\boldsymbol{A}, \boldsymbol{B}$ and for all scalars λ, μ.
(2) $\operatorname{tr} \mathbf{1}_n = n$ if $\mathbf{1}_n$ is the identity matrix on $\mathbb{K}^n$.
(3) $\operatorname{tr}\colon \operatorname{End}(V) \to \mathbb{K}$ is a linear map: $\operatorname{tr}(\lambda F+\lambda G) = \lambda \operatorname{tr} F + \mu \operatorname{tr} G$ for all endomorphisms F, G of the finite-dimensional space V and for all scalars λ, μ.
(4) $\operatorname{tr} \operatorname{Id}_V = \dim V$.

The **determinant** of a square matrix can be uniquely characterized by some properties or can be, equivalently, defined by an explicit formula. The defining properties are the following:

(1) The determinant is linear with respect to every column of the matrix.
(2) The determinant vanishes if any two column of the matrix are equal.
(3) The determinant of the identity matrix is 1.

One can show that, for every n, there is a unique map $\operatorname{Mat}_{n\times n}(\mathbb{K}) \to \mathbb{K}$ satisfying these three properties, and this map is called the determinant. In words, one says that the determinintant is the unique alternating multilinear normalized map on the columns of a square matrix. Some derived properties are the following:

(D.1) Properties (1) and (2) above hold taking rows instead of colums.
(D.2) $\det \boldsymbol{A}^\mathsf{T} = \det \boldsymbol{A}$ for every $n \times n$ matrix $\boldsymbol{A}$.
(D.3) If two columns (or two rows) are exchanged the determinant changes sign.
(D.4) If one adds to a column a linear combination of the other columns (or to a row a linear combination of the other rows) the determinant does not change.
(D.5) $\det(\lambda \boldsymbol{A}) = \lambda^n \det \boldsymbol{A}$ for every $n \times n$ matrix $\boldsymbol{A}$ and every scalar λ.
(D.6) The determinant of a diagonal matrix or of an upper triangular matrix or of a lower triangular matrix is equal to the product of its diagonal elements.
(D.7) $\det \boldsymbol{AB} = \det \boldsymbol{A} \det \boldsymbol{B}$ for all $n \times n$ matrices $\boldsymbol{A}, \boldsymbol{B}$.
(D.8) $\det \boldsymbol{A} \neq 0$ iff $\boldsymbol{A}$ is invertible. By the previous property, we also see that in this case, $\det \boldsymbol{A}^{-1} = (\det \boldsymbol{A})^{-1}$.
(D.9) A collection $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ of vectors of $\mathbb{K}^n$ is a basis iff the determinant of the matrix $\boldsymbol{S}$ whose ith column is $\boldsymbol{v}_i$ is different from zero. (Note that $\boldsymbol{v}_i = \boldsymbol{S}\boldsymbol{e}_i$ for all i.)
(D.10) The determinant of a block-diagonal matrix is the product of the determinants of its blocks: $\det \left(\begin{smallmatrix} \boldsymbol{A} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{D} \end{smallmatrix}\right) = \det \boldsymbol{A} \det \boldsymbol{D}$, where $\boldsymbol{A}$ and $\boldsymbol{D}$ are square matrices.

(D.11) More generally, $\det\left(\begin{smallmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \mathbf{0} & \boldsymbol{D}\end{smallmatrix}\right) = \det\boldsymbol{A}\det\boldsymbol{D}$, where $\boldsymbol{A}$ and $\boldsymbol{D}$ are square matrices.

The last property is derived from the others by the following remarks. We should distinguish the case when $\boldsymbol{A}$ is invertible and when it is not. In the second case, there is some nonzero vector $\boldsymbol{v}$ in the kernel of $\boldsymbol{A}$. By completing $\boldsymbol{v}$ with zeros, we get a nonzero vector in the kernel of $\det\left(\begin{smallmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \mathbf{0} & \boldsymbol{D}\end{smallmatrix}\right)$, which is then also not invertible. In this case, both sides of the equality vanish by (D.8). If, on the other hand, $\boldsymbol{A}$ is invertible, we can write $\left(\begin{smallmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \mathbf{0} & \boldsymbol{D}\end{smallmatrix}\right) = \left(\begin{smallmatrix}\boldsymbol{A} & \mathbf{0}\\ \mathbf{0} & \boldsymbol{D}\end{smallmatrix}\right)\left(\begin{smallmatrix}\mathbf{1} & \boldsymbol{A}^{-1}\boldsymbol{B}\\ \mathbf{0} & \mathbf{1}\end{smallmatrix}\right)$. The determinant of the first matrix on the right hand side is $\det\boldsymbol{A}\det\boldsymbol{D}$ by (D.10), whereas the determinant of the second is 1 by (D.6).

The determinant of an $n \times n$ matrix $\boldsymbol{A} = A^i_j$ can also be explicitly computed by the Leibniz formula

$$\det\boldsymbol{A} = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma A^1_{\sigma(1)} \dots A^n_{\sigma(n)}, \tag{1.7}$$

where $\operatorname{sgn}\sigma$ is the sign of the permutation σ.[14] In particular,

$$\det\begin{pmatrix} a & b\\ c & d\end{pmatrix} = ad - bc.$$

The determinant can also be computed in terms of the Laplace expansion along the ith row

$$\det\boldsymbol{A} = \sum_{j=1}^{n}(-1)^{i+j}A^i_j d^i_j,$$

where d^i_j is the determinant of the matrix obtained by removing the ith row and the jth column from $\boldsymbol{A}$. For example, the Laplace expansion of a 3×3 matrix along the first row is

$$\det\begin{pmatrix} a & b & c\\ d & e & f\\ g & h & i\end{pmatrix} = a\det\begin{pmatrix} e & f\\ h & i\end{pmatrix} - b\det\begin{pmatrix} d & f\\ g & i\end{pmatrix} + c\det\begin{pmatrix} d & e\\ g & h\end{pmatrix}.$$

An analogus formula for the expansion along a column also exists as a consequence of (D.2).

The determinant may also be used to compute the inverse of an invertible matrix as

$$\boldsymbol{A}^{-1} = \frac{1}{\det\boldsymbol{A}}\operatorname{adj}\boldsymbol{A},$$

[14] Every permutation may be written, in a nonunique way, as a product of transpositions, i.e., permutations that exchange exactly two elements. The parity of the number of occurring transpositions does not depend on the decomposition. The sign of a permutation is then defined as -1 to the number of occurring transpositions.

where adj $\boldsymbol{A}$ denotes the adjugate matrix of $\boldsymbol{A}$, i.e., the matrix whose (i,j) entry is $(-1)^{i+j}$ times the determinant d_i^j of the matrix obtained by removing the the jth row and and ith column from $\boldsymbol{A}$. This formula is theoretically important—it shows, e.g., that the entries of $\boldsymbol{A}^{-1}$ are rational functions of the entries of $\boldsymbol{A}$—but practically not so useful, apart from the 2×2 case, which is also easy to remember:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \tag{1.8}$$

Namely, apart from dividing by the determinant, we just have to swap the diagonal entries and change the sign of the off diagonal entries.

Determinants may also be used, via property (D.8), to establish whether a linear map $F\colon V \to W$ between finite-dimensional vector spaces is invertible. By Proposition 1.67, we know that a necessary condition is that V and W have the same dimension, say, n. A representing matrix of F is then an $n \times n$ matrix, which by (D.8) is invertible iff its determinant is different from zero.[15] On the other hand, F is invertible iff any of its representing matrices is so. Therefore, we have the

Proposition 1.73. *A linear map $F\colon V \to W$ between finite-dimensional vector spaces is invertible iff $\dim V = \dim W$ and the determinant of any of its representing matrices is different from zero.*

It follows from (D.7) and (D.8) that $\det \boldsymbol{S}^{-1}\boldsymbol{B}\boldsymbol{S} = \det \boldsymbol{B}$, so similar matrices—see Remark 1.70—have the same trace. Therefore, we can make the following

Definition 1.74. The determinant of an endomorphism F of a finite-dimensional vector space V is the determinant of any of its representing matrices: $\det F := \det F_{\mathcal{B}}$, where $\mathcal{B}$ is any basis of V.

In particular, an endomorphism is invertible iff its determinant is different from zero.

Remark 1.75 (Discriminants). Congruent matrices may on the other hand have different determinants. We see however, from (D,2) and (D.7) that $\boldsymbol{B}' = \boldsymbol{S}^{\mathsf{T}}\boldsymbol{B}\boldsymbol{S}$ implies $\det \boldsymbol{B}' = (\det \boldsymbol{S})^2 \det \boldsymbol{B}$, so the determinant changes by a factor that is the square of a nonzero scalar. The discriminant of a matrix is by definition its determinant up to such a

[15] Note that this statment is independent of the chosen bases. In fact, any two representing matrices of F are equivalent $n \times n$ matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, as in Remark 1.69, i.e., $\boldsymbol{A} = \boldsymbol{T}\boldsymbol{B}\boldsymbol{S}$, where $\boldsymbol{S}$ and $\boldsymbol{T}$ are invertible. We then have, by (D.7), $\det \boldsymbol{A} = c \det \boldsymbol{B}$, where $c = \det \boldsymbol{S} \det \boldsymbol{T}$ is a nonzero number.

factor. It follows that congruent matrices have the same discriminant and that we may define the discriminant of a bilinear form as the discriminant of any of its representing matrices. If we work over $\mathbb{C}$, where every element is a square, the only meaningful statement we can make is whether the discriminant is equal to zero or different from zero. Over $\mathbb{R}$, we can refine this and speak of strictly positive, strictly negative or zero discriminants.

CHAPTER 2

Linear ODEs and Diagonalization of Endomorphisms

In this chapter we discuss the problem of bringing an endomorphism, via a suitable choice of basis, to a diagonal representing matrix, whenever possible. We start by motivating this problem with the study of systems of linear ordinary differential equations with constant coefficients.

2.1. Differential equations

An ordinary differential equation (ODE) is an equation whose unknown is a differentiable function of one variable that appears in the equation together with its derivatives.

Newton's equation $F = ma$ is an example of an ODE. In this case, the unknown is a path $x(t)$.[1] In normal form (i.e., with the highest derivative set in evidence on the left hand side) the equation reads

$$\ddot{x} = \frac{1}{m} F(x, \dot{x}, t).$$

A solution is a specific function $x(t)$ that satisfies the equation for all t in some open interval (possibly the whole of $\mathbb{R}$). In the one-dimensional case, this is a single equation, but if we consider the problem in three space dimensions, we get a system of ODEs—one equation for each component. We also get a system is we describe the interaction of several particles.

The order of an ODE (or of a system of ODEs) corresponds to highest derivative occurring in it. For example, Newton's equation (like many fundamental equations in physics) is a second-order ODE. It is possible to reduce the order by the following trick, which we illustrate in the case of Newton's equation. Namely, we introduce the momentum $p := mv$. We can then rewrite Newton's equation as the first-order

[1]By path we mean a (twice) differentiable map whose domain is an interval.

system

$$\begin{cases} \dot{x} = \dfrac{p}{m} \\ \dot{p} = F\left(x, \dfrac{p}{m}, t\right) \end{cases}$$

where now the pair (x, p) is regarded as the unknown.

The Cauchy problem for a system of first-order ODEs consists of the system together the specification of the variables at some initial time.

The theory of ODEs is discussed in the analysis classes, and there it is proved, under some mild conditions, that a Cauchy problem has a unique solution (in an open interval around the initial time).

We will consider here only linear (systems of) first-order ODEs with constant coefficients, where methods of linear algebra can be used.

2.1.1. Linear ODEs with constant coefficients. ODEs are called linear if the unknown and its derivatives appear linearly. The ODE is called homogenous if there is no term independent of them, inhonogeneous otherwise.

A linear first-order ODE is then the equation $\dot{x} = ax$, when homogeneous, and $\dot{x} = ax + b$, when inhomogeneous, where a and b are given functions of t. We say that the equation has constant coefficients if a is constant.

NOTATION 2.1. It is a standard practice in the theory of ODEs to write (t) after a variable to specify that it is not assumed to be constant. If (t) does not appear, the variable is assumed to be a constant. The unknown function is written without (t) in the equation, as the notation $x(t)$ is reserved for writing a solution. Therefore, a linear first-order ODE constant coefficients is written as

$$\dot{x} = ax + b(t). \tag{2.1}$$

In the homogeneous case—i.e., when $b(t) = 0$—we write

$$\dot{x} = ax. \tag{2.2}$$

EXAMPLE 2.2 (Growth processes). The homogenous equation $\dot{x} = ax$ with constant a describes a growth process where the growth $\dot{x}$ is proportional to the quantity x itself (properly speaking, we have a growth when $a > 0$ and a decay when $a < 0$). Such equation is widely used: e.g., in economics to describe capital growth by compound interest, in biology to describe growth (or decline) of a population, in physics to describe radioactive decay.

To solve (2.2), we introduce $y(t) := \mathrm{e}^{-at}x(t)$. Differentianting we get $\dot{y} = \mathrm{e}^{-at}(\dot{x} - ax)$. Therefore, x is a solution to (2.2) iff $\dot{y} = 0$, i.e., $y = c$, where c is a constant. We then get the general solution $x(t) = \mathrm{e}^{at}c$. We can also rephrase this as the

PROPOSITION 2.3. *The Cauchy problem*

$$\begin{cases} \dot{x} &= ax \\ x(0) &= x_0 \end{cases}$$

for a homogenous linear ODE with constant coefficient a has the unique solution

$$x(t) = \mathrm{e}^{at}x_0. \tag{2.3}$$

The solution is defined for all $t \in \mathbb{R}$.

By the same trick, we may also study the associated nonhomogenous equation (2.1), where $b(t)$ is not necessarily assumed to be constant. Namely, we write again $y(t) := \mathrm{e}^{-at}x(t)$. In this case, x is a solution to (2.1) iff $\dot{y} = \mathrm{e}^{-at}\,b(t)$, so we can get y, and hence x, by integrating $\mathrm{e}^{-at}b(t)$. Namely, we have $y(t) = c + \int_0^t \mathrm{e}^{-as}\,b(s)\,\mathrm{d}s$, where c is a constant. Therefore, we have the general solution

$$x(t) = \mathrm{e}^{at}c + \int_0^t \mathrm{e}^{a(t-s)}\,b(s)\,\mathrm{d}s. \tag{2.4}$$

This leads to the

PROPOSITION 2.4. *The Cauchy problem*

$$\begin{cases} \dot{x} &= ax + b(t) \\ x(0) &= x_0 \end{cases}$$

for a nonhomogenous linear ODE with constant coefficient a has the unique solution

$$x(t) = \mathrm{e}^{at}x_0 + \int_0^t \mathrm{e}^{a(t-s)}\,b(s)\,\mathrm{d}s. \tag{2.5}$$

The solution is defined for all $t \in \mathbb{R}$.

2.1.2. Systems of linear ODEs with constant coefficients. A system of the form

$$\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}(t), \tag{2.6}$$

where $\boldsymbol{A}$ is a given $n \times n$ matrix (with constant entries), called the coefficient matrix, and $\boldsymbol{b}$ is a given map from an open interval to $\mathbb{R}^n$, is called a system of n linear ODEs with constant coefficients. For $\boldsymbol{b}(t)$ the zero map, we have the system

$$\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x}, \tag{2.7}$$

which is called homogeneous.

EXAMPLE 2.5 (Harmonic oscillator). Consider Newton's equation in one dimension with force $F = -kx$, where k is a given positive constant. The second-order ODE $m\ddot{x} = -kx$ may be rewritten as the system

$$\begin{cases} \dot{x} = \dfrac{p}{m} \\ \dot{p} = -kx \end{cases}$$

which can be brought in matrix form $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x}$ by setting $\boldsymbol{x} = \left(\begin{smallmatrix} x \\ p \end{smallmatrix}\right)$ and

$$\boldsymbol{A} = \begin{pmatrix} 0 & \frac{1}{m} \\ -k & 0 \end{pmatrix}.$$

EXAMPLE 2.6 (Homogenous nth-order linear ODE with constant coefficients). Consider the ODE

$$x^{(n)} + a_1 x^{(n-1)} + \cdots + a_{n-1}\dot{x} + a_n x = 0, \tag{2.8}$$

where $a_1, \ldots, a_n$ are given constants. This ODE can be rewritten as a system by defining

$$\boldsymbol{x} := \begin{pmatrix} x \\ \dot{x} \\ \vdots \\ x^{(n-2)} \\ x^{(n-1)} \end{pmatrix}.$$

In fact, we have

$$\dot{\boldsymbol{x}} = \begin{pmatrix} \dot{x} \\ \ddot{x} \\ \vdots \\ x^{(n-1)} \\ x^{(n)} \end{pmatrix} = \begin{pmatrix} \dot{x} \\ \ddot{x} \\ \vdots \\ x^{(n-1)} \\ -a_1 x^{(n-1)} - \cdots - a_{n-1}\dot{x} - a_n x \end{pmatrix}.$$

Therefore, the ODE (2.8) is equivalent to the system (2.7) with

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & 0 & 1 \\ -a_n & -a_{n-1} & \cdots & -a_2 & -a_1 \end{pmatrix}, \tag{2.9}$$

where all the nondisplayed entries are equal to zero.

EXAMPLE 2.7 (Infinite-dimensional examples). In physics one also studies equations involving functions of several variables together with their partial derivatives—these are called partial differential equations

(PDEs). Several important PDEs in physics are linear. For example, the wave equation

$$\frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} = \Delta\psi,$$

the heat equation

$$\frac{\partial\psi}{\partial t} = \alpha\Delta\psi,$$

and the Schrödinger equation

$$\frac{\mathrm{i}}{\hbar}\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\Delta\psi + V\psi.$$

In these examples, ψ is a function (real in the first two cases and complex in the third) of the time variable t and the space variables x, y, z; $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is the Laplace operator; c, α, $\hbar$, and m are real constants (respectively: velocity, diffusivity, Planck constant, mass); and V is a real function of the space variables (the potential). Each of these equations may be viewed as a system of linear ODEs with constant coefficients in the infinite-dimensional vector space $C^2(\mathbb{R}^3)$ of twice differentiable functions in the space variables. The unknown $\psi(t; x, y, z)$ is then viewed as map $\mathbb{R} \to C^2(\mathbb{R}^3)$, $t \mapsto \psi_t$, with $\psi_t(x, y, z) := \psi(t; x, y, z)$. The techniques we present in this section for finite systems of linear ODEs with constant coefficients may be extended to these infinite-dimensional systems, but we will not do it here.

Example 2.8 (Diagonal case). The homogeneous system (2.6) may be easily solved if the coefficient matrix is diagonal. Namely, suppose $\boldsymbol{A} = \boldsymbol{D}$ with $\boldsymbol{D}$ diagonal with diagonal entries $\lambda_1, \dots, \lambda_n$:

$$\boldsymbol{D} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}. \tag{2.10}$$

The system then splits into n independent equations

$$\dot{x}^1 = \lambda_1 x^1, \dots, \dot{x}^n = \lambda_n x^n.$$

The ith equation has the general solution $x^i(t) = \mathrm{e}^{\lambda_i t} c^i$, where c^i is a constant. It then follows that the associated Cauchy problem with initial condition $\boldsymbol{x}(0) = \boldsymbol{x}_0$ has the unique solution $x^i(t) = \mathrm{e}^{\lambda_i t} x_0^i$, which is defined for all $t \in \mathbb{R}$, for $i = 1, \dots, n$. If we denote by $\mathrm{e}^{\boldsymbol{D}t}$ the diagonal matrix with diagonal entries $\mathrm{e}^{\lambda_1 t}, \dots, \mathrm{e}^{\lambda_n t}$,

$$\mathrm{e}^{\boldsymbol{D}t} = \begin{pmatrix} \mathrm{e}^{\lambda_1 t} & & \\ & \ddots & \\ & & \mathrm{e}^{\lambda_n t} \end{pmatrix}, \tag{2.11}$$

we can write the unique solution as

$$\boldsymbol{x}(t) = \mathrm{e}^{\boldsymbol{D}t}\boldsymbol{x}_0. \tag{2.12}$$

DIGRESSION 2.9 (Upper triangular case). The homogenous case when the coefficient matrix is upper triangular may also be easily solved. To illustrate the idea, we consider the two-dimensional case with $\boldsymbol{A} = \begin{pmatrix} \lambda_1 & \beta \\ 0 & \lambda_2 \end{pmatrix}$. We then have the two equations

$$\dot{x}^1 = \lambda_1 x^1 + \beta x^2, \quad \dot{x}^2 = \lambda_2 x^2.$$

The second equation is independent of x^1 and has the general solution $x^2(t) = \mathrm{e}^{\lambda_2 t}c^2$, where c^2 is a constant. We may plug this solution into the first equation getting the nonhomogenous equation

$$\dot{x}^1 = \lambda_1 x^1 + \beta\, \mathrm{e}^{\lambda_2 t}c^2,$$

which can be solved using (2.4) with $b(t) = \mathrm{e}^{\lambda_2 t}\beta c^2$. If $\lambda_1 \neq \lambda_2$, we then get

$$x^1(t) = \mathrm{e}^{\lambda_1 t}c^1 + \frac{\mathrm{e}^{\lambda_2 t}}{\lambda_2 - \lambda_1}\beta c^2,$$

where c^1 is a new constant. If $\lambda_1 = \lambda_2 = \lambda$, we get instead

$$x^1(t) = \mathrm{e}^{\lambda t}c^1 + t\,\mathrm{e}^{\lambda t}\beta c^2.$$

Note that in the "degenerate case" when $\lambda_1 = \lambda_2$, the solution does not only depend on exponential functions but also has a factor t. The general case, with $\boldsymbol{A}$ upper triangular, is solved similarly. One solves the equations iteratively from the last equation, which only depends on the last component of $\boldsymbol{x}$, to the first. Every time one inserts the solution into the previous equation, which then turns out to be a nonhomogeneous linear ODE that can be solved by (2.4). By induction one sees that the inhomogenous term $b(t)$ is a linear combination of products of exponential and polynomials. In conclusion, the general solution will also be given by a linear combination of products of exponentials and polynomials. If the diagonal entries are all different, the general solution is simply a linear combination of exponentials.

2.1.3. The matrix exponential. Following the examples of the solutions (2.3) and (2.12), we now want to get a general solution to (2.7) in the form of an exponential.

The exponential of a square matrix $\boldsymbol{A}$ is defined by extending to matrices the usual series defining the exponential of a real or complex number:

$$\mathrm{e}^{\boldsymbol{A}} := \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{A}^k = \mathbf{1} + \boldsymbol{A} + \frac{1}{2}\boldsymbol{A}^2 + \cdots.$$

One can easily see that the series converges for any matrix $\boldsymbol{A}$ and, moreover, that the power series

$$\mathrm{e}^{\boldsymbol{A}t} = \sum_{k=0}^{\infty} \frac{t^k}{k!}\boldsymbol{A}^k = \mathbf{1} + t\boldsymbol{A} + \frac{t^2}{2}\boldsymbol{A}^2 + \cdots .$$

has infinite radius of convergence. It follows that it defines a smooth (i.e., infinitely often continuously differentiable) function and that taking a derivative commutes with the sum, so we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{\boldsymbol{A}t} = \mathrm{e}^{\boldsymbol{A}t}\boldsymbol{A} = \boldsymbol{A}\mathrm{e}^{\boldsymbol{A}t}. \tag{2.13}$$

Moreover, from

$$\mathrm{e}^{\boldsymbol{A}0} = \mathbf{1},$$

we see that $\boldsymbol{U}(t) := \mathrm{e}^{\boldsymbol{A}t}$ is the unique solution to the matrix Cauchy problem

$$\begin{cases} \dot{\boldsymbol{U}} & = & \boldsymbol{A}\boldsymbol{U} \\ \boldsymbol{U}(0) & = & \mathbf{1} \end{cases}$$

Remark 2.10. The matrix exponential has the following properties, which can be easily proved:

(1) As in the case of the exponential of a number,

$$\mathrm{e}^{\boldsymbol{A}(t+s)} = \mathrm{e}^{\boldsymbol{A}t}\,\mathrm{e}^{\boldsymbol{A}s} \tag{2.14}$$

for all $t, s \in \mathbb{C}$. In particular, taking $s = -t$, we see that

$$\mathbf{1} = \mathrm{e}^{\boldsymbol{A}t}\mathrm{e}^{-\boldsymbol{A}t},$$

so $\mathrm{e}^{\boldsymbol{A}t}$ is always invertible and its inverse is $\mathrm{e}^{-\boldsymbol{A}t}$.

(2) If $\boldsymbol{A}$ and $\boldsymbol{B}$ commute (i.e., $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A}$), then in the product of two exponentials we can rearrange the factors. Therefore, we have

$$\mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}} = \mathrm{e}^{\boldsymbol{A}}\,\mathrm{e}^{\boldsymbol{B}} \quad \text{if } \boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A}. \tag{2.15}$$

Note that this equation does not hold if $\boldsymbol{A}$ and $\boldsymbol{B}$ do not commute because on the right hand side all powers of $\boldsymbol{A}$ comes to the left and all powers of $\boldsymbol{B}$ to the right, whereas on the left hand side powers of $\boldsymbol{A}$ and $\boldsymbol{B}$ come in all possible orders.

(3) For every invertible matrix $\boldsymbol{S}$, we have

$$\mathrm{e}^{\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}} = \boldsymbol{S}^{-1}\mathrm{e}^{\boldsymbol{A}}\boldsymbol{S}. \tag{2.16}$$

This follows from $(\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S})^k = \boldsymbol{S}^{-1}\boldsymbol{A}^k\boldsymbol{S}$, which is easily proved for all k.

One more interesting property of the matrix exponential is given by the following

PROPOSITION 2.11. *Let $\boldsymbol{A}$ be a square matrix. Then*

$$\det \mathrm{e}^{\boldsymbol{A}t} = \mathrm{e}^{t \operatorname{tr} \boldsymbol{A}}$$

for every t.

PROOF. We consider the function $d(t) := \det \mathrm{e}^{\boldsymbol{A}t}$ and compute its derivative. Using (2.14) and the multiplicativity of the determinant—i.e, property (D.7) on page 29—we have $d(t+h) = d(t)d(h)$. Therefore, using $d(0) = 1$, we have

$$\dot{d}(t) = \lim_{h\to 0} \frac{d(t+h) - d(t)}{h} = \lim_{h\to 0} \frac{d(h) - 1}{h} d(t) = a d(t)$$

with $a := \dot{d}(0)$. By Proposition 2.3, we then have $d(t) = \mathrm{e}^{at}$. To complete the proof, we only have to show that $a = \operatorname{tr} \boldsymbol{A}$. This can be done explicitly by using the Leibniz formula (1.7):

$$\det \mathrm{e}^{\boldsymbol{A}h} = \det(\mathbf{1} + h\boldsymbol{A} + O(h^2)) = 1 + h \operatorname{tr} \boldsymbol{A} + O(h^2).$$

□

We may use the matrix exponential to solve any homogeneous linear system of ODEs with constant coefficients $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x}$ by the same trick we used in the case of a single equation. Namely, we introduce $\boldsymbol{y} := \mathrm{e}^{-\boldsymbol{A}t}\boldsymbol{x}$. Differentianting, thanks to (2.13), we get $\dot{\boldsymbol{y}} = \mathrm{e}^{-\boldsymbol{A}t}(\dot{\boldsymbol{x}} - \boldsymbol{A}\boldsymbol{x})$. Therefore, $\boldsymbol{x}$ is a solution to (2.7) iff $\dot{\boldsymbol{y}} = 0$, i.e., $\boldsymbol{y} = \boldsymbol{c}$, where $\boldsymbol{c}$ is a constant vector. We then get the general solution $\boldsymbol{x}(t) = \mathrm{e}^{\boldsymbol{A}t}\boldsymbol{c}$. We can also rephrase this as the

PROPOSITION 2.12. *The Cauchy problem*

$$\begin{cases} \dot{\boldsymbol{x}} & = & \boldsymbol{A}\boldsymbol{x} \\ \boldsymbol{x}(0) & = & \boldsymbol{x}_0 \end{cases}$$

for a homogenous linear system of ODEs with constant coefficients has the unique solution

$$\boldsymbol{x}(t) = \mathrm{e}^{\boldsymbol{A}t}\boldsymbol{x}_0. \tag{2.17}$$

The solution is defined for all $t \in \mathbb{R}$.

By the same trick, we may also study the associated nonhomogenous equation (2.6), where $\boldsymbol{b}(t)$ is a (not necessarily constant) map from an interval to $\mathbb{R}^n$. Namely, we write again $\boldsymbol{y} := \mathrm{e}^{-\boldsymbol{A}t}\boldsymbol{x}$. In this case, $\boldsymbol{x}$ is a solution to (2.6) iff $\dot{\boldsymbol{y}} = \mathrm{e}^{-\boldsymbol{A}t}\,\boldsymbol{b}(t)$, so we can get $\boldsymbol{y}$, and hence $\boldsymbol{x}$, by integrating $\mathrm{e}^{-\boldsymbol{A}t}\,\boldsymbol{b}(t)$. Namely, we have $\boldsymbol{y}(t) = \boldsymbol{c} + \int_0^t \mathrm{e}^{-\boldsymbol{A}s}\,\boldsymbol{b}(s)\,\mathrm{d}s$, where $\boldsymbol{c}$ is a constant vector, and the integral is computed componentwise. Therefore, we have the general solution

$$\boldsymbol{x}(t) = \mathrm{e}^{\boldsymbol{A}t}\boldsymbol{c} + \int_0^t \mathrm{e}^{\boldsymbol{A}(t-s)}\,\boldsymbol{b}(s)\,\mathrm{d}s. \tag{2.18}$$

2.1.4. Computing the matrix exponential. The practical problem consists in computing the exponential of a given matrix $\boldsymbol{A}$. The simplest case is when the matrix $\boldsymbol{A}$ is diagonal, $\boldsymbol{A} = \boldsymbol{D}$ with $\boldsymbol{D}$ as in (2.10). In fact, we have

$$\boldsymbol{D}^k = \begin{pmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{pmatrix},$$

so $\mathrm{e}^{\boldsymbol{D}t}$ is as in (2.11). This way we recover the solution discussed in Example 2.8.

Thanks to (2.13), we can also explictly compute the exponential of a diagonalizable matrix, i.e., a square matrix $\boldsymbol{A}$ that is similar to a diagonal matrix $\boldsymbol{D}$. Writing

$$\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{D}$$

for some invertible matrix $\boldsymbol{S}$, we get the solution to the associated Cauchy problem in the form

$$\boldsymbol{x}(t) = \boldsymbol{S}^{-1} \begin{pmatrix} \mathrm{e}^{\lambda_1 t} & & \\ & \ddots & \\ & & \mathrm{e}^{\lambda_n t} \end{pmatrix} \boldsymbol{S}\boldsymbol{x}_0,$$

where $\lambda_1, \dots, \lambda_n$ are the diagonal elements of $\boldsymbol{D}$. In the next sections we will develop methods to determine the scalars $\lambda_1, \dots, \lambda_n$ and the matrix $\boldsymbol{S}$, whenever possible.

REMARK 2.13. Not every matrix is diagonalizable. Consider, e.g., $\boldsymbol{A} = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. Pick an invertible matrix $\boldsymbol{S} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Computing $\boldsymbol{S}^{-1}$ as in (1.8), we then get

$$\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \frac{1}{ad - bc} \begin{pmatrix} dc & d^2 \\ -c^2 & -cd \end{pmatrix}.$$

Since $\boldsymbol{S}$ is invertible, the entries c and d cannot be both zero, so the right and side cannot be a diagonal matrix.

DIGRESSION 2.14. By Digression 2.9, and by (2.13), we can also explicitly compute the exponential of a matrix that is similar to an upper triangular matrix. This turns out to be always possible if we work over complex numbers. (See Section 2.4.)

EXAMPLE 2.15. Consider again the nondiagonalizable matrix $\boldsymbol{A} = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. We can compute its exponential explicitly as follows. We easily see that $\boldsymbol{A}^2 = \boldsymbol{0}$, which in turn implies $\boldsymbol{A}^n = 0$ for all $n > 1$. Therefore,

$$\mathrm{e}^{\boldsymbol{A}t} = \boldsymbol{1} + t\boldsymbol{A} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

This exponential actually solves the problem of the free particle

$$m\ddot{x} = 0,$$

which is the same as the harmonic oscillator of Example 2.5 but with $k = 0$. The coefficient matrix is in this case $\boldsymbol{A} = \begin{pmatrix} 0 & \frac{1}{m} \\ 0 & 0 \end{pmatrix}$, i.e., $\frac{1}{m}$ times the matrix considered above. In this case, we have

$$\mathrm{e}^{\boldsymbol{A}t} = \boldsymbol{1} + t\boldsymbol{A} = \begin{pmatrix} 1 & \frac{t}{m} \\ 0 & 1 \end{pmatrix}.$$

The solution to the Cauchy problem is then

$$\begin{pmatrix} x(t) \\ p(t) \end{pmatrix} = \mathrm{e}^{\boldsymbol{A}t} \begin{pmatrix} x_0 \\ p_0 \end{pmatrix} = \begin{pmatrix} 1 & \frac{t}{m} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ p_0 \end{pmatrix} = \begin{pmatrix} x_0 + \frac{t}{m} p_0 \\ p_0 \end{pmatrix}.$$

This yields the usual formula

$$x(t) = x_0 + \frac{p_0}{m} t.$$

2.2. Diagonalization of matrices

Suppose $\boldsymbol{A}$ is diagonalizable, i.e., $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{D}$ for some invertible matrix $\boldsymbol{S}$ and a diagonal matrix $\boldsymbol{D}$. If λ_i is the ith diagonal entry of $\boldsymbol{D}$, then we have $\boldsymbol{D}\boldsymbol{e}_i = \lambda_i \boldsymbol{e}_i$. Denoting by $\boldsymbol{v}_i$ the ith column of $\boldsymbol{S}$, i.e., $\boldsymbol{v}_i = \boldsymbol{S}\boldsymbol{e}_i$, we get

$$\boldsymbol{A}\boldsymbol{v}_i = \boldsymbol{A}\boldsymbol{S}\boldsymbol{e}_i = \boldsymbol{S}\boldsymbol{D}\boldsymbol{e}_i = \lambda_i \boldsymbol{S}\boldsymbol{e}_i = \lambda_i \boldsymbol{v}_i.$$

This motivates the following

Definition 2.16 (Eigenvectors and eigenvalues). A <u>nonzero</u> vector $\boldsymbol{v}$ is called an eigenvector of a square matrix $\boldsymbol{A}$ if there is a scalar λ, called the eigenvalue to the eigenvector $\boldsymbol{v}$, such that

$$\boldsymbol{A}\boldsymbol{v} = \lambda \boldsymbol{v}. \tag{2.19}$$

We then have the

Theorem 2.17. *A square matrix is diagonalizable iff it admits a basis of eigenvectors.*

Proof. If $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{D}$ is diagonal, then, by the above discussion, the vectors $\boldsymbol{v}_i = \boldsymbol{S}\boldsymbol{e}_i$ are eigenvectors. They are a basis because $\boldsymbol{S}$ is invertible. (Note that $\boldsymbol{S}$ defines the change of basis from the standard basis $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ to the basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$.)

On the other hand, if $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ is a basis of eigenvectors, we define $\boldsymbol{S}$ as the matrix whose ith column is $\boldsymbol{v}_i$. It is invertible by property (D.9) on page 29. We then have

$$\boldsymbol{D}\boldsymbol{e}_i = \boldsymbol{D}\boldsymbol{S}^{-1}\boldsymbol{v}_i = \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{v}_i = \lambda_i \boldsymbol{S}^{-1}\boldsymbol{v}_i = \lambda_i \boldsymbol{e}_i,$$

which shows that $\boldsymbol{D}$ is diagonal with diagonal entries $\lambda_1, \dots, \lambda_n$. □

REMARK 2.18 (Linear systems of ODEs). Back to our problem $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x}$, assuming we have a basis of eigenvectors $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ of $\boldsymbol{A}$, we have the expansion $\boldsymbol{x}(t) = \xi^i(t)\boldsymbol{v}_i$, with uniquely determined scalars $\xi^1(t), \dots, \xi^n(t)$ for each t. The system then becomes $\dot{\xi}^1 = \lambda_1\xi^1, \dots, \dot{\xi}^n = \lambda_n\xi^n$, which is solved by $\xi^i(t) = \mathrm{e}^{\lambda_i t}\xi_0^i$, with $(\xi_0^1, \dots, \xi_0^n)$ the components of the expansion of $\boldsymbol{x}_0$: $\boldsymbol{x}_0 = \xi_0^i\boldsymbol{v}_i$. Therefore, we get the unique solution to the Cauchy problem in the form

$$\boldsymbol{x}(t) = \sum_{i=1}^{n} \mathrm{e}^{\lambda_i t}\xi_0^i\boldsymbol{v}_i. \tag{2.20}$$

REMARK 2.19 (Choice of field). For our application to linear systems of ODE we assume $\boldsymbol{A}$ to have real (or complex) entries. On the other hand, the general problem of diagonalization, Theorem 2.17, and the rest of the discussion make sense for every ground field.

Now note that we can rewrite the eigenvector equation (2.19) as $(\boldsymbol{A} - \lambda\mathbf{1})\boldsymbol{v} = 0$. This shows the following

LEMMA 2.20. *A scalar λ is an eigenvalue of the square matrix $\boldsymbol{A}$ iff $\boldsymbol{A} - \lambda\mathbf{1}$ is <u>not</u> invertible.*

PROOF. $\boldsymbol{A}-\lambda\mathbf{1}$ is invertible iff its kernel is different from zero. This happens iff there is a nonzero vector $\boldsymbol{v}$ such that $(\boldsymbol{A} - \lambda\mathbf{1})\boldsymbol{v} = 0$. □

It follows, from property (D.8) on page 29, that the eigenvalues of $\boldsymbol{A}$ are precisely the solutions to $\det(\boldsymbol{A} - \lambda\mathbf{1}) = 0$. This motivates considering the function

$$P_{\boldsymbol{A}} := \det(\boldsymbol{A} - \lambda\mathbf{1}).$$

If $\boldsymbol{A}$ is an $n \times n$ matrix, by the Leibniz formula (1.7) we see that $P_{\boldsymbol{A}}$ is a polynomial of degree n,

$$P_A = b_0\lambda^n + b_1\lambda^{n-1} + \cdots + b_n.$$

In particular, $b_0 = (-1)^n$, $b_1 = (-1)^{n-1}\operatorname{tr}\boldsymbol{A}$, and $b_n = \det\boldsymbol{A}$.

DEFINITION 2.21 (Characteristic polynomial). The polynomial $P_{\boldsymbol{A}}$ is called the characteristic polynomial of the square matrix $\boldsymbol{A}$.

We may summarize the previous discussion as the

PROPOSITION 2.22. *The eigenvalues of a square matrix are the roots of its characteristic polynomial.*

In the quest for the diagonalization of $\boldsymbol{A}$, we first compute its eigenvalues by this proposition. Next, we proceed to the determination of its eigenvectors. That is, for each root λ of the characteristic polynomial of $\boldsymbol{A}$, we consider the system $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$ of n linear equations (with unknown the components $v^1, \dots, v^n$ of the vector $\boldsymbol{v}$). We know that this system has nontrivial solutions because $\boldsymbol{A} - \lambda\mathbf{1}$ is not invertible.

Note that for every eigenvalue there are infinitely many eigenvectors. For example, if $\boldsymbol{v}$ is a λ-eigenvalue, then so is $a\boldsymbol{v}$ for every scalar $a \neq 0$. More generally, if $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are λ-eigenvalues, then so is any nonzero linear combination of them.

In order to diagonalize a matrix, we have to determine all its eigenvalues, but there is no need to find all the corresponding eigenvectors: it is enough to find a basis of eigenvectors (if possible).

EXAMPLE 2.23. Let $\boldsymbol{A} = \left(\begin{smallmatrix} 0 & b \\ c & 0 \end{smallmatrix}\right)$. Its characteristic polynomial is $P_{\boldsymbol{A}} = \det\left(\begin{smallmatrix} -\lambda & b \\ c & -\lambda \end{smallmatrix}\right) = \lambda^2 - bc$. Assuming $bc > 0$, we have the two real distinct roots $\lambda_\pm = \pm\sqrt{bc}$. The eigenvector equation $\boldsymbol{A}\boldsymbol{v} = \lambda_+\boldsymbol{v}$ is then the system

$$\begin{cases} bv^2 = \sqrt{bc}\,v^1 \\ cv^1 = \sqrt{bc}\,v^2 \end{cases}$$

The first equation yields the relation $v^2 = \sqrt{\frac{c}{b}}\,v^1$. The second equation does not yield any new independent condition (this is a consequence of $\boldsymbol{A} - \lambda_+\mathbf{1}$ being not invertible). Therefore, we have a 1-parameter family of solutions (we can choose $v^1 \in \mathbb{R}$ as the parameter). For example, for $v^1 = 1$ we have the eigenvector $\boldsymbol{v}_+ = \left(\begin{smallmatrix} 1 \\ \sqrt{\frac{c}{b}} \end{smallmatrix}\right)$. A similar computation yields the eigenvector $\boldsymbol{v}_- = \left(\begin{smallmatrix} 1 \\ -\sqrt{\frac{c}{b}} \end{smallmatrix}\right)$ to the eigenvalue λ_-.[2] One can easily check that $(\boldsymbol{v}_+, \boldsymbol{v}_-)$ is a basis of $\mathbb{R}^2$. The transformation matrix and its inverse are then

$$\boldsymbol{S} = (\boldsymbol{v}_+\ \boldsymbol{v}_-) = \begin{pmatrix} 1 & 1 \\ \sqrt{\frac{c}{b}} & -\sqrt{\frac{c}{b}} \end{pmatrix}, \quad \boldsymbol{S}^{-1} = \frac{1}{2}\begin{pmatrix} 1 & \sqrt{\frac{b}{c}} \\ 1 & -\sqrt{\frac{b}{c}} \end{pmatrix}.$$

One can then explicitly verify that $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \left(\begin{smallmatrix} \sqrt{bc} & 0 \\ 0 & -\sqrt{bc} \end{smallmatrix}\right)$.

[2] If b and c are both positive, we could also pick $v^1 = \sqrt{b}$, getting $v^2 = \sqrt{c}$ and $\boldsymbol{v}_\pm = \left(\begin{smallmatrix} \sqrt{b} \\ \pm\sqrt{c} \end{smallmatrix}\right)$.

EXAMPLE 2.24. Consider the matrix $\boldsymbol{A}$ of (2.9) associated to a homogenous nth-order linear ODE with constant coefficients as in Example 2.6. It is a good exercise to show that in this case

$$P_{\boldsymbol{A}} = (-1)^n(\lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n).$$

Therefore, the characteristic equation $P_{\boldsymbol{A}} = 0$ may be obtained from (2.8) by formally substituting λ^k to $x^{(k)}$ for $k = 0, \ldots, n$. Equivalently, it may be obtained by inserting into (2.8) the ansatz $x(t) = \mathrm{e}^{\lambda t}$.

To be sure that the roots of the characteristic polynomial exist, we assume from now that we work over $\mathbb{C}$ and make use of the

THEOREM 2.25 (Fundamental theorem of algebra). *A nonconstant complex polynomial has a root. As a consequence, it splits into a product of linear factors.*

The characteristic polynomial $P_{\boldsymbol{A}}$ of a complex $n \times n$ matrix $\boldsymbol{A}$ factorizes as

$$P_{\boldsymbol{A}} = (-1)^n (\lambda - \lambda_1)^{s_1} \cdots (\lambda - \lambda_k)^{s_k},$$

where $\lambda_1, \ldots, \lambda_k$ are the pairwise distinct roots of $P_{\boldsymbol{A}}$. Note that $k \leq n$. The exponent s_i is called the algebraic multiplicity of λ_i. Note that we have $s_1 + \cdots + s_k = n$.

REMARK 2.26 (Linear systems of ODEs). The coefficient matrix $\boldsymbol{A}$ of a linear system of ODEs is usually assumed to be real, and we are interested in a real solution $\boldsymbol{x}(t)$. The trick is to regard $\boldsymbol{A}$ as a complex matrix; assuming it then to be diagonalizable, we may find a basis of eigenvectors and proceed as in Remark 2.18. The unique solution to the Cauchy problem is still given by (2.20), i.e.,

$$\boldsymbol{x}(t) = \sum_{i=1}^{n} \mathrm{e}^{\lambda_i t} \xi_0^i \boldsymbol{v}_i,$$

where now the λ_is, ξ_0^is, and $\boldsymbol{v}_i$s may be complex. If the initial condition $\boldsymbol{x}_0$ is real, then the unique solution is also real, which ensures that the sum of complex vectors on the right hand side of (2.20) yields a real vector. It is always possible to rearrange this sum of complex exponentials times complex vectors into a real sum involving real exponentials and trigonometric functions as we explain in the next example and, more generally, in Section 2.2.1.

EXAMPLE 2.27 (Harmonic oscillator). In the Example 2.5 of the harmonic oscillator, we have $\boldsymbol{A} = \begin{pmatrix} 0 & \frac{1}{m} \\ -k & 0 \end{pmatrix}$. The characteristic polynomial is

$$P_{\boldsymbol{A}} = \det(\boldsymbol{A} - \lambda \mathbf{1}) = \begin{pmatrix} -\lambda & \frac{1}{m} \\ -k & -\lambda \end{pmatrix} = \lambda^2 + \omega^2$$

with $\omega := \sqrt{\frac{k}{m}}$. The complex eigenvalues are then $\pm\mathrm{i}\omega$. To find the eigenvectors, we then study first the equation $\boldsymbol{A}\boldsymbol{v} = \mathrm{i}\omega\boldsymbol{v}$. Writing $\boldsymbol{v} = \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$, we get the equation

$$\frac{b}{m} = \mathrm{i}\omega a$$

(with the second equation in the system a multiple of this one). By choosing $a = 1$, we get the eigenvector $\boldsymbol{v} = \left(\begin{smallmatrix} 1 \\ \mathrm{i}m\omega \end{smallmatrix}\right)$. Similarly, one sees that $\bar{\boldsymbol{v}} = \left(\begin{smallmatrix} 1 \\ -\mathrm{i}m\omega \end{smallmatrix}\right)$ is an eigenvector for $-\mathrm{i}\omega$. (This is a general fact: if $\boldsymbol{v}$ is an eigenvector with eigenvalue λ for a <u>real</u> matrix $\boldsymbol{A}$, then $\bar{\boldsymbol{v}}$ is an eigenvector for $\bar{\lambda}$; this follows from taking the complex conjugation $\boldsymbol{A}\bar{\boldsymbol{v}} = \bar{\lambda}\bar{\boldsymbol{v}}$ of the equation $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$. See also Remark 2.36.) Since $(\boldsymbol{v}, \bar{\boldsymbol{v}})$ is a basis of $\mathbb{C}^2$, the matrix $\boldsymbol{A}$ is diagonalizable. The general <u>real</u> solution $\boldsymbol{x}$ of the associated linear system of ODEs has then the form

$$\boldsymbol{x}(t) = z\mathrm{e}^{\mathrm{i}\omega t}\begin{pmatrix} 1 \\ \mathrm{i}m\omega \end{pmatrix} + \bar{z}\mathrm{e}^{-\mathrm{i}\omega t}\begin{pmatrix} 1 \\ -\mathrm{i}m\omega \end{pmatrix}$$

for some complex constant z. In particular, the first component is

$$x(t) = z\mathrm{e}^{\mathrm{i}\omega t} + \bar{z}\mathrm{e}^{-\mathrm{i}\omega t} = A\cos(\omega t + \alpha)$$

if we write $z = \frac{A}{2}\mathrm{e}^{\mathrm{i}\alpha}$.

Remark 2.28. Note that a nonzero vector cannot be the eigenvector of two different eigenvalues. In fact, if $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$ and $\boldsymbol{A}\boldsymbol{v} = \mu\boldsymbol{v}$, we get by taking the difference that $(\lambda - \mu)\boldsymbol{v} = 0$. If $\lambda \neq \mu$, we then get $\boldsymbol{v} = 0$.

Another important observation is the following

Lemma 2.29. *A collection $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)$ of eigenvectors of an $n \times n$ matrix $\boldsymbol{A}$ corresponding to pairwise distinct eigenvalues is linearly independent.*

Proof. Suppose $\sum_{i=1}^m \alpha_i\boldsymbol{v}_i = 0$. Pick $k \in \{1, \ldots, m\}$. If we apply $\prod_{j\neq k}(\boldsymbol{A} - \lambda_j\mathbf{1})$ to the sum, all the terms with $i \neq k$ are killed. On the other hand,

$$\prod_{j\neq k}(\boldsymbol{A} - \lambda_j\mathbf{1})\boldsymbol{v}_k = \prod_{j\neq k}(\lambda_k - \lambda_j)\boldsymbol{v}_k.$$

Therefore, $\alpha_k\prod_{j\neq k}(\lambda_k - \lambda_j)\boldsymbol{v}_k = 0$. Since $\boldsymbol{v}_k \neq 0$ and $\lambda_j \neq \lambda_k$ for all $j \neq k$, we get $\alpha_k = 0$. Repeating this argument for each $k \in \{1, \ldots, m\}$, we see that every α_k has to vanish. □

This immediately implies the following useful criterion.

PROPOSITION 2.30. *If the $n \times n$ matrix $\boldsymbol{A}$ has n pairwise distinct eigenvalues, then it is diagonalizable.*

PROOF. Let $\lambda_1, \ldots, \lambda_n$ be the pairwise distinct eigenvalues of $\boldsymbol{A}$. Choose an eigenvector $\boldsymbol{v}_i$ for each eigenvalue λ_i. By Lemma 2.29 this is a basis. □

To study the diagonalization procedure in general, we need the following

DEFINITION 2.31. Let λ be an eigenvalue of $\boldsymbol{A}$. The space

$$\operatorname{Eig}(\boldsymbol{A}, \lambda) := \ker(\boldsymbol{A} - \lambda \mathbf{1})$$

is called the eigenspace of $\boldsymbol{A}$ associated with λ.[3] The dimension

$$d := \dim \operatorname{Eig}(\boldsymbol{A}, \lambda)$$

is called the geometric multiplicity of λ.

REMARK 2.32. One can show that $d \leq s$ for every eigenvalue, where s is the algebraic multiplicity and d is the geometric multiplicity. (We will prove this as Corollary 2.52 in Section 2.4.1.)

We have the following generalization of Proposition 2.30.

THEOREM 2.33. *Let $\lambda_1, \ldots, \lambda_k$ be the pairwise distinct eigenvalues of the $n \times n$ matrix $\boldsymbol{A}$ and let d_i denote the geometric multiplicity of λ_i. Then $\boldsymbol{A}$ is diagonalizable iff $d_1 + \cdots + d_k = n$. In this case, we have*

$$\mathbb{K}^n = \operatorname{Eig}(\boldsymbol{A}, \lambda_1) \oplus \cdots \oplus \operatorname{Eig}(\boldsymbol{A}, \lambda_k).$$

PROOF. We use the criterion of Remark 1.22 to show that the sum $\operatorname{Eig}(\boldsymbol{A}, \lambda_1) + \cdots + \operatorname{Eig}(\boldsymbol{A}, \lambda_k)$ is direct.

Suppose we have $\boldsymbol{v}_1 + \cdots + \boldsymbol{v}_k = 0$, $\boldsymbol{v}_i \in \operatorname{Eig}(\boldsymbol{A}, \lambda_i)$. If some of the $\boldsymbol{v}_i$s were different from zero, then we would have a zero linear combination of eigenvectors corresponding to pairwise distinct eigenvalues, which is in contradiction with Lemma 2.29. Therefore, the zero vector has a unique decomposition and the sum is direct.

By (1.2), the direct sum $\operatorname{Eig}(\boldsymbol{A}, \lambda_1) \oplus \cdots \oplus \operatorname{Eig}(\boldsymbol{A}, \lambda_k)$ has dimension $d_1 + \cdots + d_k$. By Proposition 1.49, it is then the whole space $\mathbb{K}^n$ iff $d_1 + \cdots + d_k = n$.

In this case, for $i = 1, \ldots, k$ let $(\boldsymbol{v}_{i,j})_{j=1,\ldots,d_i}$ be a basis of $\operatorname{Eig}(\boldsymbol{A}, \lambda_i)$. By Remark 1.51, the union of these bases is a basis of the whole space. Since every $\boldsymbol{v}_{i,j}$ is an eigenvector, $\boldsymbol{A}$ is then diagonalizable by Theorem 2.17. □

[3] In principle we may define $\operatorname{Eig}(\boldsymbol{A}, \lambda)$ for any scalar λ. However, if λ is not an eigenvalue, we have $\operatorname{Eig}(\boldsymbol{A}, \lambda) = 0$.

Digression 2.34. By Remark 2.32, one then also has that $\boldsymbol{A}$ is diagonalizable iff the geometric multiplicity of every eigenvalue is equal to its algebraic multiplicity.

The procedure for the diagonalization of a square matrix $\boldsymbol{A}$ is then the following:

Step 1. Find all the pairwise distinct roots $\lambda_1, \dots, \lambda_k$ of the characteristic polynomial $P_{\boldsymbol{A}}$.
Step 2. For every root λ_i choose a basis of $\operatorname{Eig}(\boldsymbol{A}, \lambda_i)$ and use it compute the dimension d_i.
Step 3. If $d_1 + \cdots + d_k = n$, then we have found a basis of eigenvectors and $\boldsymbol{A}$ is diagonalizable.

Remark 2.35. We have seen in Remark 2.13 the example $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ of a nondiagonalizable matrix. Let us see what goes wrong with the diagonalization procedure. The characteristic polynomial is $P_{\boldsymbol{A}} = \det \boldsymbol{A} = \begin{pmatrix} -\lambda & 1 \\ 0 & -\lambda \end{pmatrix} = \lambda^2$. We therefore only have the eigenvalue $\lambda = 0$, which comes with algebraic multiplicity 2. An eigenvector $\begin{pmatrix} a \\ b \end{pmatrix}$ must then satisfy $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = 0$, i.e., $b = 0$. The eigenspace of $\lambda = 0$ is then the span of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, which shows that the geometric multiplicity is 1. In particular, $\operatorname{Eig}\left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, 0\right) \subsetneq \mathbb{C}^2$.

2.2.1. Digression: The real case. Suppose now that $\boldsymbol{A}$ is a real $n \times n$ matrix—e.g., the coefficient matrix of a real system of ODEs. To proceed, we regard it as a complex matrix. Its eigenvalues may then be complex numbers.

Remark 2.36. Since $P_{\boldsymbol{A}}$ is in this case a real polynomial, any complex root comes with its complex conjugate root. Therefore, if λ is not real, we also have a distinct eigenvalue $\bar{\lambda}$. Suppose that $\boldsymbol{z} \in \mathbb{C}^n$ is an eigenvector to the eigenvalue λ, i.e., $\boldsymbol{A}\boldsymbol{z} = \lambda \boldsymbol{z}$. By taking complex conjugation, we get $\boldsymbol{A}\bar{\boldsymbol{z}} = \bar{\lambda}\bar{\boldsymbol{z}}$, so $\bar{\boldsymbol{z}}$ is an eigenvector to the eigenvalue $\bar{\lambda}$.

Now suppose that $\boldsymbol{A}$ is diagonalizable as a complex matrix. We want to show that one can find a convenient basis of real vectors associated to the complex eigenvectors so that we can bring $\boldsymbol{A}$ in a convenient normal form.

To proceed let us introduce the following notation. We denote by $\operatorname{Eig}_{\mathbb{C}}(\boldsymbol{A}, \lambda) \subseteq \mathbb{C}^n$ the eigenspace to the eigenvalue λ of $\boldsymbol{A}$ as a complex matrix. If λ is real, we denote by $\operatorname{Eig}_{\mathbb{R}}(\boldsymbol{A}, \lambda) \subseteq \mathbb{R}^n$ the eigenspace to the eigenvalue λ of $\boldsymbol{A}$ as a real matrix.

If λ is a real eigenvalue, one can show that $\dim_{\mathbb{R}} \operatorname{Eig}_{\mathbb{R}}(\boldsymbol{A}, \lambda) = \dim_{\mathbb{C}} \operatorname{Eig}_{\mathbb{C}}(\boldsymbol{A}, \lambda)$.

If λ is not real, then we write $\lambda = \alpha + \mathrm{i}\beta$, with α and β its real and imaginary parts. Let $\boldsymbol{v}$ be a λ-eigenvector. Then, by Remark 2.36, $\bar{\boldsymbol{v}}$ is a $\bar{\lambda}$-eigenvector. As $\boldsymbol{v}$ and $\bar{\boldsymbol{v}}$ belong to different eigenspaces, they are linearly independent. Now let $\boldsymbol{u}$ and $\boldsymbol{w}$ be the real and imaginary parts of $\boldsymbol{v}$, i.e., $\boldsymbol{v} = \boldsymbol{u} + \mathrm{i}\boldsymbol{w}$ with $\boldsymbol{u}$ and $\boldsymbol{w}$ real vectors. We also have $\bar{\boldsymbol{v}} = \boldsymbol{u} - \mathrm{i}\boldsymbol{w}$. Note that $(\boldsymbol{u}, \boldsymbol{w})$ is a basis of $\mathrm{Span}_{\mathbb{C}}\{\boldsymbol{v}, \bar{\boldsymbol{v}}\}$. Moreover, we have

$$\begin{aligned} \boldsymbol{A}\boldsymbol{u} &= \boldsymbol{A}\frac{\boldsymbol{v}+\bar{\boldsymbol{v}}}{2} = \frac{\lambda\boldsymbol{v}+\bar{\lambda}\bar{\boldsymbol{v}}}{2} = \alpha\boldsymbol{u} - \beta\boldsymbol{w},\\ \boldsymbol{A}\boldsymbol{w} &= \boldsymbol{A}\frac{\boldsymbol{v}-\bar{\boldsymbol{v}}}{2\mathrm{i}} = \frac{\lambda\boldsymbol{v}-\bar{\lambda}\bar{\boldsymbol{v}}}{2\mathrm{i}} = \beta\boldsymbol{u} + \alpha\boldsymbol{w}. \end{aligned}$$

Therefore, the restriction of $\boldsymbol{A}$ to $\mathrm{Span}_{\mathbb{R}}\{\boldsymbol{u}, \boldsymbol{w}\}$ has the representing matrix $\left(\begin{smallmatrix} \alpha & \beta \\ -\beta & \alpha \end{smallmatrix}\right)$ in the basis $(\boldsymbol{u}, \boldsymbol{w})$.

If we now do this for every nonreal eigenvalue $\lambda_i = \alpha_i + \mathrm{i}\beta_i$, α_i and β_i real, passing from the basis $(\boldsymbol{v}_{i,1}, \dots, \boldsymbol{v}_{i,d_i}, \bar{\boldsymbol{v}}_{i,1}, \dots, \bar{\boldsymbol{v}}_{i,d_i})$ of $\mathrm{Eig}_{\mathbb{C}}(\boldsymbol{A}, \lambda) \oplus \mathrm{Eig}_{\mathbb{C}}(\boldsymbol{A}, \bar{\lambda})$ to the basis $(\boldsymbol{u}_{i,1}, \boldsymbol{w}_{i,1}, \dots, \boldsymbol{u}_{i,d_i}, \boldsymbol{w}_{i,d_i})$, with $\boldsymbol{u}_{i,j}$ and $\boldsymbol{w}_{i,j}$ the real and imaginary parts of $\boldsymbol{v}_{i,j}$, we get the

Proposition 2.37. *Let $\boldsymbol{A}$ be a real $n \times n$ matrix that is diagonalizable as a complex matrix. Then there is an invertible real matrix $\boldsymbol{S}$ such that $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}$ has the block form*

$$\begin{pmatrix} \lambda_1 &&&&& \\ & \ddots &&&& \\ && \lambda_r &&& \\ &&& \boldsymbol{B}_1 && \\ &&&& \ddots & \\ &&&&& \boldsymbol{B}_s \end{pmatrix},$$

where $\lambda_1, \dots, \lambda_r$ are the real eigenvalues of $\boldsymbol{A}$, and the $\boldsymbol{B}_j$s are 2×2 blocks of the form

$$\boldsymbol{B}_j = \begin{pmatrix} \alpha_j & \beta_j \\ -\beta_j & \alpha_j \end{pmatrix}, \qquad \alpha_j, \beta_j \in \mathbb{R},$$

corresponding each to a pair $(\lambda_j, \bar{\lambda}_j)$ of conjugate nonreal eigenvalues.

Remark 2.38. If we want to apply this result to the solution of linear system of ODEs with constant coefficients, we have to compute the exponential of t times the block matrix in the proposition. To do so, it is enough to compute the exponential of each block. For each of the real eigenvalues, we then simply get the 1×1 block $\mathrm{e}^{\lambda_i t}$. For each block $\boldsymbol{B}_j$ we observe that $\boldsymbol{B}_j = \alpha_j \mathbf{1} + \beta_j \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. By (2.15), we then get

$$\mathrm{e}^{\boldsymbol{B}_j t} = \mathrm{e}^{\alpha_j t}\mathrm{e}^{\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)\beta_j t}.$$

It is a useful exercise to check that[4]

$$\mathrm{e}^{\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)x} = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix}.$$

Therefore,

$$\mathrm{e}^{\boldsymbol{B}_j t} = \mathrm{e}^{\alpha_j t}\begin{pmatrix} \cos(\beta_j t) & \sin(\beta_j t) \\ -\sin(\beta_j t) & \cos(\beta_j t) \end{pmatrix}.$$

This shows that the solution to a linear system of ODEs with constant coefficients whose coefficient matrix is diagonalizable as a complex matrix can be written in terms of real exponentials and trigonometric functions as announced at the end of Remark 2.26.

2.3. Diagonalization of endomorphisms

The problem of diagonalization and the results we have discussed for matrices can be generalized to endomorphisms. We present them here, also as an occasion to recapitulate what we have seen.

Definition 2.39. An endomorphism F of a vector space V is called diagonalizable if there is a basis $\mathcal{B}$ such that $F_{\mathcal{B}}$ is a diagonal matrix.

Note that if v is an element of such a basis, we have $v \neq 0$ and $Fv = \lambda v$ for some scalar λ.

Definition 2.40 (Eigenvectors and eigenvalues). A nonzero vector v is called an eigenvector of an endomorphism F if there is a scalar λ, called the eigenvalue to the eigenvector v, such that $Fv = \lambda v$.

We clearly have the

Theorem 2.41. *An endomorphism F of V is diagonalizable iff V admits a basis of eigenvectors of F.*

If V is finite-dimensional, by Definition 1.74, we may define the characteristic polynomial of an endomorphism F as

$$P_F := \det(F - \lambda\,\mathrm{Id}).$$

We then have the

Lemma 2.42. *The eigenvalues of an endomoprhism of a finite-dimensional space are the roots of its characteristic polynomial.*

[4]This can either be done by explicit resumming the exponential series or by solving the harmonic oscillator with $k = m = 1$.

We also have the following generalization of Lemma 2.29 (with essentially the same proof).

LEMMA 2.43. *A collection of eigenvectors of an endomorphism corresponding to pairwise distinct eigenvalues is linearly independent.*

This implies again the

PROPOSITION 2.44. *If an endomorphism of an n-dimensional space has n pairwise distinct eigenvalues, then it is diagonalizable.*

We also associate to an eigenvalue λ of $F \in \mathrm{End}(V)$ its eigenspace

$$\mathrm{Eig}(F, \lambda) := \ker(F - \lambda\,\mathrm{Id})$$

and its geometric multiplicity

$$d := \dim \mathrm{Eig}(F, \lambda).$$

We have $\mathrm{Eig}(F, \lambda) \cap \mathrm{Eig}(F, \mu) = 0$ for $\lambda \neq \mu$ and the following

THEOREM 2.45. *Let $\lambda_1, \ldots, \lambda_k$ be the pairwise distinct eigenvalues of an endomorphism F of an n-dimensional space V and let d_i denote the geometric multiplicity of λ_i. Then F is diagonalizable iff $d_1 + \cdots + d_k = n$. In this case, we have*

$$V = \mathrm{Eig}(\boldsymbol{A}, \lambda_1) \oplus \cdots \oplus \mathrm{Eig}(\boldsymbol{A}, \lambda_k).$$

If P_F splits into linear factor (e.g., if the ground field is $\mathbb{C}$) as

$$P_F = (-1)^n\,(\lambda - \lambda_1)^{s_1} \cdots (\lambda - \lambda_k)^{s_k},$$

then s_i is called the algebraic multiplicity of λ_i.

REMARK 2.46. For every eigenvalue λ_i, we have $d_i \leq s_i$. Therefore, F is diagonalizable iff $d_i = s_i$ for every i.

In applications it is often important to diagonalize two different endomorphisms at the same time. Of course one has first of all to assume that each of them is diagonalizable. We say that two diagonalizable endomorphisms F and G on a vector space V are simultaneously diagonalizable if they possess a common basis of eigenvectors.

PROPOSITION 2.47 (Simultaneous diagonalization). *Two diagonalizable endomorphisms F and G on a vector space V are simultaneously diagonalizable iff they commute, i.e., $FG = GF$.*

PROOF. See Exercise 2.10 □

In the case of matrices, the above proposition reads more explictly as follows.

COROLLARY 2.48. *Two diagonalizable matrices* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *commute (i.e.,* $\boldsymbol{AB} = \boldsymbol{BA}$*) iff there is an invertible matrix* $\boldsymbol{S}$ *such that* $\boldsymbol{S}^{-1}\boldsymbol{AS} = \boldsymbol{D}$ *and* $\boldsymbol{S}^{-1}\boldsymbol{BS} = \boldsymbol{D}'$ *where* $\boldsymbol{D}$ *and* $\boldsymbol{D}'$ *are diagonal matrices.*

2.3.1. The spectral decomposition. Suppose $F \in \mathrm{End}(V)$ is diagonalizable. If we decompose $V = \mathrm{Eig}(\boldsymbol{A}, \lambda_1) \oplus \cdots \oplus \mathrm{Eig}(\boldsymbol{A}, \lambda_k)$, we have, as in every direct sum, a unique decomposition of every $v \in V$ as $v = w_1 + \cdots + w_k$ with $w_i \in \mathrm{Eig}(\boldsymbol{A}, \lambda_i)$. We let $P_i \colon V \to V$ be the linear map that assigns to a vector v its ith component w_i. The P_is are a complete system of mutually transversal projections; i.e.,

$$P_i^2 = P_i \ \forall i, \qquad P_iP_j = P_jP_i = 0 \ \forall i \neq j, \quad \sum_{i=1}^{k} P_i = \mathrm{Id}\,.$$

Since the ith component of a vector is an eigenvector to λ_i, we have $FP_iv = Fw_i = \lambda_i w_i = \lambda_i P_i v$ for every $v \in V$. As an identity of maps, this reads $FP_i = \lambda_i P_i$. Summing over i, we then get

$$F = \sum_{i=1}^{k} \lambda_i P_i. \tag{2.21}$$

The set $\{\lambda_1, \ldots, \lambda_k\}$ of the pairwise distinct eigenvalues of F is called its spectrum and (2.21) is called the spectral decomposition of F.[5]

2.3.2. The infinite-dimensional case. In Example 2.7, we have seen that some important PDEs in physics are linear and can be viewed as linear ODEs on some infinite-dimensional space.

We will not treat this case here in general, but we will consider an example: the wave equation on a one-dimensional space interval, a.k.a. the vibrating string. Namely, we want to study the PDE

$$\frac{1}{c^2}\frac{\partial^2 \psi}{\partial t^2} = \frac{\partial^2 \psi}{\partial x^2},$$

where the unknown ψ is a function on $\mathbb{R} \times [0, L] \ni (t, x)$.[6] We assume that the string endpoints are fixed:

$$\psi(t, 0) = \psi(t, L) = 0 \quad \text{for all } t$$

[5]This terminology ultimately comes from physics, namely from the fact that the spectral lines of an atom are computed in quantum mechanics by taking differences of eigenvalues of a certain endomorphism, the Hamiltonian operator of the electrons in the atom.

[6]More generally, we could consider an interval $[a, b]$, but translating it to $[0, L]$, with $L = b - a$ its length, simplifies the discussion.

(this models, e.g., the string of a musical instrument). We then introduce the infinite-dimensional vector space

$$V := \{\phi \in C^\infty([0, L]) \mid \phi(0) = \phi(L) = 0\}$$

and regard ψ as a map $\mathbb{R} \to V$.

The right hand side of the wave equation uses the endomorphism $F := \frac{\mathrm{d}^2}{\mathrm{d}x^2}$ of V. We now want to find its eigenvectors, i.e., $\phi \in V \setminus \{0\}$ such that

$$\phi'' = \lambda\phi$$

for some complex scalar λ. Denoting by $\pm\alpha$ the two square roots of λ, we see that the general solution to this equation is

$$\phi(x) = A\mathrm{e}^{\alpha x} + B\mathrm{e}^{-\alpha x},$$

where A and B are complex constants. The endpoint conditions $\phi(0) = \phi(L) = 0$ amount to the linear system

$$\begin{cases} A & + & B & = & 0 \\ A\mathrm{e}^{\alpha L} & + & B\mathrm{e}^{-\alpha L} & = & 0 \end{cases}$$

which has a nontrivial solution ($B = -A \neq 0$) iff the coefficient matrix $\left(\begin{smallmatrix} 1 & 1 \\ \mathrm{e}^{\alpha L} & \mathrm{e}^{-\alpha L} \end{smallmatrix}\right)$ is degenerate. Since its determinant is $\mathrm{e}^{-\alpha L} - \mathrm{e}^{\alpha L}$, we may have a nontrivial solution iff $\mathrm{e}^{2\alpha L} = 1$, i.e., $\alpha = \frac{\mathrm{i}\pi k}{L}$ with k an integer.

The case $k = 0$ yields $\phi = 0$, which is not an eigenvector. For $k \neq 0$, we take $A = \frac{1}{2\mathrm{i}}$ (and hence $B = -\frac{1}{2\mathrm{i}}$), so we have the real eigenvector

$$\phi_k(x) = \sin\left(\frac{\pi k x}{L}\right)$$

corresponding to the eigenvalue $\lambda = -\frac{\pi^2 k^2}{L^2}$. Note that $\phi_{-k} = -\phi_k$, so they are not linearly independent. Therefore, we only consider $k > 0$.

By Lemma 2.43, the collection $(\phi_k)_{k\in\mathbb{Z}_{>0}}$ is linearly independent in V. Therefore, on the subspace V' spanned by this collection, we have that the set of eigenvalues (a.k.a. the spectrum) is

$$\left\{-\frac{\pi^2 k^2}{L^2},\ k \in \mathbb{N}_{>0}\right\}.$$

If the two initial conditions, $\psi|_{t=0}$ and $\frac{\partial\psi}{\partial t}|_{t=0}$, for the wave equation are linear combinations of the ϕ_ks, we may then write the unique solution to the Cauchy problem as a linear combination of the ϕ_ks with time-dependent coefficients. In fact, suppose

$$\psi|_{t=0} = \sum_{k=1}^{\infty} b_{k0}\phi_k,$$
$$\frac{\partial \psi}{\partial t}|_{t=0} = \sum_{k=1}^{\infty} v_{k0}\phi_k,$$

where only finitely many of the b_{k0}s and of the v_{k0}s are different from zero. We can then consider a solution of the form

$$\psi = \sum_{k=1}^{\infty} b_k\phi_k,$$

where the coefficients b_k are now functions of time, and the sum is restricted to the ks for which $b_{k0} \neq 0$ or $v_{k0} \neq 0$. The wave equation then yields separate ODEs for each of these ks, which we can assemble into the Cauchy problems

$$\begin{cases} \ddot{b}_k &= -\frac{\pi^2 c^2 k^2}{L^2} b_k \\ b_k(0) &= b_{k0} \\ \dot{b}_k(0) &= v_{k0} \end{cases}$$

Interestingly, it turns out that one can also make sense of infinite linear combinations $\sum_{k=1}^{\infty} b_k \sin\left(\frac{\pi k x}{L}\right)$, where (b_k) is a sequence of real numbers with appropriate decaying conditions for $k \to \infty$. This is an example of Fourier series. We will return to this in Example 3.38, where we will also learn a method to compute the coefficients b_k of an expansion.

2.4. Trigonalization

Even though not every matrix can be diagonalized, it turns out complex matrices can be brought to a nice upper triangular form. More precisely, we have the

THEOREM 2.49. *Let $\boldsymbol{A}$ be a $n \times n$ matrix whose characteristic polynomial splits into linear factors (e.g., a complex matrix),*

$$P_{\boldsymbol{A}} = (-1)^n (\lambda - \lambda_1)^{s_1} \cdots (\lambda - \lambda_k)^{s_k}. \tag{2.22}$$

Then there is an invertible matrix $\boldsymbol{S}$ such that

$$\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{D} + \boldsymbol{N},$$

where $\boldsymbol{D}$ is a diagonal matrix with the eigenvalues of $\boldsymbol{A}$ as its diagonal entries, $\boldsymbol{N}$ is an upper triangular matrix with zeros on the diagonal, and $\boldsymbol{D}\boldsymbol{N} = \boldsymbol{N}\boldsymbol{D}$.

Before we prove the theorem, let us see its consequences for the exponential of a real or complex matrix $\boldsymbol{A}$ (and hence for the associated system of ODEs). By (2.16), we have that $\mathrm{e}^{\boldsymbol{A}} = \boldsymbol{S}\mathrm{e}^{\boldsymbol{D}+\boldsymbol{N}}\boldsymbol{S}^{-1}$. By (2.15), we have $\mathrm{e}^{\boldsymbol{D}+\boldsymbol{N}} = \mathrm{e}^{\boldsymbol{D}}\mathrm{e}^{\boldsymbol{N}}$. We already know how to compute the exponential of a diagonal matrix, so we are only left with the exponential of $\boldsymbol{N}$. Observe that $\boldsymbol{N}$ applied to a vector whose last $k < n$ components are equal to 0 yields a vector whose last $k+1$ components are equal to 0. Therefore, $\boldsymbol{N}^n$ applied to any vector yields the zero vector. In conclusion, $\boldsymbol{N}^m = 0$ for all $m \geq n$. One says that $\boldsymbol{N}$ is nilpotent (meaning that it has a vanishing power). It follows that $\mathrm{e}^{\boldsymbol{N}}$ is a finite sum. These results do not change if we multiply $\boldsymbol{A}$ by t, so we get

$$\mathrm{e}^{\boldsymbol{A}t} = \boldsymbol{S}^{-1}\mathrm{e}^{\boldsymbol{D}t}\left(\sum_{r=0}^{n-1}\frac{t^r}{r!}\boldsymbol{N}^r\right)\boldsymbol{S}.$$

In particular, this means that a solution of the associated system of ODEs is a combination of exponentials and polynomials.

Also note that $\mathrm{e}^{\boldsymbol{N}}$ is an upper triangular matrix with 1s on the diagonal, so $\det \mathrm{e}^{\boldsymbol{N}} = 1$. Therefore, $\det \mathrm{e}^{\boldsymbol{A}t} = \det \mathrm{e}^{\boldsymbol{D}t}$. Since $\boldsymbol{D}$ is diagonal, we obviously have $\det \mathrm{e}^{\boldsymbol{D}t} = \mathrm{e}^{t \operatorname{tr} \boldsymbol{D}}$. On the other hand, we have $\operatorname{tr} \boldsymbol{A} = \operatorname{tr} \boldsymbol{D}$. In conclusion,

$$\det \mathrm{e}^{\boldsymbol{A}t} = \mathrm{e}^{t \operatorname{tr} \boldsymbol{A}}$$

for every t. This is a purely algebraic proof of Proposition 2.11. (In the theorem we assume that $P_{\boldsymbol{A}}$ splits into linear factors, which might not be the case if $\boldsymbol{A}$ is real. In this case, we can however regard $\boldsymbol{A}$ as a complex matrix, apply the theorem, and get the last identity; finally, we observe that both the left and the right hand side are defined over $\mathbb{R}$.)

2.4.1. Proof of Theorem 2.49. The eigenspace of $\boldsymbol{A}$ associated with the eigenvalue λ may be viewed as the largest subspace on which the restriction of $\boldsymbol{A} - \lambda\mathbf{1}$ is zero. Since we are looking for a basis in which the representing matrix minus the diagonal matrix with the eigenvalues as its diagonal entries is nilpotent, we define the generalized eigenspace associated with the eigenvalue λ as

$$\widetilde{\mathrm{Eig}}[\boldsymbol{A}, \lambda] := \{\boldsymbol{v} \mid \exists m \in \mathbb{N}\ (\boldsymbol{A} - \lambda\mathbf{1})^m\boldsymbol{v} = 0\}.$$

More precisely, we say that $\boldsymbol{v}$ is a generalized λ-eigenvector of rank $m > 0$ if $(\boldsymbol{A} - \lambda\mathbf{1})^m\boldsymbol{v} = 0$ but $(\boldsymbol{A} - \lambda\mathbf{1})^{m-1}\boldsymbol{v} \neq 0$. Note that every nonzero vector in $\widetilde{\mathrm{Eig}}[\boldsymbol{A}, \lambda]$ is a generalized λ-eigenvector with a well-defined rank. In particular, an eigenvector in the original sense is a generalized eigenvector of rank 1, so

$$\operatorname{Eig}(\boldsymbol{A}, \lambda) \subseteq \widetilde{\operatorname{Eig}}[\boldsymbol{A}, \lambda].$$

Also observe that $\boldsymbol{A} - \lambda\mathbf{1}$ applied to a generalized λ-eigenvector of rank $m > 1$ yields a generalized λ-eigenvector of rank $m - 1$.

Finally, note that $\widetilde{\operatorname{Eig}}[\boldsymbol{A}, \lambda]$ is $\boldsymbol{A}$-invariant, so also $(\boldsymbol{A} - \lambda\mathbf{1})$-invariant. We denote by $\boldsymbol{N}_\lambda$ the restriction of $\boldsymbol{A} - \lambda\mathbf{1}$ to $\widetilde{\operatorname{Eig}}[\boldsymbol{A}, \lambda]$. By the definition of generalized eigenspace, $\boldsymbol{N}_\lambda$ is nilpotent.

Let $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_d)$ be a basis of $\widetilde{\operatorname{Eig}}[\boldsymbol{A}, \lambda]$. We order the basis in such a way that the rank of $\boldsymbol{v}_j$ is less than or equal to the rank of $\boldsymbol{v}_{j+1}$ for $j = 1, \dots, d-1$. It follows that $\boldsymbol{N}_\lambda$ is represented in this basis by an upper triangular matrix with zeros on the diagonal. Moreover, $\boldsymbol{N}_\lambda$ clearly commutes with $\lambda \operatorname{Id}_{\widetilde{\operatorname{Eig}}[\boldsymbol{A},\lambda]}$. We are then done after proving the following

Proposition 2.50. *Under the assumptions of Theorem 2.49:*

1. *the algebraic multiplicity s_i of λ_i is equal to the dimension δ_i of $\widetilde{\operatorname{Eig}}[\boldsymbol{A}, \lambda_i]$ for every i, and*
2. *we have the decomposition $\mathbb{K}^n = \widetilde{\operatorname{Eig}}[\boldsymbol{A}, \lambda_1] \oplus \cdots \oplus \widetilde{\operatorname{Eig}}[\boldsymbol{A}, \lambda_k]$.*

In fact, it is enough to choose a basis of each generalized eigenspace, ordered by rank as above. In the basis of $\mathbb{K}^n$ given by the union of these bases, $\boldsymbol{A}$ is represented by a matrix $\boldsymbol{D} + \boldsymbol{N}$ as in the theorem.[7]

We start by proving the

Lemma 2.51. *The sum $\widetilde{\operatorname{Eig}}[\boldsymbol{A}, \lambda_1] + \cdots + \widetilde{\operatorname{Eig}}[\boldsymbol{A}, \lambda_k]$ is direct.*

Proof. We use the criterion of Remark 1.22. Suppose we have $\boldsymbol{v}_1 + \cdots + \boldsymbol{v}_k = 0$, $\boldsymbol{v}_i \in \widetilde{\operatorname{Eig}}[\boldsymbol{A}, \lambda_i]$.

Suppose, by contradiction, that some $\boldsymbol{v}_i$ is different from zero, and let $m_i > 0$ be its rank. Since $\boldsymbol{w}_i := (\boldsymbol{A} - \lambda_i\mathbf{1})^{m_i-1}\boldsymbol{v}_i \neq 0$ is an eigenvector for λ_i, we have

$$\prod_{j=1}^{k}(\boldsymbol{A} - \lambda_j\mathbf{1})^{m_j-1}\boldsymbol{v}_i = \prod_{j\neq i}(\boldsymbol{A} - \lambda_j\mathbf{1})^{m_j-1}\boldsymbol{w}_i = \prod_{j\neq i}(\lambda_i - \lambda_j)^{m_j-1}\boldsymbol{w}_i =: \boldsymbol{z}_i.$$

Since $\prod_{j\neq i}(\lambda_i - \lambda_j)^{m_j-1}$ is different from zero, we get that $\boldsymbol{z}_i \neq 0$.

If we now apply $\prod_{j=1}^{k}(\boldsymbol{A} - \lambda_j\mathbf{1})^{m_j}$ to $\boldsymbol{v}_1 + \cdots + \boldsymbol{v}_k = 0$, we get $\sum_{r:\boldsymbol{v}_r\neq 0} \boldsymbol{z}_r = 0$. Since the $\boldsymbol{z}_r$s in the sum are also different from zero, this is in contradiction with Lemma 2.29.

[7] In particular, the restriction of $\boldsymbol{D}$ to $\widetilde{\operatorname{Eig}}[\boldsymbol{A}, \lambda_j]$ is $\lambda_j \operatorname{Id}_{\widetilde{\operatorname{Eig}}[\boldsymbol{A},\lambda_j]}$, whereas the restriction of $\boldsymbol{N}$ is $\boldsymbol{N}_{\lambda_j}$.

Therefore, the zero vector has a unique decomposition and the sum is direct.

□

PROOF OF PROPOSITION 2.50. Note that part (1) of the statement implies part (2) using Lemma 2.51. In fact,

$$\dim(\widetilde{\mathrm{Eig}}[\boldsymbol{A},\lambda_1]\oplus\cdots\oplus\widetilde{\mathrm{Eig}}[\boldsymbol{A},\lambda_k]) = \delta_1+\cdots+\delta_k = s_1+\dots s_k = n,$$

so the direct sum is the whole space $\mathbb{K}^n$.

Therefore, we will only prove statement (1) by induction on the dimension n. For $n = 1$ there is nothing to prove, as in this case $\boldsymbol{A} = (\lambda)$ and $\widetilde{\mathrm{Eig}}[\boldsymbol{A},\lambda] = \mathrm{Eig}(\boldsymbol{A},\lambda) = \mathbb{K}$.

Next, we assume we have proved (1), and hence (2), for dimensions up to $n-1$. We pick an $i \in \{1,\dots,k\}$, denote by $\boldsymbol{A}_i$ the restriction of $\boldsymbol{A}$ to $\widetilde{\mathrm{Eig}}[\boldsymbol{A},\lambda_i]$, which is $\boldsymbol{A}$-invariant, and choose a complement W to $\widetilde{\mathrm{Eig}}[\boldsymbol{A},\lambda_i]$ in $\mathbb{K}^n$. With respect to the decomposition $\mathbb{K}^n = \widetilde{\mathrm{Eig}}[\boldsymbol{A},\lambda_i]\oplus W$, $\boldsymbol{A}$ has the form

$$\begin{pmatrix}\boldsymbol{A}_i & \boldsymbol{B}\\ \boldsymbol{0} & \boldsymbol{C}\end{pmatrix},$$

where $\boldsymbol{B}$ is the composition of the restriction of $\boldsymbol{A}$ to W with the projection to $\widetilde{\mathrm{Eig}}[\boldsymbol{A},\lambda_i]$ and $\boldsymbol{C}$ is the composition of the restriction of $\boldsymbol{A}$ to W with the projection to W.

By property (D.11) of the determinant, we have $P_{\boldsymbol{A}} = P_{\boldsymbol{A}_i}P_{\boldsymbol{C}}$. If we take a rank-ordered basis of $\widetilde{\mathrm{Eig}}[\boldsymbol{A},\lambda_i]$, we have that $\boldsymbol{A}_i$ is represented by an upper triangular matrix with diagonal entries equal to λ_i. Therefore, by property (D.6), $P_{\boldsymbol{A}_i} = (-1)^{\delta_i}(\lambda-\lambda_i)^{\delta_i}$. In conclusion, $P_{\boldsymbol{A}} = (-1)^{\delta_i}(\lambda-\lambda_i)^{\delta_i}P_{\boldsymbol{C}}$. Comparing with (2.22), we have that

$$P_{\boldsymbol{C}} = (-1)^{\dim W}(\lambda-\lambda_1)^{s_1'}\cdots(\lambda-\lambda_k)^{s_k'}$$

with $s_j' = s_j$ for $j \neq i$ and $s_i' = s_i - \delta_i$. We claim that $s_i' = 0$, which in particular implies $\delta_i = s_i$. Since this can be done for every i, this completes the proof of the proposition.

We now prove the claim that $s_i' = 0$. Since $\dim W < n$ and $P_{\boldsymbol{C}}$ splits into linear factors, we may apply the induction hypothesis, so[8] $W = \widetilde{\mathrm{Eig}}[\boldsymbol{C},\lambda_1]\oplus\cdots\oplus\widetilde{\mathrm{Eig}}[\boldsymbol{C},\lambda_k]$.

Consider the space $V = \widetilde{\mathrm{Eig}}[\boldsymbol{A},\lambda_i]\oplus\widetilde{\mathrm{Eig}}[\boldsymbol{C},\lambda_i]$. For $\boldsymbol{v}$ in the first summand, we have $\boldsymbol{A}\boldsymbol{v} \in \widetilde{\mathrm{Eig}}[\boldsymbol{A},\lambda_i]$; for $\boldsymbol{v}$ in the second summand, we have $\boldsymbol{A}\boldsymbol{v} = \boldsymbol{B}\boldsymbol{v} + \boldsymbol{C}_i\boldsymbol{v}$, where $\boldsymbol{C}_i$ denotes the restriction of $\boldsymbol{C}$ to $\widetilde{\mathrm{Eig}}[\boldsymbol{C},\lambda_i]$, which is a $\boldsymbol{C}$-invariant subspace. Therefore, the restriction

[8] We actually choose a basis to identify W with $\mathbb{K}^{\dim W}$ and apply statement (2) to the representing matrix of $\boldsymbol{C}$.

$\boldsymbol{A}_V$ of $\boldsymbol{A}$ to V has the following form with respect to the decomposition:[9]

$$\begin{pmatrix} \boldsymbol{A}_i & \boldsymbol{B} \\ \boldsymbol{0} & \boldsymbol{C}_i \end{pmatrix}.$$

One easily proves, by induction on r, that

$$(\boldsymbol{A}_V - \lambda_i \boldsymbol{1})^r = \begin{pmatrix} \boldsymbol{A}_i - \lambda_i \boldsymbol{1} & \boldsymbol{B} \\ \boldsymbol{0} & \boldsymbol{C}_i - \lambda_i \boldsymbol{1} \end{pmatrix}^r = \begin{pmatrix} (\boldsymbol{A}_i - \lambda_i \boldsymbol{1})^r & \boldsymbol{B}_r \\ \boldsymbol{0} & (\boldsymbol{C}_i - \lambda_i \boldsymbol{1})^r \end{pmatrix}$$

for some matrix $\boldsymbol{B}_r$. If $\boldsymbol{v} \in \widetilde{\mathrm{Eig}}[\boldsymbol{C}, \lambda_i]$ has rank r, we then have $(\boldsymbol{A} - \lambda_i \boldsymbol{1})^r \boldsymbol{v} = \boldsymbol{B}_r \boldsymbol{v}$. Since this is now an element of $\widetilde{\mathrm{Eig}}[\boldsymbol{A}, \lambda_i]$, there is an s such that $(\boldsymbol{A} - \lambda_i \boldsymbol{1})^s \boldsymbol{B}_r \boldsymbol{v} = 0$. This means that $(\boldsymbol{A} - \lambda_i \boldsymbol{1})^{r+s} \boldsymbol{v} = 0$, i.e., $\boldsymbol{v} \in \widetilde{\mathrm{Eig}}[\boldsymbol{A}, \lambda_i]$. Finally, since $\widetilde{\mathrm{Eig}}[\boldsymbol{C}, \lambda_i]$ belongs to a complement of $\widetilde{\mathrm{Eig}}[\boldsymbol{A}, \lambda_i]$, the only vector in their intersection is 0. We have thus proved that $\widetilde{\mathrm{Eig}}[\boldsymbol{C}, \lambda_i] = 0$ and, hence, that $s'_i = 0$. □

Another interesting consequence of part (1) of Proposition 2.50 is the following result, announced in Remark 2.32.

COROLLARY 2.52. *The geometric multiplicity of every eigenvalue is less than or equal to its algebraic multiplicity.*

PROOF. Since $\mathrm{Eig}(\boldsymbol{A}, \lambda_i) \subseteq \widetilde{\mathrm{Eig}}[\boldsymbol{A}, \lambda_i]$, we have $\dim \mathrm{Eig}(\boldsymbol{A}, \lambda_i) \leq \delta_i = s_i$. □

2.5. Digression: The Jordan normal form

By choosing a more suitable basis of each $\widetilde{\mathrm{Eig}}[\boldsymbol{A}, \lambda_i]$, the strictly upper triangular matrix $\boldsymbol{N}$ of Theorem 2.49 may be put in a "canonical" form that can be more easily dealt with.

For this we need the notion of a Jordan block (after the French mathematician CAMILLE JORDAN). For a scalar λ and a positive integer m, the Jordan block $\boldsymbol{J}_{\lambda,m}$ is the $m \times m$ upper triangular matrix whose diagonal entries are equal to λ, the entries right above the diagonal are equal to 1, and all other entries vanish; e.g.,

$$\boldsymbol{J}_{\lambda,1} = \begin{pmatrix} \lambda \end{pmatrix}, \quad \boldsymbol{J}_{\lambda,2} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad \boldsymbol{J}_{\lambda,3} = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix},$$

[9]By abuse of notation, we keep writing $\boldsymbol{B}$, but what appears here is actually the restriction of $\boldsymbol{B}$ to $\widetilde{\mathrm{Eig}}[\boldsymbol{C}, \lambda_i]$.

and

$$\boldsymbol{J}_{\lambda,4} = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}.$$

A Jordan matrix is a block diagonal matrix whose diagonal blocks are Jordan blocks.

THEOREM 2.53. *Let $\boldsymbol{A}$ be a matrix as in Theorem 2.49. Then there is a basis in which $\boldsymbol{A}$ is represented by a Jordan matrix. More precisely, for each eigenvalue λ_i, the generalized eigenspace $\widetilde{\mathrm{Eig}}[\boldsymbol{A}, \lambda_i]$ has a basis in which the restriction of $\boldsymbol{A}$ is represented by a Jordan matrix of the form*

$$\begin{pmatrix} \boldsymbol{J}_{\lambda_i,m_1} & & \\ & \ddots & \\ & & \boldsymbol{J}_{\lambda_i,m_k} \end{pmatrix},$$

with $0 < m_1 \leq \cdots \leq m_k$ and $m_1 + \cdots + m_k = s_i$.

To prove the theorem, it is enough to prove the statement for each generalized eigenspace separately. By definition of generalized eigenspace, the restriction $\boldsymbol{N}_i$ of $\boldsymbol{A} - \lambda\mathbf{1}$ to $\widetilde{\mathrm{Eig}}[\boldsymbol{A}, \lambda_i]$ is nilpotent. It follows that it is enough to prove the following

PROPOSITION 2.54. *Let N be a nilpotent operator on an n-dimensional vector space V. Then there is a basis of V in which N is represented by a Jordan matrix of the form*

$$\begin{pmatrix} \boldsymbol{J}_{0,m_1} & & \\ & \ddots & \\ & & \boldsymbol{J}_{0,m_k} \end{pmatrix},$$

with $0 < m_1 \leq \cdots \leq m_k$ and $m_1 + \cdots + m_k = n$.

To prove the proposition, we need some preliminary remarks. The first remark is that a nilpotent operator is not invertible, so it has a nonzero kernel.

We say that a vector v in V has rank $m > 0$ if $N^m v = 0$ but $N^{m-1} v \neq 0$. For a vector v of rank m we define

$$v_j := N^{m-j} v$$

for $j = 1, \ldots, m$. The vectors $v_1, \ldots, v_m$ are called a Jordan chain (note that $v_m = v$ and that v_1 is in the kernel of N). We denote by $\mathrm{Jor}(v)$ their span:

$$\mathrm{Jor}(v) := \mathrm{Span}\{v_1, \ldots, v_m\}.$$

The next remark is that $\operatorname{Jor}(v)$ is N-invariant and that the vectors $v_1, \dots, v_m$ form a basis. That they generate $\operatorname{Jor}(v)$ is obvious by definition, so we only have to check that they are linearly independent. Suppose $\alpha^i v_i = 0$ for some scalars $\alpha^1, \dots, \alpha^m$. Applying N^{m-1} to this linear combination yields $\alpha^m N^{m-1} v = 0$, which implies $\alpha^m = 0$. If we then apply N^{m-2} to the linear combination, knowing that $\alpha^m = 0$ we get $\alpha^{m-1} N^{m-1} v = 0$ which implies $\alpha^{m-1} = 0$, and so on. In particular, this shows that $\dim \operatorname{Jor}(v) = m$.

The final remark is that, by construction, the restriction of N to $\operatorname{Jor}(v)$ in the basis $(v_1, \dots, v_m)$ is represented by the Jordan block $\boldsymbol{J}_{0,m}$.

The strategy to prove Proposition 2.54 consists then in decomposing V into spans of Jordan chains.

Note that the vector v of rank m we started with to define $\operatorname{Jor}(v)$ might be in the image of N, say, $v = Nw$. In this case, we may extend the Jordan chain $\operatorname{Jor}(v)$ to the Jordan chain $\operatorname{Jor}(w) \supsetneq \operatorname{Jor}(v)$ (note that $w_i = v_i$ for $j = 1, \dots, m$ and that $w_{m+1} = w$). If v is not in the image of N, we say that $v_1, \dots, v_m$ is a maximal Jordan chain and that v is a lead vector for it. Proposition 2.54 is then a consequence of the following

Lemma 2.55. *Let N be a nilpotent operator on an n-dimensional vector space V. Then there is a collection $v_{(1)}, \dots, v_{(k)}$ of lead vectors*[10] *of ranks $m_1, \dots, m_k$, respectively, such that $V = \operatorname{Jor}(v_{(1)}) \oplus \cdots \oplus \operatorname{Jor}(v_{(k)})$.*

Note that we may arrange the lead vectors $v_{(1)}, \dots, v_{(k)}$ so that $0 < m_1 \leq \cdots \leq m_k$ as in Proposition 2.54.

Proof of the lemma. We prove the lemma by induction on the dimension n of V. If $n = 0$, there is nothing to prove.

Next, assume we have proved the lemma up to dimension $n - 1$. Let $W \subseteq V$ be the image of N. Note that W is N-invariant. By the dimension formula (1.4), we have $\dim W = n - \dim \ker N < n$, since N has nonzero kernel, so we can apply the induction assumption to W. Namely, we can find vectors $v_{(1)}, \dots, v_{(k)}$ in W such that $W = \operatorname{Jor}(v_{(1)}) \oplus \cdots \oplus \operatorname{Jor}(v_{(k)})$. (The $v_{(i)}$s are lead vectors in W but not in V.)

We now have two cases to consider. The first case is when $W \cap \ker N = 0$.[11] In this case $V = W \oplus \ker N$. A basis $(z_{(1)}, \dots, z_{(l)})$ of $\ker N$ produces the decomposition $\ker N = \operatorname{Jor}(z_{(1)}) \oplus \operatorname{Jor}(z_{(l)})$ (note

[10] We write $v_{(i)}$ for the lead vectors and $v_{(i),j}$ for the vectors in the corresponding Jordan chain.

[11] This case in particular happens when $n = 1$.

that $\operatorname{Jor}(z_{(i)}) = \mathbb{K}z_{(i)}$, for $z_{(i)}$ is in the kernel of N). This concludes the proof in this case.

The other case is when $W \cap \ker N \neq 0$. Let W' be a complement of $W + \ker N$ in V. In particular, the restriction of N to W' is injective. Therefore, we have uniquely determined $w_{(1)}, \dots, w_{(k)}$ in W' satisfying $v_{(i)} = Nw_{(i)}$ for $i = 1, \dots, k$.

We claim that $(w_{(1)}, \dots, w_{(k)})$ is a basis of W'. This completes the proof of the lemma, since it yields the decomposition $V = \operatorname{Jor}(w_{(1)}) \oplus \cdots \oplus \operatorname{Jor}(w_{(k)}) \oplus \operatorname{Jor}(z_{(1)}) \oplus \operatorname{Jor}(z_{(l)})$, where $(z_{(1)}, \dots, z_{(l)})$ is a basis of a complement of W in $W + \ker N$.

To prove the claim, take some v in W'. We have to show that it has a unique decomposition in $w_{(1)}, \dots, w_{(k)}$. Since Nv is in W, we may expand it in the basis $(N^{j_i}v_{(i)})_{i=1,\dots,k,\ j_i=0,\dots,m_i'-1}$, where m_i' is the rank of $v_{(i)}$ in W. Note that all these vectors but $v_{(1)}, \dots, v_{(k)}$ are in the image of N^2. Therefore, we have uniquely determined scalars $\alpha^1, \dots, \alpha^k$ and some vector w such that $Nv = \sum_i \alpha^i v_{(i)} + N^2 w$. Setting $\widetilde{v} := v - \sum_i \alpha^i w_{(i)}$, we get $N\widetilde{v} = N^2 w$, so $\widetilde{v} - Nw \in \ker N$ and hence $\widetilde{v} \in W + \ker N$. Since, however, $\widetilde{v} \in W'$, which is a complement of $W + \ker N$, we get $\widetilde{v} = 0$ and hence $v = \sum_i \alpha^i w_{(i)}$. □

Exercises for Chapter 2

2.1. Applying the formula $\mathrm{e}^{\boldsymbol{A}t} = \boldsymbol{S}\mathrm{e}^{\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}t}\boldsymbol{S}^{-1}$, compute $\mathrm{e}^{\boldsymbol{A}t}$ for $\boldsymbol{A} = \left(\begin{smallmatrix} 0 & 2 \\ 8 & 0 \end{smallmatrix}\right)$ using $\boldsymbol{S} = \left(\begin{smallmatrix} 1 & 1 \\ 2 & -2 \end{smallmatrix}\right)$.

2.2. Let $\boldsymbol{A} = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$.
(a) Compute $\boldsymbol{A}^n$ for all integers $n > 0$.
Hint: Distinguish the cases n even and n odd.
(b) Using the above result and writing the exponential series as

$$\mathrm{e}^{\boldsymbol{A}t} = \sum_{s=0}^{\infty} \frac{1}{(2s)!} t^{2s} \boldsymbol{A}^{2s} + \sum_{s=0}^{\infty} \frac{1}{(2s+1)!} t^{2s+1} \boldsymbol{A}^{2s+1},$$

show that

$$\mathrm{e}^{\boldsymbol{A}t} = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}.$$

Hint: Use the series expansions of sin and cos.

2.3. Let $\boldsymbol{A} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and $\boldsymbol{B} = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. Compute $\mathrm{e}^{\boldsymbol{A}t}$, $\mathrm{e}^{\boldsymbol{B}t}$, $\mathrm{e}^{\boldsymbol{A}t}\mathrm{e}^{\boldsymbol{B}t}$, $\mathrm{e}^{\boldsymbol{B}t}\mathrm{e}^{\boldsymbol{A}t}$, and $\mathrm{e}^{(\boldsymbol{A}+\boldsymbol{B})t}$. (Hint: Proceed as in the previous exercise.)

2.4. Determine the characteristic polynomial of the following matrices, and find their eigenvalues and a basis of eigenvectors:

$$\boldsymbol{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \boldsymbol{C} = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix}.$$

2.5. Let $\boldsymbol{A} = \begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix}$.

(a) Find the similarity transformation that diagonalizes $\boldsymbol{A}$, i.e., find a matrix $\boldsymbol{S}$ such that $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{D}$ is a diagonal matrix and compute $\boldsymbol{D}$ explicitly.

(b) Using your results just obtained, find the solution to the Cauchy problem given by

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{A}\boldsymbol{x}, \qquad \boldsymbol{x}(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

2.6. In this exercise we prove that $P_{\boldsymbol{A}}(\boldsymbol{A}) = \boldsymbol{0}$ for any 2×2 matrix,[12] where $P_{\boldsymbol{A}}(\lambda)$ is the characteristic polynomial of $\boldsymbol{A}$ in λ. For this we can proceed as follows:

(a) Show that for any 2×2 matrix $\boldsymbol{A}$ the characteristic polynomial can be written as

$$P_{\boldsymbol{A}}(\lambda) = \lambda^2 - \mathrm{tr}(\boldsymbol{A})\lambda + \det(\boldsymbol{A}).$$

(b) We will now interpret it as a polynomial in the matrix $\boldsymbol{A}$. Show that

$$\boldsymbol{A}^2 - \mathrm{tr}(\boldsymbol{A})\,\boldsymbol{A} + \det(\boldsymbol{A})\,\mathbf{1} = \mathbf{0}.$$

2.7. Let J be an endomorphism of V satisfying $J^2 = \mathrm{id}$.

(a) Show that $\lambda = \pm 1$ are the only possible eigenvalues of J.

(b) Using

$$v = \frac{v + Jv}{2} + \frac{v - Jv}{2},$$

show that J is diagonalizable.

---

[12] The result actually holds for matrices of any size and is known as the Cayley–Hamilton theorem.

2.8. Motivated by the study of the vibrating string where the endpoints are free to slide frictionless in the vertical direction, we consider the endomorphism $F := \frac{\mathrm{d}^2}{\mathrm{d}x^2}$ of the vector space

$$V := \{\phi \in C^\infty([0,\ell]) \mid \phi'(0) = \phi'(\ell) = 0\}.$$

Find all eigenvalues of F and a corresponding linearly independent system of eigenvectors.

2.9. Let

$$\boldsymbol{A} = \begin{pmatrix} 5 & -1 & -2 \\ -1 & 5 & 2 \\ 0 & 0 & 6 \end{pmatrix}.$$

(a) Find all eigenvalues of $\boldsymbol{A}$.
(b) Find linearly independent eigenvectors corresponding to all eigenvalues.
(c) Find an invertible matrix $\boldsymbol{S}$ such that $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}$ is upper triangular.
Hint: Find an appropriate basis for each generalized eigenspace.

2.10. The goal of this exercise is to show that two <u>diagonalizable</u> endomorphisms F and G on a vector space[13] V are simultaneously diagonalizable—i.e., possess a common basis of eigenvectors—iff they commute—i.e., $FG = GF$.
(a) Assume that F and G have a a common basis of eigenvectors. Show that they commute.
(b) Now assume that F and G commute.
(i) Show that every eigenspace of F is G-invariant.
(ii) Let $(\lambda_1, \dots, \lambda_k)$ be the pairwise distinct eigenvalues of F. Let v be an eigenvector of G with eigenvalue μ. Let $v = v_1 + \cdots + v_k$ be the unique decomposition of v with $v_i \in \mathrm{Eig}(F, \lambda_i)$. Show that $Gv_i = \mu v_i$ for every i.
Hint: Use point (a) and the uniqueness of the decomposition.
(iii) Show that $\mathrm{Eig}(G,\mu) = \bigoplus_{i=1}^k (\mathrm{Eig}(F,\lambda_i) \cap \mathrm{Eig}(G,\mu))$.[14]
Hint: Use point (b).

[13]For notational simplicity, we assume V to be finite-dimensional.
[14]Some of these intersections might be the zero space.

(iv) Conclude that

$$V = \bigoplus_{\substack{i=1,\dots,k \\ j=1,\dots,l}} \mathrm{Eig}(F, \lambda_i) \cap \mathrm{Eig}(G, \mu_j),$$

where $(\mu_1, \dots, \mu_l)$ are the pairwise distinct eigenvalues of G.

(v) Conclude that F and G have a common basis of eigenvectors.

CHAPTER 3

Inner Products

In this chapter we discuss inner products, a generalization of the familiar dot product, a.k.a. the scalar product.

REMARK 3.1 (Terminology). The term scalar product is unfortunately used both as a synonym of dot product and as a synonym of inner product. For this reason, we will avoid using this term. We will only speak of dot product (for the well-known special case) and of inner product (for the generalization, which comprises the dot product as a special case).

3.1. The dot product

3.1.1. The dot product on the plane. Recall that the dot product of two vectors $\boldsymbol{v} = \binom{v^1}{v^2}$ and $\boldsymbol{w} = \binom{w^1}{w^2}$ on the plane is defined as

$$\boldsymbol{v} \cdot \boldsymbol{w} := v^1 w^1 + v^2 w^2. \tag{3.1}$$

Also recall that the length $\|\boldsymbol{v}\|$ of $\boldsymbol{v}$, also known as the norm of the vector $\boldsymbol{v}$, is defined by the Pythagorean theorem as $\|\boldsymbol{v}\| = \sqrt{(v^1)^2 + (v^2)^2}$, so we have

$$\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}. \tag{3.2}$$

If $\boldsymbol{v}$ and $\boldsymbol{w}$ are different from zero, we may find the oriented angle θ

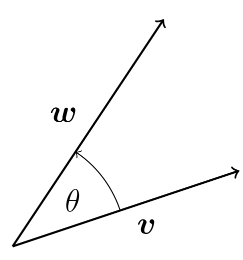


FIGURE 3.1. The oriented angle between two vectors

between them, see Figure 4.1, and also write

$$\boldsymbol{v} \cdot \boldsymbol{w} = \|\boldsymbol{v}\| \|\boldsymbol{w}\| \cos\theta. \tag{3.3}$$

Computing the sum of the vectors as in Figure 3.2, by the law of cosines

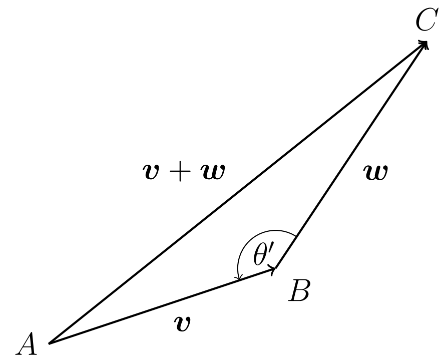


FIGURE 3.2. The law of cosines

(the generalization of the Pythagorean theorem), we have

$$\overline{AC}^2 = \overline{AB}^2 + \overline{BC}^2 - 2\overline{AB}\,\overline{BC}\cos\theta'.$$

By writing the length of the sides of the triangle in terms of the norms of the corresponding vectors, and observing that $\theta' = \pi - \theta$, we get

$$\|\boldsymbol{v}+\boldsymbol{w}\|^2 = \|\boldsymbol{v}\|^2 + \|\boldsymbol{w}\|^2 + 2\|\boldsymbol{v}\|\|\boldsymbol{w}\|\cos\theta. \tag{3.4}$$

Since $\cos\theta \leq 1$, we get, after taking the square roots of both sides, the triangle inequality

$$\|\boldsymbol{v}+\boldsymbol{w}\| \leq \|\boldsymbol{v}\| + \|\boldsymbol{w}\| \tag{3.5}$$

stating that the length of one side of a triangle cannot exceed the sum of the lengths of the two other sides (the inequality is saturated—i.e., it becomes an equality—iff the triangle is degenerate).

3.1.2. The dot product in n dimensions. We now generalize the above considerations to the n-dimensional space $\mathbb{R}^n$. The dot product of two n-dimensional vectors

$$\boldsymbol{v} = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \qquad \boldsymbol{w} = \begin{pmatrix} w^1 \\ \vdots \\ w^n \end{pmatrix}$$

is defined componentwise generalizing (3.1):

$$\boldsymbol{v}\cdot\boldsymbol{w} := \sum_{i=1}^{n} v^i w^i.$$

Note that, using transposition, this can also be written as

$$\boldsymbol{v}\cdot\boldsymbol{w} = \boldsymbol{v}^{\mathsf{T}}\boldsymbol{w}, \tag{3.6}$$

where we use matrix multiplication on the right hand side. The length (or norm) $\|\boldsymbol{v}\|$ of $\boldsymbol{v}$ is defined by the n-dimensional extension of the

Pythagorean theorem as $\|\boldsymbol{v}\| = \sqrt{\sum_{i=1}^n (v^i)^2}$, so we have again (3.2). If $\boldsymbol{v}$ and $\boldsymbol{w}$ are different from zero, they span a plane and inside this plane we find the oriented angle θ between them, so we have again (3.3), the cosine law (3.4), and the triangle inequality (3.5).

3.2. Inner product spaces

We now want to generalize all the above to general vector spaces over $\mathbb{R}$.

Definition 3.2 (Inner products). A positive-definite symmetric bilinear form on a real vector space is called an inner product. A real vector space endowed with an inner product is called an inner product space (a.k.a. a euclidean space).

More explicitly, if V is a vector space over $\mathbb{R}$, an inner product is a map $V \times V \to \mathbb{R}$, usually denoted by $(\ ,\ )$, satisfying the following three properties.

Bilinearity: For all $v, v_1, v_2, w, w_1, w_2 \in V$ and all $\lambda_1, \lambda_2 \in \mathbb{R}$, we have

$$(\lambda_1 v_1 + \lambda_2 v_2, w) = \lambda_1 (v_1, w) + \lambda_2 (v_2, w),$$
$$(v, \lambda_1 w_1 + \lambda_2 w_2) = \lambda_1 (v, w_1) + \lambda_2 (v, w_2).$$

Symmetry: For all $v, w \in V$, we have

$$(v, w) = (w, v).$$

Positivity: For all $v \in V \setminus \{0\}$, we have

$$(v, v) > 0.$$

Remark 3.3. Because of symmetry, it is enough to verify linearity on one of the two arguments, as that on the other follows. Moreover, bilinearity implies

$$(\lambda v, \lambda v) = \lambda^2 (v, v) \tag{3.7}$$

for all $v \in V$ and all $\lambda \in \mathbb{R}$. In particular, setting $\lambda = 0$, we get $(0, 0) = 0$. By positivity, we then get $(v, v) \geq 0$ for every $v \in V$.

Example 3.4 (Dot product). The dot product $\boldsymbol{v} \cdot \boldsymbol{w}$ on $\mathbb{R}^n$ is an example of inner product. We will see (Theorem 3.84) that, upon choosing an appropriate basis, an inner product on a finite-dimensional vector space can always be brought to this form.

Example 3.5 (Subspaces). If W is a subspace of an inner product space V, we may restrict the inner product to elements of W. This makes W itself into an inner product space.

EXAMPLE 3.6. On $\mathbb{R}^n$ we may also define

$$(\boldsymbol{v}, \boldsymbol{w}) := \sum_{i=1}^{n} \lambda_i v^i w^i$$

for a given choice of real numbers $\lambda_1, \dots, \lambda_n$. This is clearly bilinear and symmetric. One can easily verify that it is positive definite iff $\lambda_i > 0$ for all $i = 1, \dots, n$.

EXAMPLE 3.7. More generally, on $\mathbb{R}^n$ we may define

$$(\boldsymbol{v}, \boldsymbol{w}) := \boldsymbol{v}^\mathsf{T} \boldsymbol{g} \boldsymbol{w} = v^i g_{ij} w^j,$$

where $\boldsymbol{g}$ is a given $n \times n$ real matrix, and we have used Einstein's convention in the last term. (The previous example is the case when $\boldsymbol{g}$ is diagonal.) Bilinearity is clear. Symmetry is satisfied iff $\boldsymbol{g}$ is symmetric.[1] A symmetric matrix $\boldsymbol{g}$ is called positive definite if the corresponding symmetric bilinear form is positive definite, i.e., if $\boldsymbol{v}^\mathsf{T} \boldsymbol{g} \boldsymbol{v} > 0$ for every nonzero vector $\boldsymbol{v}$.

REMARK 3.8 (Representing matrix). If we have an inner product $(\ ,\ )$ on a finite-dimensional space V with a basis $\mathcal{B} = (e_1, \dots, e_n)$, we define the representing matrix $\boldsymbol{g}$ with entries

$$g_{ij} := (e_i, e_j).$$

Upon using the isomorphism $\phi_\mathcal{B} \colon \mathbb{R}^n \to V$ of Remark 1.52, we get on $\mathbb{R}^n$ the inner product of Example 3.7.

The following two examples are (the real version of examples that are) important for quantum mechanics.

EXAMPLE 3.9 (Continuous functions on a compact interval). We consider the vector space $V = C^0([a, b])$ of real-valued functions on the interval $[a, b]$. Then

$$(f, g) := \int_a^b fg \, \mathrm{d}x \tag{3.8}$$

is an inner product on V. Bilinearity and symmetry are obvious. As for positivity, note that, if $f \neq 0$, there is some $x_0 \in [a, b]$ with $f(x_0) = c_0 \neq 0$. By continuity there is some open interval $(c, d) \subset [a, b]$ containing x_0 such that $f(x)^2 > c_0^2/2$ for all $x \in (c, d)$. We may write

$$(f, f) = \int_a^c f^2 \, \mathrm{d}x + \int_c^d f^2 \, \mathrm{d}x + \int_d^b f^2 \, \mathrm{d}x.$$

[1] Recall that a matrix $\boldsymbol{A}$ is called symmetric if $\boldsymbol{A}^\mathsf{T} = \boldsymbol{A}$.

Since the first and the last integral are nonnegative and the middle one is larger than or equal to $\frac{c_0^2}{2}(d-c)$, and hence positive, we get $(f, f) > 0$.

EXAMPLE 3.10 (Compactly supported continuous functions). Denote by $V = C_c^0(\mathbb{R})$ the vector space of real-valued functions on $\mathbb{R}$ with compact support: i.e., f belongs to $C_c^0(\mathbb{R})$ iff it is continuous and there is an interval $[a, b]$ outside of which f vanishes. We define[2]

$$(f, g) := \int_{-\infty}^{\infty} fg \, \mathrm{d}x.$$

This can be proved to be an inner product as in the previous example.

The following is also an important infinite-dimensional example.

EXAMPLE 3.11. A sequence $a = (a_1, a_2, \dots)$ of real numbers is called finite if only finitely many a_is are different from zero (equivalently, if there is an N such that $a_i = 0$ for all $i > N$). We denote by $\mathbb{R}^\infty$ the vector space of all finite real sequences, with vector space operations

$$\lambda(a_1, a_2, \dots) = (\lambda a_1, \lambda a_2, \dots)$$
$$(a_1, a_2, \dots) + (b_1, b_2, \dots) = (a_1 + b_1, a_2 + b_2, \dots).$$

It is an inner product space with

$$(a, b) := \sum_{i=1}^{\infty} a_i b_i,$$

where the right hand side clearly converges because it is a finite sum.

3.2.1. Nondegeneracy. The positivity condition of an inner product $(\ ,\ )$ on V implies in particular the nondegeneracy condition

$$(v, w) = 0 \ \forall w \iff v = 0.$$

In fact, the condition has to be satisfied in particular for $w = v$, so we have $(v, v) = 0$ and hence $v = 0$. A further consequence of this is that the linear map

$$\begin{array}{ccc} V & \to & V^* \\ v & \mapsto & L(v) \end{array},$$

with $L(v)(w) := (v, w)$, is injective.

[2]Note that the integral converges. In fact, f vanishes outside some interval $[a, b]$ and g outside some interval $[a', b']$. Let $[c, d]$ be some interval that contains both $[a, b]$ and $[a', b']$. Then $\int_{-\infty}^{\infty} fg \, \mathrm{d}x = \int_c^d fg \, \mathrm{d}x$ which converges since f and g are continuous.

Remark 3.12 (The induced isomorphism). If V is finite-dimensional, then this map is also surjective: in summary, we get an isomorphism L between V and V^*. We denote by R its inverse.

Example 3.13. In the case of the dot product on $\mathbb{R}^n$—$(\boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{v} \cdot \boldsymbol{w} = \boldsymbol{v}^\mathsf{T}\boldsymbol{w}$ as in (3.6)—the map L is just the usual transposition map $\mathbb{R}^n \to (\mathbb{R}^n)^*$:

$$L(\boldsymbol{v}) = \boldsymbol{v}^\mathsf{T}.$$

Remark 3.14 (Lowering and raising indices). In the case of Example 3.7—$(\boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{v}^\mathsf{T}\boldsymbol{g}\boldsymbol{w}$ on $\mathbb{R}^n$—the map L is given by

$$L(\boldsymbol{v}) = \boldsymbol{v}^\mathsf{T}\boldsymbol{g}.$$

As $(\ ,\ )$ is positive definite, the map L is an isomorphism. This implies that the matrix $\boldsymbol{g}$ is invertible. Note that $L(\boldsymbol{v})$ is the row vector $\boldsymbol{\alpha}$ with components

$$\alpha_j = v^i g_{ij}, \quad j = 1, \ldots, n.$$

For this reason, applying L is also called the operation of lowering indices. It is customary, especially in the physics literature, to denote by g^{ij} (note the upper indices!) the entries of the inverse matrix $\boldsymbol{g}^{-1}$. That is,

$$g^{ij} g_{jl} = \delta^i_l,$$

where we have used the Kronecker delta. The map $R\colon (\mathbb{R}^n)^* \to \mathbb{R}^n$, inverse to L, maps a row vector $\boldsymbol{\alpha}$ to the column vector $\boldsymbol{v} = R(\boldsymbol{\alpha})$ whose components are

$$v^i = g^{ij}\alpha_j, \quad i = 1, \ldots, n.$$

For this reason, applying R is also called the operation of raising indices.

Remark 3.15. For a general finite-dimensional inner product space $(V, (\ ,\ ))$ with basis $(e_1, \ldots, e_n)$ and representing matrix with entries $g_{ij} = (e_i, e_j)$ as in Remark 3.8, the maps L and R are also described in terms of lowering and raising indices:

$$L(v)_j = v^i g_{ij}, \quad R(\alpha)^i = g^{ij}\alpha_j,$$

with $v = v^i e_i$ and $\alpha = \alpha_j e^j$. Here we have denoted by $(e^1, \ldots, e^n)$ the dual basis on V^* and by g^{ij} the entries of the inverse matrix of $\boldsymbol{g} = (g_{ij})$.

3.3. The norm

Let $(V, (\ ,\ ))$ be an inner product space. Positivity implies—see Remark 3.3—that (v, v) is a nonnegative real number. Therefore, we can compute its square root. We use this to define the norm of a vector v, generalizing (3.2), as

$$\|v\| := \sqrt{(v, v)}.$$

The norm has three important properties. We start considering the first two, which follow immediately from the properties of the inner product:

(N.1) $\|v\| \geq 0$ for all $v \in V$, and $\|v\| = 0$ iff $v = 0$.
(N.2) $\|\alpha v\| = |\alpha|\|v\|$ for all $v \in V$ and for all $\alpha \in \mathbb{R}$.

For the third property—the triangle inequality—we need first the following

Theorem 3.16 (Cauchy–Schwarz inequality). *Let $(V, (\ ,\ ))$ be an inner product space, and let $\|\ \|$ denote the induced norm. Then all $v, w \in V$ satisfy the Cauchy–Schwarz inequality*

$$|(v, w)| \leq \|v\|\|w\| \tag{3.9}$$

with equality saturated iff v and w are linearly dependent.

Proof. We start proving the inequality. If $w = 0$, both sides vanish, so the inequality is satisfied. If $w \neq 0$, we consider the function

$$f(\lambda) := \|v + \lambda w\|^2 = (v + \lambda w, v + \lambda w) = \|v\|^2 + 2\lambda(v, w) + \lambda^2\|w\|^2.$$

The function satisfies the following three properties:

(1) $f(\lambda) \geq 0$ for all λ.
(2) $f'(\lambda) = 2(v, w) + 2\lambda\|w\|^2$.
(3) $f''(\lambda) = 2\|w\|^2 > 0$ for all λ.

Property (2) implies that f has a unique critical point, which by property (3) is a minimum, located at

$$\lambda_{\min} = -\frac{(v, w)}{\|w\|^2}.$$

Property (1) implies that this minimum $f_{\min}$ is nonnegative. Therefore,

$$0 \leq f_{\min} = f(\lambda_{\min}) = \|v\|^2 - \frac{(v, w)^2}{\|w\|^2}.$$

This inequality may be rewritten as $|(v, w)|^2 \leq \|v\|^2\|w\|^2$. Taking the square root yields the Cauchy–Schwarz inequality (3.9).

Next, assume that v and w are linearly dependent. Upon exchanging them if necessary, we have $w = \alpha v$ for some real number α. We then have $(v, w) = \alpha\|v\|^2$, by linearity with respect to the second argument, and $\|w\| = |\alpha|\|v\|$, by property (N.2). This shows that we have an equality in (3.9).

Vice versa, suppose that we have an equality in (3.9). If $w = 0$, the vectors are obviously linearly dependent. If $w \neq 0$, we consider the function f as above. The equality in (3.9) implies $f_{\min} = 0$. This means $\|v + \lambda_{\min} w\| = 0$, so $v + \lambda_{\min} w = 0$ by property (N.1). Therefore, v and w are linearly dependent. □

We then have the following generalization of (3.5):

PROPOSITION 3.17 (The triangle inequality). *Let $(V, (\ ,\ ))$ be an inner product space, and let $\|\ \|$ denote the induced norm. Then all $v, w \in V$ satisfy the* **triangle inequality**

$$\|v + w\| \leq \|v\| + \|w\|. \tag{3.10}$$

PROOF. We have

$$\|v + w\|^2 = (v + w, v + w) = \|v\|^2 + 2(v, w) + \|w\|^2.$$

By taking the absolute value (and using the triangle inequality on $\mathbb{R}$), we get

$$\|v + w\|^2 = |\|v + w\|^2| \leq \|v\|^2 + 2|(v, w)| + \|w\|^2.$$

By the Cauchy–Schwarz inequality, we then have

$$\|v + w\|^2 \leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2.$$

By taking the square root, we get the triangle inequality (3.10). □

As a consequence, we have the following

THEOREM 3.18 (Properties of the norm). *Let $(V, (\ ,\ ))$ be an inner product space. Then the induced norm $\|\ \|$ satisfies the following three properties*

(N.1) $\|v\| \geq 0$ *for all* $v \in V$, *and* $\|v\| = 0$ *iff* $v = 0$.
(N.2) $\|\alpha v\| = |\alpha|\|v\|$ *for all* $v \in V$ *and for all* $\alpha \in \mathbb{R}$.
(N.3) $\|v + w\| \leq \|v\| + \|w\|$ *for all* $v, w \in V$.

DIGRESSION 3.19 (Normed spaces). A **norm** on a real vector space V is a function $\|\ \|\colon V \to \mathbb{R}$ satisfying properties (N.1), (N.2), and (N.3). A real vector space endowed with a norm is called a **normed space**. The above theorem shows that an inner product space is automatically a normed space as well. On the other hand, there are norms that are not defined in terms of an inner product. For example, on $\mathbb{R}^n$ one can show that

$$\|\boldsymbol{v}\|_p := \left(\sum_{i=1}^n |v^i|^p\right)^{\frac{1}{p}}$$

defines a norm for every real number $p \geq 1$. The following is also a norm:

$$\|\boldsymbol{v}\|_\infty := \max\{|v^1|, \ldots, |v^n|\}.$$

REMARK 3.20 (The other triangle inequality). In plane geometry we know that not only is the length of one side of a triangle shorter than the sum of the lengths of the other two sides, but also that it is longer than their difference. Analogously, in a normed space, in addition to the triangle inequality (3.10) we also

$$|\|v\| - \|w\|| \leq \|v - w\| \tag{3.11}$$

for all v and w. To prove this just apply the usual triangle equality to w and $v - w$:

$$\|v\| = \|w + (v - w)\| \leq \|w\| + \|v - w\|.$$

Therefore, $\|v\| - \|w\| \leq \|v - w\|$. Exchanging v and w yields $\|w\| - \|v\| \leq \|v - w\|$. The two inequalities together imply (3.11).

REMARK 3.21 (Angle between vectors). Returning to the Cauchy–Schwarz inequality (3.9), we observe that for nonzero vectors v and w we may also write $\frac{|(v,w)|}{\|v\|\|w\|} \leq 1$. This implies that there is an angle θ, unique in $[0, \pi]$, such that

$$\cos\theta = \frac{(v, w)}{\|v\|\|w\|}.$$

This formula generalizes (3.3). We also have

$$\|v + w\|^2 = \|v\|^2 + 2(v, w) + \|w\|^2 = \|v\|^2 + \|w\|^2 + 2\|v\|\|w\|\cos\theta,$$

which generalizes the law of cosines (3.4).

REMARK 3.22 (Infinite-dimensional spaces). Note that our proofs of the Cauchy–Schwarz and the triangle inequalities also hold in the case of infinite-dimensional spaces. In particular, in the case of Example 3.9 of continuous functions on a compact interval $[a, b]$ with inner product as in (3.8) the induced norm is

$$\|f\| = \left(\int_a^b f^2 \, \mathrm{d}x\right)^{\frac{1}{2}},$$

so the Cauchy–Schwarz inequality explictly reads

$$\left|\int_a^b fg \,\mathrm{d}x\right| \le \left(\int_a^b f^2 \,\mathrm{d}x\right)^{\frac{1}{2}} \left(\int_a^b g^2 \,\mathrm{d}x\right)^{\frac{1}{2}}, \tag{3.12}$$

whereas the triangle inequality reads

$$\left(\int_a^b (f+g)^2 \,\mathrm{d}x\right)^{\frac{1}{2}} \le \left(\int_a^b f^2 \,\mathrm{d}x\right)^{\frac{1}{2}} + \left(\int_a^b g^2 \,\mathrm{d}x\right)^{\frac{1}{2}}. \tag{3.13}$$

On the inner product space $\mathbb{R}^\infty$ of Example 3.11, the norm is

$$\|a\| = \sqrt{\sum_{i=1}^{\infty} (a_i)^2},$$

the Cauchy–Schwarz inequality reads

$$\left|\sum_{i=1}^{\infty} a_i b_i\right| \le \sqrt{\sum_{i=1}^{\infty} (a_i)^2} \sqrt{\sum_{i=1}^{\infty} (b_i)^2}, \tag{3.14}$$

and the triangle inequality is

$$\sqrt{\sum_{i=1}^{\infty} (a_i + b_i)^2} \le \sqrt{\sum_{i=1}^{\infty} (a_i)^2} + \sqrt{\sum_{i=1}^{\infty} (b_i)^2}. \tag{3.15}$$

3.3.1. Square-integrable continuous functions. The inequalities (3.12) and (3.13) allow for the construction of a more interesting inner product space (whose complex version is) important for quantum mechanics, namely, the space of square-integrable continuous functions.

We start with a digression on improper Riemann integrals of continuous functions.[3] Recall that one defines

$$\int_{-\infty}^{\infty} f \,\mathrm{d}x := \lim_{a,b\to+\infty} \int_{-a}^{b} f \,\mathrm{d}x,$$

where f is a continuous function, if the limit on the right hand side exists. For further convenience, we define

$$I_f(a,b) := \int_{-a}^{b} f \,\mathrm{d}x,$$

so $\int_{-\infty}^{\infty} f \,\mathrm{d}x := \lim_{a,b\to+\infty} I_f(a,b)$. Note that I_f is a continuous function of a and b.

[3] What we discuss here has far reaching generalizations to much larger classes of functions via the Lebesgue integral.

The situation is better behaved when $f \geq 0$ because, in this case, I_f is monotonically increasing, so

$$\lim_{a,b\to+\infty} I_f(a,b) = \sup\{I_f(a,b),\ (a,b) \in (\mathbb{R}_{\geq 0})^2\}$$

and the limit exists, although it can be infinite.

We can always reduce to this well-behaved case by introducing

$$f_+ := \frac{|f|+f}{2} \quad \text{and} \quad f_- := \frac{|f|-f}{2}.$$

Note that, for f continuous, also $f_\pm$ are continuous. They are moreover nonnnegative, so the integrals $\int_{-\infty}^{\infty} f_\pm \, \mathrm{d}x$ exist, although they can be infinite. Observe that we have

$$f = f_+ - f_-.$$

If the imporper integrals of $f_\pm$ are finite, then there difference is well-defined (and finite), so the limit defining the improper integral of f converges.

Definition 3.23. A continuos function $f\colon \mathbb{R} \to \mathbb{R}$ is called integrable if the improper integrals $\int_{-\infty}^{\infty} f_\pm \, \mathrm{d}x$ are finite.

By the above consideration, we have

$$\int_{-\infty}^{\infty} f \, \mathrm{d}x = \int_{-\infty}^{\infty} f_+ \, \mathrm{d}x - \int_{-\infty}^{\infty} f_- \, \mathrm{d}x$$

for an integrable continuous function f.

Now observe that

$$|f| = f_+ + f_-,$$

If f is continuous, we then have

$$\int_{-\infty}^{\infty} |f| \, \mathrm{d}x = \int_{-\infty}^{\infty} f_+ \, \mathrm{d}x + \int_{-\infty}^{\infty} f_- \, \mathrm{d}x,$$

where each of the three improper integrals is possibly infinite. Observe however that, if the integral on the left hand side is finite, then both integrals on the right hand side must be finite, since they are both nonnegative.

Definition 3.24. A continuos function $f\colon \mathbb{R} \to \mathbb{R}$ is called absolutely integrable if the improper integral $\int_{-\infty}^{\infty} |f| \, \mathrm{d}x$ is finite.

By the above consideration, we have the

Lemma 3.25. *An absolutely integrable continuous function is also integrable.*

We next move to the main object of interest for us:

DEFINITION 3.26. A continuos function $f\colon \mathbb{R} \to \mathbb{R}$ is called square integrable if the improper integral $\int_{-\infty}^{\infty} f^2 \,\mathrm{d}x$ is finite.

We first have the

LEMMA 3.27. *If f and g are square-integrable continuous functions on $\mathbb{R}$, then so is $f+g$.*

PROOF. We use the triangle inequality (3.13), writing $-a$ instead of a:

$$\left(\int_{-a}^{b} (f+g)^2 \,\mathrm{d}x\right)^{\frac{1}{2}} \leq \left(\int_{-a}^{b} f^2 \,\mathrm{d}x\right)^{\frac{1}{2}} + \left(\int_{-a}^{b} g^2 \,\mathrm{d}x\right)^{\frac{1}{2}}.$$

Since f and g are square-integrable, the two summands on the right hand side have finite limits for $a, b \to +\infty$. This implies that $I_{(f+g)^2}(a,b)$ is bounded from above, so its limit—which exists because $(f+g)^2$ is nonnegative—is also finite. □

Since multipliying a square-integrable continuous function by a real constant clearly yields again a square-integrable continuous function, the lemma has the following

COROLLARY 3.28. *The set $L^{2,0}(\mathbb{R})$ of square-integrable continuous functions on $\mathbb{R}$ is a subspace of the real vector space $C^0(\mathbb{R})$ of continuous functions on $\mathbb{R}$.*

Next, we have the

LEMMA 3.29. *If f and g are square-integrable continuous functions on $\mathbb{R}$, then the product fg is absolutely integrable, and hence integrable.*

PROOF. In this case we use the Cauchy–Schwarz inequality (3.12), writing $-a$ instead of a, applied to the nonnegative continuous functions $|f|$ and $|g|$:

$$\left|\int_{-a}^{b} |f||g| \,\mathrm{d}x\right| \leq \left(\int_{-a}^{b} |f|^2 \,\mathrm{d}x\right)^{\frac{1}{2}} \left(\int_{-a}^{b} |g|^2 \,\mathrm{d}x\right)^{\frac{1}{2}}.$$

Since $|f||g| = |fg|$, $|f|^2 = f^2$, and $|g|^2 = g^2$, we can rewrite it as

$$\int_{-a}^{b} |fg| \,\mathrm{d}x \leq \left(\int_{-a}^{b} f^2 \,\mathrm{d}x\right)^{\frac{1}{2}} \left(\int_{-a}^{b} g^2 \,\mathrm{d}x\right)^{\frac{1}{2}},$$

where we have removed the absolute value around the left hand side which is clearly nonnegative.

Since f and g are square integrable, the two factors on the right hand side have finite limits for $a, b \to +\infty$. This implies that $I_{|fg|}(a,b)$

is bounded from above, so its limit—which exists because $|fg|$ is non-negative—is also finite. □

We can summarize these results in the

THEOREM 3.30. *On the real vector space $L^{2,0}(\mathbb{R})$ of square-integrable continuous functions on $\mathbb{R}$, we have the inner product*

$$(f,g) := \int_{-\infty}^{+\infty} fg \, \mathrm{d}x.$$

PROOF. We have already proved that the integral defining the inner product converges. It is clearly bilinear and symmetric. Positivity is proved as in Example 3.9. □

Finally, observe that the induced norm—called the L^2-norm—is

$$\|f\| = \left(\int_{-\infty}^{+\infty} f^2 \, \mathrm{d}x\right)^{\frac{1}{2}}$$

Since $L^{2,0}(\mathbb{R})$ is an inner product space, we have the Cauchy–Schwarz and triangle inequalities. Explicitly, they say that for square-integrable continuous functions f and g on $\mathbb{R}$, we have

$$\left|\int_{-\infty}^{+\infty} fg \, \mathrm{d}x\right| \leq \left(\int_{-\infty}^{+\infty} f^2 \, \mathrm{d}x\right)^{\frac{1}{2}} \left(\int_{-\infty}^{+\infty} g^2 \, \mathrm{d}x\right)^{\frac{1}{2}},$$

$$\left(\int_{-\infty}^{+\infty} (f+g)^2 \, \mathrm{d}x\right)^{\frac{1}{2}} \leq \left(\int_{-\infty}^{+\infty} f^2 \, \mathrm{d}x\right)^{\frac{1}{2}} + \left(\int_{-\infty}^{+\infty} g^2 \, \mathrm{d}x\right)^{\frac{1}{2}}.$$

REMARK 3.31. The space $C_c^0(\mathbb{R})$ of compactly supported continuous functions of Example 3.10 is a subspace of $L^{2,0}(\mathbb{R})$, and its inner product is the restriction of the inner product on $L^{2,0}(\mathbb{R})$.

3.3.2. Square-summable sequences. Analogously to the space of square-integrable functions, we may study the space of square-summable sequences (this is actually easier, so we leave many details to the reader).

DEFINITION 3.32. A sequence $a = (a_1, a_2, \dots)$ of real numbers is called square summable if the series $\sum_{i=1}^{\infty}(a_i)^2$ converges.

If a is a sequence, we denote by $a^{(N)}$ its N-truncation, i.e., the sequence whose first N terms are the same as in a, whereas the others are equal to zero. Note that, for every N, $a^{(N)}$ belongs to the inner product space $\mathbb{R}^\infty$ of finite real sequences introduced in Example 3.11.

Using the triangle inequality (3.15) on $\mathbb{R}^\infty$ for truncated sequences and taking the limit for N going to infinity, one shows that the sum of two square-summable sequences is again square-summable.

By the Cauchy–Schwarz inequality (3.14) on $\mathbb{R}^\infty$, one shows that, if a and b are square summable, then $\sum_{i=1}^\infty |a_i b_i|$ converges. A fortiori, the right hand side of

$$(a, b) := \sum_{i=1}^{\infty} a_i b_i$$

also converges.

The inner product space of square-summable real sequences is denoted by ℓ^2 (or, more precisely, by $\ell^2_{\mathbb{R}}$ to stress that we are considering real sequences).

3.4. Orthogonality

Let $(V, (\ ,\ ))$ be an inner product space. Two vectors v and w are called orthogonal if $(v, w) = 0$. In this case one writes $v \perp w$.

A collection $(e_i)_{i \in S}$ of <u>nonzero</u> vectors in V is called an orthogonal system if $e_i \perp e_j$ for all $i \neq j$ in S.

Lemma 3.33. *An orhogonal system is linearly independent.*

Proof. Suppose $\sum_i \lambda^i e_i = 0$ for some scalars λ^i (only finitely many of which are different from zero). For every j, we then have

$$\left(e_j, \sum_i \lambda^i e_i\right) = 0.$$

On the other hand, using the linearity of the inner product and the orthogonality of the system, we get[4]

$$\left(e_j, \sum_i \lambda^i e_i\right) = \lambda^j \|e_j\|^2.$$

Since $e_j \neq 0$, we get $\lambda^j = 0$. □

If $(e_i)_{i \in S}$ is in addition a system of generators, then it is called an orthogonal basis.

An orthogonal system $(e_i)_{i \in S}$ is called an orthonormal system if in addition $\|e_i\|^2 = 1$ for all $i \in S$. Succinctly,

$$(e_i, e_j) = \delta_{ij}, \tag{3.16}$$

[4]No sum is understood on the right hand side.

where δ_{ij} is the Kronecker delta:

$$\delta_{ij} := \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

If $(e_i)_{i\in S}$ is in addition a system of generators, then it is called an orthonormal basis.

EXAMPLE 3.34. The standard basis $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ of $\mathbb{R}^n$ is an orthonormal basis for the dot product.

REMARK 3.35. Note that (3.16) implies that the representing matrix of an inner product on a finite-dimensional space in an orthonormal basis is the identity matrix. We will see (Theorem 3.44) that every finite-dimensional space admits an orthonormal basis, so eventually we always go back to the case of the dot product.

An orthogonal system $(e_i)_{i\in S}$ can be transformed into an orthonormal system $(\widetilde{e}_i)_{i\in S}$ with the same span simply by normalizing the vectors: $\widetilde{e}_i := \frac{e_i}{\|e_i\|}$ for all $i \in S$. We will therefore mainly consider orthonormal systems/bases.

If $v = \sum_i v^i e_i$ is in the span of an orthonormal system $(e_i)_{i\in S}$, we can get the coefficients of the expansion by the formula

$$v^i = (e_i, v). \tag{3.17}$$

We can also rewrite the expansion as

$$v = \sum_i (e_i, v) e_i.$$

Moreover, note that, if $w = \sum_i w^i e_i$, then $(v, w) = \sum_{ij} v^i w^j (e_i, e_j)$, so, by (3.16),

$$(v, w) = \sum_i v^i w^i,$$

and, in particular,

$$\|v\|^2 = \sum_i (v^i)^2. \tag{3.18}$$

REMARK 3.36 (Einstein's convention). The above formulae do not fit with Einstein's convention as introduced in Definition 1.33. The way out is to raise and lower indices as in Remark 3.14, now using the three different notations δ_{ij}, δ_i^j, and δ^{ij} for the Kronecker delta (each of them being 1 for $i = j$ and 0 otherwise). We then define, in Einstein's notation, $e^i = \delta^{ij} e_j$ and $v_i = \delta_{ij} v^j$. The formulae above now read

$$\begin{aligned} v &= v^i e_i, \\ v_i &= (e_i, v), \\ v &= (e^i, v) e_i, \\ (v, w) &= v^i \delta_{ij} w^j, \\ \|v\|^2 &= v^i \delta_{ij} v^j. \end{aligned}$$

REMARK 3.37 (Bessel's inequality). The generalization of Pythagoras' theorem given by (3.18) does not hold if we have an orthonormal system $(e_i)_{i\in S}$ that is not a basis. Actually, if S is infinite, then (3.18) may not be a finite sum, as there could be infinitely many nonvanishing coefficients $v_i = (e_i, v)$ for a given vector v.[5] Suppose now that $S = \mathbb{N}_{>0}$. In particular, for each N, $(e_1, \dots, e_N)$ is an orthonormal system, so we have[6]

$$\sum_{i=1}^{N} (v_i)^2 \leq \|v\|^2.$$

Therefore, the limit for $N \to \infty$ converges, and we have

$$\sum_{i=1}^{\infty} (v_i)^2 \leq \|v\|^2, \tag{3.19}$$

which is known as Bessel's inequality.

EXAMPLE 3.38 (Sine series). Consider the space

$$V := \{\phi \in C^0([0, L]) \mid \phi(0) = \phi(L) = 0\}$$

which extends that of Section 2.3.2 to all continuous functions. Let

$$e_k(x) := \sqrt{\frac{2}{L}} \sin\left(\frac{\pi k x}{L}\right).$$

---

[5] If $(e_i)_{i\in S}$ is a basis, the v^is are the coefficients of the expansion of v, so only finitely many of them are different from zero, by definition of a basis.

[6] Observe that

$$0 \leq \left\| v - \sum_{i=1}^{N} v^i e_i \right\|^2 = \left(v - \sum_{i=1}^{N} v^i e_i, v - \sum_{i=1}^{N} v^i e_i \right) = \\ = \|v\|^2 - 2 \sum_{i=1}^{N} (v_i)^2 + \sum_{i=1}^{N} (v_i)^2 = \|v\|^2 - \sum_{i=1}^{N} (v_i)^2.$$

One can easily verify that $(e_k)_{k\in\mathbb{N}_{>0}}$ is an orthonormal system on V with the inner product $(f,g) := \int_0^L fg \,\mathrm{d}x$ of Example 3.9. If f is a linear combination of the e_ks,

$$f(x) = \sqrt{\frac{2}{L}}\sum_{k=1}^{\infty} b_k \sin\left(\frac{\pi k x}{L}\right) \tag{3.20}$$

with only finitely many b_ks different from zero, then we can recover the coefficients b_k of the expansion via (3.17):[7]

$$b_k = (e_k, f) = \sqrt{\frac{2}{L}}\int_0^L f(x)\sin\left(\frac{\pi k x}{L}\right)\,\mathrm{d}x. \tag{3.21}$$

Note that the integral on the right hand side actually converges for every f in V, so we can define the coefficients b_k also for functions that are not finite linear combinations of the e_ks. Bessel's inequality (4.10) in this case reads

$$\sum_{k=1}^{\infty}(b_k)^2 \leq \|f\|^2. \tag{3.22}$$

In this particular case one can show—but this is beyond the scope of these notes—that Bessel's inequality is actually saturated:

$$\sum_{k=1}^{\infty}(b_k)^2 = \|f\|^2. \tag{3.23}$$

This equality is known as Parseval's identity. Finally, it turns out that also the series (3.20) actually converges, in an appropriate sense, to the original function f. This is an example of a Fourier series, called a sine series.

3.4.1. The orthogonal projection. Let w be a nonzero vector in V. Any vector v can then be decomposed in a component parallel to w,

$$v_\parallel = (v,w)\frac{w}{\|w\|^2},$$

[7] One often prefers to write the expansion as

$$f(x) = \sum_{k=1}^{\infty} b_k \sin\left(\frac{\pi k x}{L}\right)$$

without the prefactor, i.e., using a nonnormalized orthogonal system. In this case, the coefficients b_k are obtained by

$$b_k = \frac{2}{L}\int_0^L f(x)\sin\left(\frac{\pi k x}{L}\right)\,\mathrm{d}x.$$

and in one orthogonal to it,

$$v_\perp = v - v_\parallel.$$

In fact, $(v_\parallel, w) = (v, w)$, so $(v_\perp, w) = 0$. This is an example of an orthogonal decomposition.

EXAMPLE 3.39. Consider $\mathbb{R}^2$ with the dot product. Let $\boldsymbol{v} = \left(\begin{smallmatrix} 2.5 \\ 2.5 \end{smallmatrix}\right)$ and $\boldsymbol{w} = \left(\begin{smallmatrix} 6 \\ 2 \end{smallmatrix}\right)$ as in Figure 3.3. We then have $\|\boldsymbol{w}\|^2 = 40$ and $\boldsymbol{v}\cdot\boldsymbol{w} = 20$, so $\boldsymbol{v}_\parallel = \frac{1}{2}\boldsymbol{w} = \left(\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right)$ and $\boldsymbol{v}_\perp = \left(\begin{smallmatrix} 2.5 \\ 2.5 \end{smallmatrix}\right) - \left(\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -.5 \\ 1.5 \end{smallmatrix}\right)$.

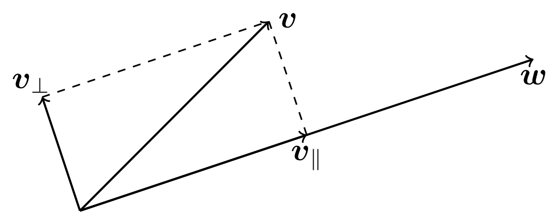


FIGURE 3.3. The orthogonal decomposition

EXAMPLE 3.40. On $\mathbb{R}^n$ with the dot product, the formula for the parallel component of $\boldsymbol{v}$ along a unit vector $\boldsymbol{w}$ reads

$$\boldsymbol{v}_\parallel = (\boldsymbol{w}\cdot\boldsymbol{v})\,\boldsymbol{w}.$$

On $\mathbb{R}^3$ one can write the orthogonal component also using the cross product (see exercise 3.16) as

$$\boldsymbol{v}_\perp = (\boldsymbol{w}\times\boldsymbol{v})\cdot\boldsymbol{w}.$$

Therefore, for $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^3$ with $\|\boldsymbol{w}\| = 1$, the decomposition reads

$$\boldsymbol{v} = (\boldsymbol{w}\cdot\boldsymbol{v})\,\boldsymbol{w} + (\boldsymbol{w}\times\boldsymbol{v})\cdot\boldsymbol{w}.$$

We denote by P_w the endomorphism that assigns to a vector v its component $v_\parallel$ parallel to w. We rewrite the definition for future use:

$$P_w v = (v, w)\frac{w}{\|w\|^2}. \tag{3.24}$$

Note that, for any scalar λ, we have

$$P_w(\lambda w) = (\lambda w, w)\frac{w}{\|w\|^2} = \lambda w.$$

Therefore, P_w restricted to the span of w acts as the identity. This also means that $P_w^2 = P_w$, so P_w is a projection.

As a consequence, the endomorphism $P'_w = \mathrm{Id} - P_w$ is also a projection. Its image, which we denote by $w^\perp$, is called the orthogonal complement to the span of w. Explicitly, we have

$$w^\perp = \{v \in V \mid v \perp w\}. \tag{3.25}$$

PROOF. For any v, we have seen that $P'_w v = v_\perp$ is orthogonal to w, so it belongs to $w^\perp$. This shows $\operatorname{im} P'_w \subseteq w^\perp$.

On the other hand, if $v \in w^\perp$, we get $P_w v = 0$, so we have $P'_w v = v$, which shows that v is in the image of P'_w. Therefore, $w^\perp \subseteq \operatorname{im} P'_w$. □

EXAMPLE 3.41. Let $V = C^0([0,1])$ be the space of continuous functions on the interval $[0,1]$ with inner product $(f,g) = \int_0^1 fg \, \mathrm{d}x$ (see Example 3.9). Let $f(x) = x^3$ and $g(x) = x$. We have $\|g\|^2 = \int_0^1 x^2 \, \mathrm{d}x = \frac{1}{3}$ and $(f,g) = \int_0^1 x^4 \, \mathrm{d}x = \frac{1}{5}$, so $P_g f = \frac{3}{5} g$. Therefore, we get

$$(P_g f)(x) = \frac{3}{5} x \quad \text{and} \quad (P'_g f)(x) = x^3 - \frac{3}{5} x.$$

REMARK 3.42. Note that P_w and P'_w only depend on the direction of w but not on its norm:

$$P_{\lambda w} = P_w \quad \text{and} \quad P'_{\lambda w} = P'_w$$

for every $\lambda \neq 0$. In particular, if w is normalized (i.e., $\|w\| = 1$), then we get the simpler formula

$$P_w v = (v,w)w.$$

3.4.2. The Gram–Schmidt process. We want to show that a finite (or countable) linearly independent system may be transformed into an orthonormal one.

The central step of the contruction is the following: Suppose we have an orthonormal system $(e_1, \dots, e_k)$ and a vector v that is not a linear combination of the e_is. Then we can make v orhogonal to them by substracting all its parallel components:

$$\widetilde{v} := v - \sum_{i=1}^{k} (v, e_i) e_i.$$

Note that $(\widetilde{v}, e_i) = 0$ for $i = 1, \dots, k$. Moreover, $\widetilde{v} \neq 0$ because v is not a linear combination of the e_is. We can then normalize it to

$$e_{k+1} := \frac{\widetilde{v}}{\|\widetilde{v}\|}.$$

As a result $(e_1, \dots, e_{k+1})$ is also an orthonormal system, and in particular it is linearly independent by Lemma 3.33. Also note that

$$\operatorname{Span}\{e_1, \dots, e_{k+1}\} = \operatorname{Span}\{e_1, \dots, e_k, v\}.$$

In fact, let w be a linear combination of $(e_1, \dots, e_{k+1})$. Since e_{k+1} is a linear combination of $(e_1, \dots, e_k, v)$, we get that w is so, too. Vice versa, let w be a linear combination of $(e_1, \dots, e_k, v)$. Since v is a linear combination of $(e_1, \dots, e_{k+1})$, we get that w is so, too.

We then get the following

Proposition 3.43 (Gram–Schmidt process). *Let $(v_1, \dots, v_k)$ be a linearly independent system in an inner product space V. Then there is an orthonormal system $(e_1, \dots, e_k)$ with the same span. This system is determined by the following process:*

$$\begin{aligned}
\widetilde{v}_1 &:= v_1 & e_1 &:= \frac{\widetilde{v}_1}{\|\widetilde{v}_1\|}\\
\widetilde{v}_2 &:= v_2 - (v_2, e_1)e_1 & e_2 &:= \frac{\widetilde{v}_2}{\|\widetilde{v}_2\|}\\
\widetilde{v}_3 &:= v_3 - (v_3, e_1)e_1 - (v_3, e_2)e_2 & e_3 &:= \frac{\widetilde{v}_3}{\|\widetilde{v}_3\|}\\
&\vdots & &\vdots\\
\widetilde{v}_k &:= v_k - \sum_{i=1}^{k-1}(v_k, e_i)e_i & e_k &:= \frac{\widetilde{v}_k}{\|\widetilde{v}_k\|}
\end{aligned}$$

Proof. The proof goes by induction on the number k of vectors. For $k = 1$, we simply normalize the vector: $e_1 := \frac{v_1}{\|v_1\|}$.

Next suppose we have proved the statement for k vectors and we want to prove it for $k + 1$. Consider the subcollection $(v_1, \dots, v_k)$ of $(v_1, \dots, v_{k+1})$. By the induction assumption, we can replace it by the orthonormal system $(e_1, \dots, e_k)$, given by the construction in the proposition, which has the same span. In particular, v_{k+1} is not a linear combination of the e_is, $i = 1, \dots, k$. We can therefore apply the construction just before the proposition to get the vector e_{k+1}. As we observed, $(e_1, \dots, e_{k+1})$ is an orthonormal system with the same span as $(e_1, \dots, e_k, v_{k+1})$. As the latter has the same the span as $(v_1, \dots, v_{k+1})$, the proof is complete.

The displayed list of assignments summarizes this process. □

If $(v_1, \dots, v_n)$ is a basis of V, then the Gram–Schmidt process yields an orthonormal basis $(e_1, \dots, e_n)$. Therefore, we have the

Theorem 3.44 (Orthonormal bases). *A finite-dimensional or countably infinite-dimensional inner product space has an orthonormal basis.*

Example 3.45 (A countably infinite-dimensional example). Consider the inner product space $V = C^0([0, 1])$ of Example 3.9. Let W be the subspace of polynomial functions. This has the basis $(1, x, x^2, x^3, \dots)$. We write $v_k(x) = x^k$. The Gram–Schmidt process will then turn $(v_k)_{k\in\mathbb{N}}$ into an orthonormal basis $(e_k)_{k\in\mathbb{N}}$ of W. Let us see the first steps. We have $\|v_0\|^2 = \int_0^1 1\,\mathrm{d}x = 1$, so $e_0 = v_0$. Next we compute

$(v_1, e_0) = \int_0^1 x\,\mathrm{d}x = \frac{1}{2}$, so $\widetilde{v}_1(x) = x-\frac{1}{2}$. From $\|\widetilde{v}_1\|^2 = \int_0^1 \left(x-\frac{1}{2}\right)^2 \mathrm{d}x = \frac{1}{12}$, we get $e_1(x) = 2\sqrt{3}\left(x-\frac{1}{2}\right)$. Next we compute $(v_2, e_0) = \int_0^1 x^2\,\mathrm{d}x = \frac{1}{3}$ and $(v_2, e_1) = 2\sqrt{3}\int_0^1 x^2\left(x-\frac{1}{2}\right)\mathrm{d}x = \frac{1}{2\sqrt{3}}$, so $\widetilde{v}_2(x) = x^2 - \frac{1}{3} - \left(x-\frac{1}{2}\right) = x^2 - x + \frac{1}{6}$. From $\|\widetilde{v}_2\|^2 = \int_0^1 \left(x^2 - x + \frac{1}{6}\right)^2 \mathrm{d}x = \frac{1}{180}$, we get $e_2(x) = 6\sqrt{5}\left(x^2 - x + \frac{1}{6}\right)$. We then have the beginning of the orthonormal basis

$$\left(1, 2\sqrt{3}\left(x-\frac{1}{2}\right), 6\sqrt{5}\left(x^2 - x + \frac{1}{6}\right), \dots\right)$$

of W.

Example 3.46 (Hermite polynomials). This is another countably infinite-dimensional example, which is relevant for the quantum harmonic oscillator. Let V be the space of real polynomial functions (you may regard it as a subspace of $C^0(\mathbb{R})$). One can show that

$$(f, g) := \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2} f g\,\mathrm{d}x$$

defines an inner product on V. The basis $(1, x, x^2, x^3, \dots)$ of V can then be turned by the Gram–Schmidt process into an orthonormal basis $(e_0, e_1, \dots)$. Its elements (up to appropriate factors) are the Hermite polynomials H_n:

$$e_n = \frac{1}{\sqrt{\sqrt{\pi}2^n n!}} H_n, \qquad H_n := (-1)^n \mathrm{e}^{x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \mathrm{e}^{-x^2}.$$

As we have already observed, see equation (3.16) and Remark 3.35, the matrix representing the inner product in an orthonormal basis is the identity matrix. In particular, we get the following

Corollary 3.47. *A symmetric matrix* $\boldsymbol{g}$ *is positive definite iff it is of the form* $\boldsymbol{g} = \boldsymbol{E}^\mathsf{T}\boldsymbol{E}$ *with* $\boldsymbol{E}$ *an invertible matrix.*

Proof. Let $(\boldsymbol{v}, \boldsymbol{w}) := \boldsymbol{v}^\mathsf{T}\boldsymbol{g}\boldsymbol{w}$.

If $\boldsymbol{g}$ is symmetric and positive definite, then $(\ ,\ )$ is an inner product. By the Gram–Schmidt process, we have an orthonormal basis. In such a basis, see Remark 3.35, the inner product is represented by the identity matrix. Taking $\boldsymbol{E}$ to be the matrix representing the change of basis, we then get $\boldsymbol{g} = \boldsymbol{E}^\mathsf{T}\boldsymbol{E}$.

Vice versa, if $\boldsymbol{g} = \boldsymbol{E}^\mathsf{T}\boldsymbol{E}$, we get that $\boldsymbol{g}^\mathsf{T} = \boldsymbol{g}$. Moreover, for every $\boldsymbol{v}$ we have

$$(\boldsymbol{v}, \boldsymbol{v}) = \boldsymbol{v}^\mathsf{T}\boldsymbol{g}\boldsymbol{v} = \boldsymbol{v}^\mathsf{T}\boldsymbol{E}^\mathsf{T}\boldsymbol{E}\boldsymbol{v} = (\boldsymbol{E}\boldsymbol{v})^\mathsf{T}\boldsymbol{E}\boldsymbol{v} = (\boldsymbol{E}\boldsymbol{v})^\mathsf{T} \cdot \boldsymbol{E}\boldsymbol{v}.$$

If $\boldsymbol{v}$ is different from zero, then so is also $\boldsymbol{E}\boldsymbol{v}$, since $\boldsymbol{E}$ invertible. By the positivity of the dot product, we then get the positivity of $(\, , \,)$,

$$\boldsymbol{v} \neq 0 \implies \boldsymbol{E}\boldsymbol{v} \neq 0 \implies (\boldsymbol{E}\boldsymbol{v})^\mathsf{T} \cdot \boldsymbol{E}\boldsymbol{v} > 0 \implies (\boldsymbol{v}, \boldsymbol{v}) > 0,$$

and hence of $\boldsymbol{g}$. □

REMARK 3.48. As is made clear in the proof, a matrix $\boldsymbol{E}$ such that $\boldsymbol{g} = \boldsymbol{E}^\mathsf{T}\boldsymbol{E}$ is obtained via the change of basis to an orthonormal basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ for the inner product $(\boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{v}^\mathsf{T}\boldsymbol{g}\boldsymbol{w}$. In particular, if we denote by $\boldsymbol{F}$ the matrix whose columns are these basis vectors, we get

$$\boldsymbol{F}^\mathsf{T}\boldsymbol{g}\boldsymbol{F} = \begin{pmatrix} \boldsymbol{v}_1^\mathsf{T} \\ \vdots \\ \boldsymbol{v}_n^\mathsf{T} \end{pmatrix} \boldsymbol{g} \left(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\right) = \begin{pmatrix} \boldsymbol{v}_1^\mathsf{T} \\ \vdots \\ \boldsymbol{v}_n^\mathsf{T} \end{pmatrix} \left(\boldsymbol{g}\boldsymbol{v}_1, \dots, \boldsymbol{g}\boldsymbol{v}_n\right) = \mathbf{1},$$

so $\boldsymbol{g} = \boldsymbol{F}^{\mathsf{T},-1}\boldsymbol{F}^{-1}$. Therefore, we get the factorization $\boldsymbol{g} = \boldsymbol{E}^\mathsf{T}\boldsymbol{E}$ with $\boldsymbol{E} = \boldsymbol{F}^{-1}$. As $\boldsymbol{F}$ consists of basis elements, it is also called a frame. For this reason, its inverse $\boldsymbol{E}$ is called a coframe.

REMARK 3.49. The matrix $\boldsymbol{E}$ in Corollary 3.47 is not uniquely determined (as we can choose different orthonormal bases). In particular, suppose that $\boldsymbol{E}'$ is also an invertible matrix with $\boldsymbol{g} = \boldsymbol{E}'^\mathsf{T}\boldsymbol{E}'$. Then we have $\boldsymbol{E}'^\mathsf{T}\boldsymbol{E}' = \boldsymbol{E}^\mathsf{T}\boldsymbol{E}$, or, equivalently, $\boldsymbol{E}'\boldsymbol{E}^{-1} = \boldsymbol{E}'^{\mathsf{T},-1}\boldsymbol{E}^\mathsf{T}$. Since $\boldsymbol{E}'^{\mathsf{T},-1}\boldsymbol{E}^\mathsf{T} = (\boldsymbol{E}'\boldsymbol{E}^{-1})^{\mathsf{T},-1}$, we get that the invertible matrix $\boldsymbol{O} := \boldsymbol{E}'\boldsymbol{E}^{-1}$ satisfies

$$\boldsymbol{O}^\mathsf{T} = \boldsymbol{O}^{-1}.$$

A matrix with this property is called an orthogonal matrix (see more on this in Section 3.5). Note that $\boldsymbol{E}' = \boldsymbol{O}\boldsymbol{E}$. In conclusion, any two matrices occurring in a factorization of the same positive definite matrix are related by an orthogonal matrix.

The fact that a positive definite matrix $\boldsymbol{g}$ is necessarily of the form $\boldsymbol{g} = \boldsymbol{E}^\mathsf{T}\boldsymbol{E}$ for some invertible matrix $\boldsymbol{E}$ implies that $\det \boldsymbol{g} = (\det \boldsymbol{E})^2 > 0$, so we have the

LEMMA 3.50. *A positive definite matrix necessarily has positive determinant.*

This fact can actually be improved to a useful criterion to check whether a matrix is positive definite. We need the following terminology.

DEFINITION 3.51. Let $\boldsymbol{g}$ be an $n \times n$ matrix. For every $k = 1, \dots, n$, we denote by $\boldsymbol{g}_{(k)}$ the $k \times k$ upper left part of $\boldsymbol{g}$.[8] The determinant of $\boldsymbol{g}_{(k)}$ is called the kth leading principal minor of $\boldsymbol{g}$.

[8] That is, if $\boldsymbol{g} = (g_{ij})_{i,j=1,\dots,n}$, then $\boldsymbol{g}_{(k)} = (g_{ij})_{i,j=1,\dots,k}$.

Then we have the

COROLLARY 3.52. *If $\boldsymbol{g}$ is a positive definite matrix, then all its leading principal minors are necessarily positive.*

PROOF. Let W_k be the span of $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_k)$. The restriction to W_k of the inner product defined by $\boldsymbol{g}$ is also an inner product. Moreover, for any $\boldsymbol{v}, \boldsymbol{w} \in W_k$, we have $(\boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{v}^\mathsf{T} \boldsymbol{g} \boldsymbol{w} = \boldsymbol{v}^\mathsf{T} \boldsymbol{g}_{(k)} \boldsymbol{w}$. This shows that $\boldsymbol{g}_{(k)}$ is a positive definite matrix, so $\det \boldsymbol{g}_{(k)} > 0$ by Lemma 3.50. □

The converse to Corollary 3.52 also holds. (For the proof, see exercise 3.12.)

LEMMA 3.53. *If all the leading principal minors of a real symmetric matrix $\boldsymbol{g}$ are positive, then $\boldsymbol{g}$ is positive definite.*

We can summarize the results of Corollary 3.52 and of Lemma 3.53 as the following

THEOREM 3.54 (Sylvester's criterion). *A real symmetric matrix is positive definite iff all its leading principal minors are positive.*

3.4.3. Orthogonal complements. The following is a very useful generalization of the concept of the orthogonal complement $w^\perp$ of a nonzero vector w introduced in (3.25).

DEFINITION 3.55. Let W be a subspace of an inner product space V. The orthogonal subspace associated to W is the subspace

$$W^\perp := \{v \in V \mid v \perp w \ \forall w \in W\}$$

of all vectors orthogonal to vectors in W.

If $W = \mathbb{R}w$, for a nonzero vector w, then $W^\perp$ is the same as $w^\perp$ as introduced in (3.25).

EXAMPLE 3.56. Let $V = \mathbb{R}^3$ endowed with the dot product. Let $\boldsymbol{w} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and $W = \mathbb{R}\boldsymbol{w}$ the x-axis. Then $W^\perp$ is the yz-plane. Another example is the following: Let $\boldsymbol{v}, \boldsymbol{w}$ be linearly independent vectors in $\mathbb{R}^3$ and let W be the plane they generate. Then $W^\perp$ is the line through the origin orthogonal to this plane. We can write

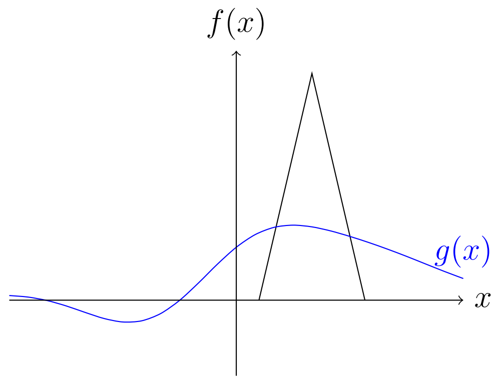
$$W^\perp = \mathbb{R}(\boldsymbol{v} \times \boldsymbol{w}),$$

where we have used the cross product.

EXAMPLE 3.57. Let $V = C^0([-1, 1])$ be the space of continuous functions on the interval $[-1, 1]$ with inner product $(f, g) = \int_{-1}^1 fg \, \mathrm{d}x$ (see Example 3.9). Let

$$W := \{f \in C^0([-1, 1]) \mid f(0) = 0\}.$$

We claim that $W^\perp = \{0\}$. To show this, assume that g is not the zero function. We will show that we can find an f in W such that $(f, g) \neq 0$. In fact, let $x_0 \in [-1, 1]\backslash\{0\}$ be a point with $g(x_0) \neq 0$.[9] We may assume $x_0 > 0$ and $g(x_0) > 0$ (the proof for the other cases is analogous). By continuity, there is an $\epsilon > 0$ such that $(x_0 - \epsilon, x_0 + \epsilon) \subset (0, 1)$ and $g(x) > 0$ for all $x \in (x_0 - \epsilon, x_0 + \epsilon)$. Consider



$$f(x) := \begin{cases} 0 & \text{for } x < x_0 - \epsilon, \\ 1 + \frac{x - x_0}{\epsilon} & \text{for } x_0 - \epsilon \leq x \leq x_0, \\ 1 - \frac{x - x_0}{\epsilon} & \text{for } x_0 \leq x \leq x_0 + \epsilon, \\ 0 & \text{for } x > x_0 + \epsilon. \end{cases}$$

We have $f \in W$ and

$$(f, g) = \int_{x_0 - \epsilon}^{x_0 + \epsilon} fg \, \mathrm{d}x$$

because f vanishes outside of the interval $(x_0 - \epsilon, x_0 + \epsilon)$. The integral is strictly positive because both f and g are strictly positive in the interval $(x_0 - \epsilon, x_0 + \epsilon)$.

Without any assumption on dimensionality, we have the following

PROPOSITION 3.58. *Let V be an inner product space, and let W and Z be subspaces of V. Then the following hold:*

(1) $\{0\}^\perp = V$.
(2) $V^\perp = \{0\}$.
(3) $W \cap W^\perp = \{0\}$.
(4) $W \subseteq Z \implies Z^\perp \subseteq W^\perp$.
(5) $W \subseteq W^{\perp\perp}$.

[9] If $g(x) = 0$ for all $x \neq 0$, then, by continuity, we also have $g(0) = 0$, so g would be the zero function.

(6) $W^{\perp\perp\perp} = W^{\perp}$.

PROOF.

(1) Every vector is orthogonal to the zero vector.
(2) If $v \in V^{\perp}$, then $(v, w) = 0$ for every $w \in V$. In particular, we may take $w = v$, so $(v, v) = 0$, which implies $v = 0$.
(3) If $v \in W^{\perp}$, then $(v, w) = 0$ for every $w \in W$. If $v \in W$, then, in particular, we may take $w = v$, so $(v, v) = 0$, which implies $v = 0$.
(4) If $v \in Z^{\perp}$, then $(v, z) = 0$ for every $z \in Z$. In particular, $(v, z) = 0$ for every $z \in W$, so $v \in W^{\perp}$.
(5) Let $w \in W$. Then, for every $v \in W^{\perp}$, we have $(v, w) = 0$. But this means that w is orthogonal to every vector in $W^{\perp}$, so $w \in W^{\perp\perp}$.
(6) Applying (4) to (5), we get $W^{\perp\perp\perp} \subseteq W^{\perp}$. On the other hand, (5) for $W^{\perp}$ reads $W^{\perp} \subseteq (W^{\perp})^{\perp\perp} = W^{\perp\perp\perp}$.

□

Property (3) of the above proposition shows that the sum of W and $W^{\perp}$ is a direct sum. If this happens to exhaust V, i.e., if

$$W \oplus W^{\perp} = V,$$

then we say that $W^{\perp}$ is the orthogonal complement of W. Note that in this case we have

$$W^{\perp\perp} = W.$$

In fact, by (3) applied to $W^{\perp}$, we have $W^{\perp} \cap W^{\perp\perp} = \{0\}$, which then implies $W^{\perp\perp} \subseteq W$. From (5) we have the other inclusion.

The orthogonal space $W^{\perp}$ might happen not to be a complement to W, as Example 3.57 shows. However, we have the following

PROPOSITION 3.59. *If W is a finite-dimensional subspace of an inner product space V, then $W^{\perp}$ is a complement of W, called the orthogonal complement:*

$$V = W \oplus W^{\perp}.$$

To stress that this is a direct sum of orthogonal spaces, we may also use the notation

$$V = W \bigoplus W^{\perp}.$$

PROOF. We prove this by generalizing the construction in Section 3.4.1.

Namely, let $(w_1, \ldots, w_k)$ be an orthonormal basis of W (which exists by Theorem 3.44). Then we define $P_W \in \operatorname{End}(V)$ as

$$P_W(v) := \sum_{i=1}^{k} (v, w_i) w_i. \tag{3.26}$$

Note that $P_W(w_j) = w_j$ for all j, so P_W is a projection (i.e., $P_W^2 = P_W$) with image W. Also note that for every $v, z \in V$, we have

$$(P_W(v), z) = \sum_{i=1}^{k} (v, w_i)(z, w_i),$$

so

$$(P_W(v), z) = (v, P_W(z))$$

for all $v, z \in V$. In particular, we have

$$(P_W(v), w) = (v, w) \quad \forall v \in V,\ \forall w \in W. \tag{3.27}$$

We then define $P'_W := \operatorname{Id} - P_W$, i.e.,

$$P'_W(v) = v - \sum_{i=1}^{k} (v, w_i) w_i.$$

This is also a projection (see Remark 1.38). We claim that its image is $W^\perp$. Clearly, if $v \in W^\perp$, then $P'_W(v) = v$, so $W^\perp \subseteq \operatorname{im} P'_W$. On the other hand, for every $v \in V$ and every $w \in W$, by (3.27) we have

$$(P'_W(v), w) = (v, w) - (P_W(v), w) = 0,$$

so $\operatorname{im} P'_W \subseteq W^\perp$.

Finally, observe that every $v \in V$ can be written as

$$v = P_W(v) + P'_W(v).$$

Since the first summand is in W, the second is in $W^\perp$, and, by (3) in Proposition 3.58, $W \cap W^\perp = \{0\}$, the proof is complete. □

Remark 3.60. Note that, since the decomposition of a vector in a direct sum is unique, the projections P_W and $P_{W^\perp}$ are canonically defined. That is, formula (3.26) is just a convenient way to write P_W when we are given an orthonormal basis $(w_1, \ldots, w_k)$, but it is independent of its choice.

In particular, Proposition 3.59 implies the following

Theorem 3.61. *Let V be a finite-dimensional inner product space. Then, for every subspace W we have the* orthogonal complement *$W^\perp$. In particular,*

$$\dim W + \dim W^\perp = \dim V,$$

and

$$W^{\perp\perp} = W.$$

We may generalize the orthogonal decomposition $W \oplus W^{\perp}$, when $W^{\perp}$ is a complement, to more general "orthogonal sums." First, we need the

DEFINITION 3.62. Two subspaces W_1 and W_2 of an inner product space V are called orthogonal, and we write

$$W_1 \perp W_2,$$

if every vector in W_1 is orthogonal to every vector in W_2.

For example, W and $W^{\perp}$ are orthogonal subspaces. Note that $W_1 \perp W_2$ implies $W_1 \cap W_2 = \{0\}$, as a vector in the intersection is orthogonal to itself, so it has to be zero. In particular, a sum of two orthogonal spaces W_1 and W_2 is automatically a direct sum, which we may also denote as $W_1 \bigoplus\!\!\!\!\perp W_2$. This generalizes to collections:

DEFINITION 3.63. Let $(W_i)_{i\in S}$ be a collection of subspaces of an inner product space V. The collection is called orthogonal if

$$W_i \perp W_j \text{ for all } i \neq j.$$

PROPOSITION 3.64. *If $(W_i)_{i\in S}$ is an orthogonal collection of subspaces, then the sum of the W_is is direct.*

REMARK 3.65 (Orthogonal sums). To stress that the summands of such a direct sum are othogonal to each other, we may also use the notation

$$\bigoplus\!\!\!\!\perp_{i\in S} W_i.$$

PROOF. Suppose we have $\sum_i w_i = 0$, $w_i \in W_i$, and only finitely many w_is different from zero. Taking the inner product with w_j yields

$$0 = \left(w_j, \sum_i w_i \right) = \|w_j\|^2,$$

so $w_j = 0$. As we can do this for every j, we get that the zero vector, and hence every vector, has a unique decomposition, so the sum is direct. □

DEFINITION 3.66. If $(W_i)_{i\in S}$ is an orthogonal collection of subspaces of V and their sum is the whole of V, then

$$V = \bigoplus\!\!\!\!\perp_{i\in S} W_i$$

is called an orthogonal decomposition of V.

REMARK 3.67. Suppose we have an orthogonal decomposition $V = \bigoplus_{i\in S} W_i$. Let P_i denote the projection to the W_i-component. Then, for any $v, v' \in V$, we have

$$(P_i v, v') = (P_i v, P_i v').$$

In fact, since $P_i v \in W_i$, it is orthogonal to every vector in the other W_js, so it only sees the W_i-component of v'. Similarly, we have $(v, P_i v') = (P_i v, P_i v')$. Therefore,

$$(P_i v, v') = (v, P_i v')$$

for all $v, v' \in V$. This is an example of a symmetric operator (more on this in Section 3.5.4).

3.5. Orthogonal operators

An endomorphism F of an inner product space V is called an orthogonal operator if

$$(Fv, Fw) = (v, w)$$

for every $v, w \in V$.

EXAMPLE 3.68 (Orthogonal matrices). In the case of the dot product on $\mathbb{R}^n$, the endomorphism defined by an $n \times n$ matrix $\boldsymbol{A}$ is orthogonal iff $\boldsymbol{v} \cdot \boldsymbol{w} = (\boldsymbol{A}\boldsymbol{v}) \cdot (\boldsymbol{A}\boldsymbol{w}) = \boldsymbol{v}^\mathsf{T}\boldsymbol{A}^\mathsf{T}\boldsymbol{A}\boldsymbol{w}$ for all $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$. Taking $\boldsymbol{v} = \boldsymbol{e}_i$ and $\boldsymbol{w} = \boldsymbol{e}_j$, the condition implies $(\boldsymbol{A}^\mathsf{T}\boldsymbol{A})_{ij} = \delta_{ij}$ for all $i, j = 1, \dots, n$. Therefore, we see that the endomorphism defined by $\boldsymbol{A}$ is orthogonal iff

$$\boldsymbol{A}^\mathsf{T}\boldsymbol{A} = \mathbf{1},$$

which is the usual definition of an orthogonal matrix. Note that the condition implies that $\boldsymbol{A}$ is invertible and also that $\boldsymbol{A}^{-1}$ is itself an orthogonal matrix. Moreover, if $\boldsymbol{A}$ and $\boldsymbol{B}$ are orthogonal, then so is their product $\boldsymbol{A}\boldsymbol{B}$.

REMARK 3.69. In particular, we have that an endomorphism of a finite-dimensional inner product space is orthogonal iff its representing matrix in any orthonormal basis is orthogonal.

REMARK 3.70. Let $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ be the columns of an $n \times n$ matrix $\boldsymbol{A}$. Then $(\boldsymbol{v}_1^\mathsf{T}, \dots, \boldsymbol{v}_n^\mathsf{T})$ are the rows of $\boldsymbol{A}^\mathsf{T}$. We then see that

$$\boldsymbol{A} \text{ is orthogonal} \iff (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n) \text{ is an orthonormal system.}$$

In particular, the matrix representing the change of basis from the standard basis to another orthonormal basis (w.r.t. to the dot product) is orthogonal. By the observations at the end of Remark 3.68, we then get the

PROPOSITION 3.71. *The matrix representing the change of basis between any two orthonormal bases on* $(\mathbb{R}^n, \cdot)$ *is orthogonal.*

REMARK 3.72. If we consider on $\mathbb{R}^n$ the inner product $(\boldsymbol{v}, \boldsymbol{w}) := \boldsymbol{v}^\mathsf{T} \boldsymbol{g} \boldsymbol{w}$, where $\boldsymbol{g}$ is a positive definite symmetric matrix as in Example 3.7, the condition for the endomorphisms defined by an $n \times n$ matrix $\boldsymbol{A}$ to be orthogonal is

$$\boldsymbol{A}^\mathsf{T} \boldsymbol{g} \boldsymbol{A} = \boldsymbol{g}.$$

If we write $\boldsymbol{g} = \boldsymbol{E}^\mathsf{T} \boldsymbol{E}$ as in Corollary 3.47, the condition becomes $\boldsymbol{A}^\mathsf{T} \boldsymbol{E}^\mathsf{T} \boldsymbol{E} \boldsymbol{A} = \boldsymbol{E}^\mathsf{T} \boldsymbol{E}$, i.e.,

$$\boldsymbol{B}^\mathsf{T} \boldsymbol{B} = \mathbf{1} \quad \text{with} \quad \boldsymbol{B} = \boldsymbol{E} \boldsymbol{A} \boldsymbol{E}^{-1}.$$

That is, upon conjugation by $\boldsymbol{E}$, we get the usual condition for an orthogonal matrix.

EXAMPLE 3.73 (An infinite-dimensional example). Consider the vector space V of square-integrable continuous functions on $\mathbb{R}$ as in Section 3.3.1 (or, for simplicity, consider the vector space V of compactly supported continuous functions on $\mathbb{R}$ of Example 3.10). For a given $a \in \mathbb{R}$, consider the endomorphism F of V defined by

$$(Ff)(x) = f(x + a).$$

Since the integral on $\mathbb{R}$ is translation-invariant, we get

$$(Ff, Fg) = \int_{-\infty}^{+\infty} f(x+a) g(x+a) \, \mathrm{d}x = \int_{-\infty}^{+\infty} f(x) g(x) \, \mathrm{d}x = (f, g),$$

so F is orthogonal.

Here is a useful characterization of orthogonal operators:

THEOREM 3.74. *Let F be an endomomorphism of an inner product space V. Then the following are equivalent:*

(1) F is orthogonal.
(2) F preserves all norms and angles.
(3) F preserves all norms.

The condition that F preserves all norms means

$$\|Fv\| = \|v\|$$

for every $v \in V$. Also recall that the angle $\theta_{v,w}$ between two nonzero vectors $v, w \in V$ is defined as the unique angle in $[0, \pi]$ such that

$$\cos \theta_{v,w} = \frac{(v, w)}{\|v\| \|w\|}.$$

PROOF OF THEOREM 3.74. By the definition of norm and angle, it is clear that (1) implies (2). It is also obvious that (2) implies (3), which contains only one of the two conditions.

We then only have to show that (3) implies (1). This is an immediate consequence of the formula

$$(v, w) = \frac{1}{2}(\|v + w\|^2 - \|v\|^2 - \|w\|^2)$$

which holds for every $v, w \in V$. □

Condition (3) implies that an orthogonal operator is injective. In fact, suppose $Fv = 0$. Then, from $0 = \|Fv\| = \|v\|$, we get that also $v = 0$. This has the following immediate corollary.[10]

COROLLARY 3.75. *An orthogonal operator on a finite-dimensional inner product space is invertible.*

One can easily prove that the composition of two orthogonal operators is an orthogonal operator, that the inverse of an orthogonal operator is an orthogonal operator, and that the identity map is an orthogonal operator. As a consequence of this and of Corollary 3.75, we have the

PROPOSITION 3.76. *The set* $\mathrm{O}(V)$ *of orthogonal operators on a finite-dimensional inner product space* V *is a group, called the* ***orthogonal group*** *of* V*.*

REMARK 3.77. In the case of $\mathbb{R}^n$ with the dot product, we write $\mathrm{O}(n)$ for the corresponding group of orthogonal matrices

$$\mathrm{O}(n) = \{\boldsymbol{A} \in \mathrm{Mat}_{n\times n}(\mathbb{R}) \mid \boldsymbol{A}^\mathsf{T}\boldsymbol{A} = \mathbf{1}\},$$

called the **orthogonal group**.

REMARK 3.78. On an infinite-dimensional inner product space an orthogonal operator may fail to be surjective. Consider, e.g., the orthogonal operator

$$(a_1, a_2, \dots) \mapsto (0, a_1, a_2, \dots)$$

on the space $\mathbb{R}^\infty$ of finite real sequences introduced in Example 3.11.

[10]The corollary also follows from Remark 3.69 and Example 3.68.

3.5.1. Digression: Linear conformal maps. Note that just preserving all angles is not enough for an endomorphism to be orthogonal.[11]

One says that an endomorphism F of an inner product space V is conformal if it is injective and preserves all angles, i.e.,

$$\frac{(Fv, Fw)}{\|Fv\|\|Fw\|} = \frac{(v, w)}{\|v\|\|w\|}$$

for all $v, w \in V \setminus 0$ (injectivity is necessary to make sense of the left hand side).

Lemma 3.79. *A linear conformal map can be uniquely written as the composition of a rescaling ($\lambda\,\mathrm{Id}$, $\lambda > 0$) and an orthogonal operator G, i.e., $F = \lambda G$.*

Proof. It is enough to show that $\frac{\|Fv\|}{\|v\|}$ takes the same value, which we denote by λ, for every nonzero vector v. In fact, this implies

$$\|Gv\| = \left\|\frac{Fv}{\lambda}\right\| = \frac{\|Fv\|}{\lambda} = \|v\|,$$

so G preserves all norms and is therefore orthogonal.

First observe that, if v is any nonzero vector and $v' := \frac{v}{\|v\|}$ its corresponding unit vector, we have $\|Fv\| = \|F(\|v\|v')\| = \|v\|\|F(v')\|$, so $\frac{\|Fv\|}{\|v\|} = \|Fv'\|$ (we have only used the linearity of F here).

Therefore, it is enough to show that $\|Fv\|$ takes the same value, which we denote by λ, for every unit vector v.

Let v and w be any two unit vectors. Then $(v + w, v - w) = 0$. Since F is conformal, we have

$$0 = (F(v + w), F(v - w)) = \|Fv\|^2 - \|Fw\|^2.$$

□

Remark 3.80. By inspection in the proof, we see that actually any injective endomorphism F that preserves orthogonality, i.e., such that

$$v \perp w \implies Fv \perp Fw,$$

is of the form λG, with $\lambda > 0$ and G orthogonal, and therefore conformal.

[11] Consider for example the endomorphism $F = \lambda\,\mathrm{Id}$ with $\lambda \neq 0, 1$.

3.5.2. Isometries. The notion of orthogonal operator may be generalized to linear maps between different spaces. Let $(V, (\ ,\ )_V)$ and $(W, (\ ,\ )_W)$ be inner product spaces. A linear map $F\colon V \to W$ is called an isometry (more precisely, a linear isometry) if

$$(Fv_1, Fv_2)_W = (v_1, v_2)_V$$

for all $v_1, v_2 \in V$. Following verbatim the proof of Theorem 3.74, we see that F is an isometry iff it preserves all norms and angles and that the latter happens iff F preserves all norms. We also see that an isometry is always injective. If V and W are finite-dimensional, we then have $\dim V \leq \dim W$, and F is an isomorphisms iff $\dim V = \dim W$.

Example 3.81. The inclusion map of a subspace, with the restriction of the inner product as in Example 3.5, is an isometry.

Example 3.82 (Sine series and spaces of sequences). Consider the space

$$V := \{\phi \in C^0([0, L]) \mid \phi(0) = \phi(L) = 0\}$$

of Example 3.38 with the orthonormal system provided there. To each function $f \in V$ we may assign the real sequence $(b_1, b_2, \dots)$ with $b_k := (e_k, f)$ as in (3.21). If f is in the span V' of the sine functions, b_k is the coefficient of its expansion in the basis $(e_k)_{k\in\mathbb{N}_{>0}}$ of V', so only finitely many b_ks are different from zero. For a general $f \in V$, however, infinitely many b_ks may be different from zero. Still we have Bessel's inequality (3.22) which shows that $(b_1, b_2, \dots)$ is a square-summable sequence (see Section 3.3.2). Therefore, we have a linear map

$$\begin{array}{rcl} F\colon \{\phi \in C^0([0, L]) \mid \phi(0) = \phi(L) = 0\} & \to & \ell^2_{\mathbb{R}} \\ f & \mapsto & \left(\sqrt{\frac{2}{L}} \int_0^L f(x) \sin\left(\frac{\pi k x}{L}\right) \mathrm{d}x\right)_{k\in\mathbb{N}_{>0}} \end{array}$$

Thanks to Parseval's identity (3.23), F is actually an isometry. It is not surjective though.[12]

Example 3.83. If $(e_1, \dots, e_n)$ is an orthonormal basis of V, then the linear map $F\colon V \to \mathbb{R}^n$ that assigns to a vector v the column vector with components its coefficients $v^i = (e_i, v)$ is a bijective isometry (in the notations of Remark 1.52, $F = \Phi_{\mathcal{B}}^{-1}$ with $\mathcal{B} = (e_1, \dots, e_n)$).

From this observation and from Theorem 3.44, we get the

Theorem 3.84. *Every n-dimensional inner product space possesses a bijective isometry with $\mathbb{R}^n$ endowed with the dot product.*

[12] We may enlarge V to contain noncontinuous square-integrable functions. All the above works. By the injectivity of the isometry F, we then get square-summable sequences of coefficients beyond those coming from continuous functions.

3.5.3. The orthogonal groups. In this section we analyze the group $\mathrm{O}(n)$ of orthogonal matrices, introduced in Remark 3.77, in particular for $n = 2$ and $n = 3$.

The first remark is that the condition $\boldsymbol{A}^\mathsf{T}\boldsymbol{A} = \mathbf{1}$ implies $(\det \boldsymbol{A})^2 = 1$. Therefore,

$$\det \boldsymbol{A} = \pm 1.$$

Orthogonal matrices with determinant 1 form a subgroup of $\mathrm{O}(n)$ called the special orthogonal group:

$$\mathrm{SO}(n) := \{\boldsymbol{A} \in \mathrm{Mat}_{n\times n}(\mathbb{R}) \mid \boldsymbol{A}^\mathsf{T}\boldsymbol{A} = \mathbf{1} \text{ and } \det \boldsymbol{A} = 1\}.$$

We also write

$$\mathrm{O}^-(n) := \{\boldsymbol{A} \in \mathrm{Mat}_{n\times n}(\mathbb{R}) \mid \boldsymbol{A}^\mathsf{T}\boldsymbol{A} = \mathbf{1} \text{ and } \det \boldsymbol{A} = -1\}.$$

We have $\mathrm{O}(n) = \mathrm{SO}(n) \sqcup \mathrm{O}^-(n)$.[13]

Note that $\mathrm{O}^-(n)$ is not a subgroup. Also note that for every $\boldsymbol{A}, \boldsymbol{B} \in \mathrm{O}^-(n)$, we have $\boldsymbol{AB} \in \mathrm{SO}(n)$. On the other hand, for every $\boldsymbol{A} \in \mathrm{O}^-(n)$ and every $\boldsymbol{B} \in \mathrm{SO}(n)$, we have that $\boldsymbol{AB}$ and $\boldsymbol{BA}$ lie in $\mathrm{O}^-(n)$.

REMARK 3.85. An example of a matrix in $\mathrm{O}^-(n)$ is

$$\boldsymbol{S} := \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & -1 \end{pmatrix};$$

i.e., $\boldsymbol{Se}_i = \boldsymbol{e}_i$ for $i < n$ and $\boldsymbol{Se}_n = -\boldsymbol{e}_n$. The matrix $\boldsymbol{S}$ acts on $\mathbb{R}^n$ as the reflection through the plane $\mathrm{Span}\{e_1, \ldots, e_{n-1}\}$. Note that for every $\boldsymbol{A} \in \mathrm{O}^-(n)$, we have $\boldsymbol{B} := \boldsymbol{AS} \in \mathrm{SO}(n)$; equivalently, $\boldsymbol{A} = \boldsymbol{BS}$. This means that, in order to describe the elements of $\mathrm{O}(n)$, it is enough to describe those of $\mathrm{SO}(n)$ and then apply $\boldsymbol{S}$.

REMARK 3.86. If n is odd, then also $-\mathbf{1}$ is an element of $\mathrm{O}^-(n)$, so we can write each $\boldsymbol{A} \in \mathrm{O}^-(n)$ as $(-\mathbf{1})(-\boldsymbol{A})$, with $-\boldsymbol{A} \in \mathrm{SO}(n)$. Note that, in any dimension, $-\mathbf{1}$ acts on $\mathbb{R}^n$ as the reflection through the origin.

3.5.3.1. *The groups* $\mathrm{O}(2)$ *and* $\mathrm{SO}(2)$. Let $\boldsymbol{A} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SO}(2)$. Its inverse, by (1.8), is $\boldsymbol{A} = \left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$. Imposing this to be equal to $\boldsymbol{A}^\mathsf{T} = \left(\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right)$ yields

$$d = a \quad \text{and} \quad b = -c,$$

so $\boldsymbol{A} = \left(\begin{smallmatrix} a & -c \\ c & a \end{smallmatrix}\right)$. Since $1 = \det \boldsymbol{A} = a^2 + c^2$, we may parametrize the entries by $a = \cos\theta$ and $c = \sin\theta$. We have proved the

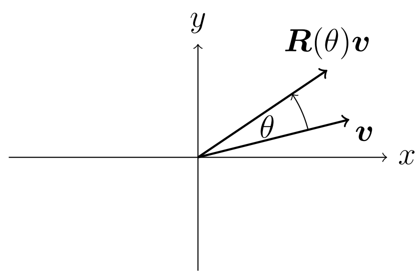
[13]The symbol $\sqcup$ denotes disjoint union.

Proposition 3.87. *A matrix in* SO(2) *has the form*

$$\boldsymbol{R}(\theta) := \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

The angle θ is uniquely determined if we take it in the interval $\theta \in [0, 2\pi)$.

Remark 3.88. If $\boldsymbol{v}$ is a vector in $\mathbb{R}^2$, then $\boldsymbol{R}(\theta)\boldsymbol{v}$ is the vector rotated counterclockwise by the angle θ. From this observation, or by direct computation, we get $\boldsymbol{R}(\theta)\boldsymbol{R}(\theta') = \boldsymbol{R}(\theta + \theta')$. The group SO(2) may then be interpreted as the group of rotations on the plane centered at the origin.



Remark 3.89. Note that $\boldsymbol{R}(0) = \mathbf{1}$ and $\boldsymbol{R}(\pi) = -\mathbf{1}$. On the other hand, for any $\theta \in (0, \pi) \cup (\pi, 2\pi)$, $\boldsymbol{R}(\theta)$ does not preserve any line, so it cannot have any eigenvalue and is therefore not diagonalizable (over the reals). We can see this also via the characteristic polynomial

$$P_{\boldsymbol{R}(\theta)} = \lambda^2 - 2\lambda\cos\theta + 1,$$

whose roots are $\lambda_\pm = \mathrm{e}^{\pm i\theta}$.

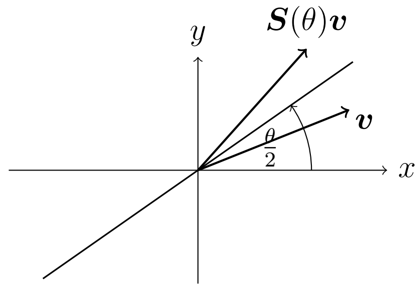
As for $\boldsymbol{A} \in \mathrm{O}^-(2)$, it follows—see also Remark 3.85—that it can then be written as $\boldsymbol{R}(\theta)\boldsymbol{S}$ with $\boldsymbol{S} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ for some θ. Therefore, we have the

Proposition 3.90. *A matrix in* $\mathrm{O}^-(2)$ *has the form*

$$\boldsymbol{S}(\theta) := \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

The angle θ is uniquely determined if we take it in the interval $\theta \in [0, 2\pi)$.

Note that $\boldsymbol{S}(\theta)^2 = \mathbf{1}$ for every θ. Geometrically, $\boldsymbol{S}(\theta)$ may be interpreted as a reflection.



PROPOSITION 3.91. *The matrix $\boldsymbol{S}(\theta)$ is diagonalizable and similar to $\left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right)$. Geometrically, $\boldsymbol{S}(\theta)$ acts on $\mathbb{R}^2$ as the reflection through the line that passes through the origin and forms an angle $\frac{\theta}{2}$ with the x-axis.*

We give two proofs, one with a more geometric flavor and the other fully algebraic.

GEOMETRIC PROOF. As $\boldsymbol{S}(\theta)$ is given by $\boldsymbol{R}(\theta)\boldsymbol{S}$, its action on a vector $\boldsymbol{v}$ is given by the successive application of $\boldsymbol{S}$ and $\boldsymbol{R}(\theta)$.

Now, if $\boldsymbol{v}$ is a nonzero vector that forms an angle α with the x-axis, then $\boldsymbol{S}\boldsymbol{v}$ is the vector of the same length that forms the angle $-\alpha$ with the x-axis, as $\boldsymbol{S}$ is the reflection through the x-axis. Applying the rotation $\boldsymbol{R}(\theta)$ next gives the vector of the same length that forms the angle $\theta - \alpha$ with the x-axis.

A nonzero vector $\boldsymbol{v}$ is then kept fixed by $\boldsymbol{S}(\theta)$ iff $\alpha = \theta - \alpha \mod 2\pi$, i.e., $2\alpha = \theta \mod 2\pi$. For $\alpha, \theta \in [0, \pi)$, this yields $\alpha = \frac{\theta}{2}$ or $\alpha = \pi + \frac{\theta}{2}$, which shows that the line $L_{\frac{\theta}{2}}$ that forms the angle $\frac{\theta}{2}$ with the x-axis is invariant.

It is then convienient to write $\alpha = \frac{\theta}{2} + \beta$. A nonzero vector forming this angle with the x-axis is then sent to the vector of the same length that forms the angle $\frac{\theta}{2} - \beta$ with the x-axis. If we now measure angles w.r.t. the line $L_{\frac{\theta}{2}}$, we see that a nonzero vector forming an angle β with it is mapped by $\boldsymbol{S}(\theta)$ to the vector of the same length that forms the angle $-\beta$ with it. This shows that $\boldsymbol{S}(\theta)$ is the reflection through $L_{\frac{\theta}{2}}$. □

ALGEBRAIC PROOF. Note that $\boldsymbol{S}(0) = \left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right)$ and $\boldsymbol{S}(\pi) = \left(\begin{smallmatrix}-1 & 0\\ 0 & 1\end{smallmatrix}\right)$, so the proposition is immediately proved in these two cases.

Assume then $\theta \in (0, \pi) \cup (\pi, 2\pi)$. The characteristic polynomial

$$P_{\boldsymbol{S}(\theta)} = \lambda^2 + 1,$$

has the two distinct real roots ± 1, so $\boldsymbol{S}(\theta)$ is diagonalizable and similar to $\left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right)$. We now look for a basis of eigenvectors $(\boldsymbol{v}_+, \boldsymbol{v}_-)$.

The eigenvector equation $\boldsymbol{S}(\theta)\boldsymbol{v}_+ = \boldsymbol{v}_+$, with $\boldsymbol{v}_+ = \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$, yields

$$a\cos\theta + b\sin\theta = a,$$

so

$$b = a\frac{1-\cos\theta}{\sin\theta} = a\tan\frac{\theta}{2}.$$

That is, the line L_+ that is fixed by $\boldsymbol{S}(\theta)$ has pendency $\frac{\theta}{2}$.

Similarly, the eigenvector equation $\boldsymbol{S}(\theta)\boldsymbol{v}_- = -\boldsymbol{v}_-$, with $\boldsymbol{v}_- = \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$, yields

$$a\cos\theta + b\sin\theta = -a,$$

so

$$b = -a\frac{1+\cos\theta}{\sin\theta} = a\tan\left(\frac{\theta}{2} + \frac{\pi}{2}\right).$$

Therefore, the line L_- that gets reflected through its origin is rotated by $\frac{\pi}{2}$ with respect to L_+, so it is orthogonal to it. □

3.5.3.2. *The groups* O(3) *and* SO(3). Analogously to what we have done in two dimensions, we now want to determine the normal form of a matrix $\boldsymbol{A}$ in O(3).

The first remark is that the characteristic polynomial of a real 3×3 matrix is a real cubic polynomial and has therefore at least one real root. The second remark is that, if $\boldsymbol{A}$ is orthogonal, then $\|\boldsymbol{A}\boldsymbol{v}\| = \|\boldsymbol{v}\|$ for any $\boldsymbol{v}$. In particular, if $\boldsymbol{v}$ is an eigenvector for a real eigenvalue λ, we then get $|\lambda|\|\boldsymbol{v}\| = \|\boldsymbol{v}\|$, so $|\lambda| = 1$. We have then proved the

Lemma 3.92. *A matrix $\boldsymbol{A} \in \mathrm{O}(3)$ has at least one real eigenvalue, which can only be* 1 *or* -1.[14]

Now, let $\boldsymbol{v}$ be an eigenvector of $\boldsymbol{A} \in \mathrm{O}(3)$. Then the plane $W := \boldsymbol{v}^\perp$ orthogonal to $\boldsymbol{v}$ is $\boldsymbol{A}$-invariant. In fact, if $\boldsymbol{w} \in W$, then $0 = (\boldsymbol{w}, \boldsymbol{v}) = (\boldsymbol{A}\boldsymbol{w}, \boldsymbol{A}\boldsymbol{v}) = \lambda(\boldsymbol{A}\boldsymbol{w}, \boldsymbol{v})$, so $\boldsymbol{A}\boldsymbol{w} \in W$ (recall that $\lambda = \pm 1$). Moreover, for every $\boldsymbol{w}, \boldsymbol{w}' \in W$, we still have $(\boldsymbol{A}\boldsymbol{w}, \boldsymbol{A}\boldsymbol{w}') = (\boldsymbol{w}, \boldsymbol{w}')$. Therefore, the restriction F of $\boldsymbol{A}$ to W is an orthogonal operator. By choosing an orthonormal basis $(\boldsymbol{w}_1, \boldsymbol{w}_2)$ of W, we then represent F as an orthogonal 2×2 matrix $\boldsymbol{B}$. Observe that, in the orthonormal basis $\left(\frac{\boldsymbol{v}}{\|\boldsymbol{v}\|}, \boldsymbol{w}_1, \boldsymbol{w}_2\right)$, our matrix $\boldsymbol{A}$ becomes

$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & & \\ 0 & \multicolumn{2}{c}{\boldsymbol{B}} \end{pmatrix}.$$

Therefore, $\det \boldsymbol{A} = \lambda \det \boldsymbol{B}$.

Consider first the case $\boldsymbol{A} \in \mathrm{SO}(3)$. Since $\boldsymbol{B} \in \mathrm{O}(2)$, we have $\det \boldsymbol{B} = \pm 1$. If $\boldsymbol{B} \in \mathrm{SO}(2)$, we then have $\lambda = 1$. If $\boldsymbol{B} \in \mathrm{O}^-(2)$,

[14]The lemma immediately generalizes to $\mathrm{O}(2k+1)$ for any k.

then by Proposition 3.5.3.1, we know that $\boldsymbol{B}$ has an eigenvector with eigenvalue 1. Therefore, we have the

PROPOSITION 3.93. *A matrix* $\boldsymbol{A} \in \mathrm{SO}(3)$ *has at least one eigenvector with eigenvalue* 1.

The line spanned by this eigenvector is called a principal axis of $\boldsymbol{A}$.

If we now let $\boldsymbol{v}$ be an eigenvector of $\boldsymbol{A} \in \mathrm{SO}(3)$ with eigenvalue 1, we get that the orthogonal matrix $\boldsymbol{B}$ that represents the restriction of $\boldsymbol{A}$ to $\boldsymbol{v}^\perp$ in an orthonormal basis has determinant 1. From Remark 3.88, we get therefore the following geometric description.

THEOREM 3.94. *A matrix* $\boldsymbol{A} \in \mathrm{SO}(3)$ *acts on* $\mathbb{R}^3$ *as a rotation around its principal axis.*[15]

The group SO(3) may then be interpreted as the group of space rotations centered at the origin.

By choosing an orthonormal basis $(\boldsymbol{v}, \boldsymbol{w}_1, \boldsymbol{w}_2)$, with $\boldsymbol{v}$ a unit vector on the principal axis and $(\boldsymbol{w}_1, \boldsymbol{w}_2)$ an orthonormal basis of $\boldsymbol{v}^\perp$ such that $\boldsymbol{w}_1 \times \boldsymbol{w}_2 = \boldsymbol{v}$, we then get, thanks to Proposition 3.71, the following theorem, where $\boldsymbol{S}$ is the matrix with colums $\boldsymbol{v}, \boldsymbol{w}_1, \boldsymbol{w}_2$.

THEOREM 3.95. *For every matrix* $\boldsymbol{A} \in \mathrm{SO}(3)$, *there is a matrix* $\boldsymbol{S} \in \mathrm{SO}(3)$ *such that*

$$\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \begin{pmatrix} 1 & 0 \quad 0 \\ 0 & \\ 0 & \boldsymbol{R}(\theta) \end{pmatrix}$$

for some angle θ.

If $\boldsymbol{A} \in \mathrm{O}^-(3)$, then $-\boldsymbol{A} \in \mathrm{SO}(3)$, which can be written in the above form. Note that $-\boldsymbol{R}(\theta) = \boldsymbol{R}(\theta + \pi)$. Therefore, now writing θ instead of $\theta + \pi$, we have the

THEOREM 3.96. *For every matrix* $\boldsymbol{A} \in \mathrm{O}^-(3)$, *there is a matrix* $\boldsymbol{S} \in \mathrm{SO}(3)$ *such that*

$$\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \begin{pmatrix} -1 & 0 \quad 0 \\ 0 & \\ 0 & \boldsymbol{R}(\theta) \end{pmatrix}$$

for some angle θ.

[15]The principal axis is uniquely determined if $\boldsymbol{A} \neq \mathbf{1}$.

3.5.4. Symmetric and skew-symmetric operators. We conclude with some related concepts.

DEFINITION 3.97. An endomorphism F of an inner product space V is called symmetric if

$$(Fv, w) = (v, Fw)$$

for every $v, w \in V$.

EXAMPLE 3.98. In the case of the dot product on $\mathbb{R}^n$, the endomorphism defined by an $n \times n$ matrix $\boldsymbol{A}$ is symmetric iff the matrix $\boldsymbol{A}$ is symmetric, i.e., $\boldsymbol{A}^\mathsf{T} = \boldsymbol{A}$. In particular, we have that an endomorphism of a finite-dimensional inner product space is symmetric iff its representing matrix in some orthonormal basis is symmetric.

EXAMPLE 3.99. If we have an orthogonal decomposition

$$V = \bigoplus_{i \in S} W_i,$$

then the projection P_i to the W_i-component is symmetric, as shown in Remark 3.67.

Symmetric operators occur in several applications. As an example, note that $B(v, w) := (v, Fw)$ is a bilinear symmetric form iff F is symmetric. We can then define a new inner product on V, provided F is positive definite, i.e., $(v, Fv) > 0$ for every $v \in V \setminus \{0\}$.

DEFINITION 3.100. An endomorphism F of an inner product space V is called skew-symmetric (or antisymmetric) if

$$(Fv, w) = -(v, Fw)$$

for every $v, w \in V$.

EXAMPLE 3.101. In the case of the dot product on $\mathbb{R}^n$, the endomorphism defined by an $n \times n$ matrix $\boldsymbol{A}$ is skew-symmetric iff the matrix $\boldsymbol{A}$ is skew-symmetric, i.e., $\boldsymbol{A}^\mathsf{T} = -\boldsymbol{A}$. In particular, we have that an endomorphism of a finite-dimensional inner product space is skew-symmetric iff its representing matrix in some orthonormal basis is skew-symmetric.

Skew-symmetric operators are closely related to orthogonal operators. One relation is the following. Suppose $\boldsymbol{O}(t)$ is a differentiable map $\mathbb{R} \to \mathrm{O}(n)$. Then, differentiating the identity $\boldsymbol{O}^\mathsf{T}\boldsymbol{O} = \mathbf{1}$, we get $\dot{\boldsymbol{O}}^\mathsf{T}\boldsymbol{O} + \boldsymbol{O}^\mathsf{T}\dot{\boldsymbol{O}} = \mathbf{0}$. Therefore, the matrix $\boldsymbol{A} := \dot{\boldsymbol{O}}\boldsymbol{O}^{-1}$ is skew-symmetric.

Another relation is the following. Let $\boldsymbol{A}$ be skew-symmetric. Define $\boldsymbol{O}(t) := \mathrm{e}^{\boldsymbol{A}t}$. Then we have $\boldsymbol{O}^\mathsf{T} = \boldsymbol{O}^{-1}$, so $\boldsymbol{O}(t)$ is orthogonal for every t. Finally, note that the trace of a skew-symmetric matrix vanishes. Therefore, by Proposition 2.11, $\det \boldsymbol{O}(t) = 1$ for every t. As a consequence, $\mathrm{e}^{\boldsymbol{A}t}$ is a differentiable map $\mathbb{R} \to \mathrm{SO}(n)$.

Remark 3.102. The vector space of skew-symmetric $n\times n$ matrices is denoted by $\mathfrak{so}(n)$:

$$\mathfrak{so}(n) := \{\boldsymbol{A} \in \mathrm{Mat}_{n\times n}(\mathbb{R}) \mid \boldsymbol{A}^\mathsf{T} = -\boldsymbol{A}\}.$$

This notation helps remembering that we have the exponential map

$$\begin{array}{rccc} \exp\colon & \mathfrak{so}(n) & \to & \mathrm{SO}(n) \\ & \boldsymbol{A} & \mapsto & \mathrm{e}^{\boldsymbol{A}} \end{array}$$

Remark 3.103. Note that $\boldsymbol{R}(\theta) = \mathrm{e}^{\boldsymbol{\rho}(\theta)}$ with $\boldsymbol{\rho}(\theta) = \left(\begin{smallmatrix} 0 & -\theta \\ \theta & 0 \end{smallmatrix}\right)$, so $\exp\colon \mathfrak{so}(2) \to \mathrm{SO}(2)$ is surjective.

Remark 3.104. Note that

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & & \\ 0 & \multicolumn{2}{c}{\boldsymbol{R}(\theta)} \end{pmatrix} = \mathrm{e}^{\left(\begin{smallmatrix} 0 & 0\ 0 \\ 0 & \\ 0 & \boldsymbol{\rho}(\theta) \end{smallmatrix}\right)}$$

with $\boldsymbol{\rho}(\theta)$ as above. From Theorem 3.95, we then have that for every matrix $\boldsymbol{O} \in \mathrm{SO}(3)$, there is a matrix $\boldsymbol{S} \in \mathrm{SO}(3)$ such that

$$\boldsymbol{O} = \boldsymbol{S}\mathrm{e}^{\left(\begin{smallmatrix} 0 & 0\ 0 \\ 0 & \\ 0 & \boldsymbol{\rho}(\theta) \end{smallmatrix}\right)}\boldsymbol{S}^{-1} = \mathrm{e}^{\boldsymbol{A}}$$

with

$$\boldsymbol{A} = \boldsymbol{S}\begin{pmatrix} 0 & 0 & 0 \\ 0 & & \\ 0 & \multicolumn{2}{c}{\boldsymbol{\rho}(\theta)} \end{pmatrix}\boldsymbol{S}^{-1} = \boldsymbol{S}\begin{pmatrix} 0 & 0 & 0 \\ 0 & & \\ 0 & \multicolumn{2}{c}{\boldsymbol{\rho}(\theta)} \end{pmatrix}\boldsymbol{S}^\mathsf{T},$$

which is skew-symmetric, since $\boldsymbol{\rho}(\theta)$ is so. Therefore, $\exp\colon \mathfrak{so}(3) \to \mathrm{SO}(3)$ is surjective.

Remark 3.105. We will see in Corollary 4.114 that for every n the exponential map $\exp\colon \mathfrak{so}(n) \to \mathrm{SO}(n)$ is surjective.

3.5.5. Digression: Minkowski space. It was observed by Poincaré and by Minkowski that Lorentz boosts (a part of the transformations, discovered by Lorentz, that preserve Maxwell's equation) are like rotations but for a "squared norm" given by $x^2 + y^2 + z^2 - c^2t^2$, where c is the speed of light.[16] As this is central to Einstein's special relativity, we will briefly digress on it.

[16] More precisely, Poincaré was thinking in terms of the usual euclidean norm but proposed to consider time as an imaginary coordinate, i.e., he considered the squared norm $x^2 + y^2 + z^2 + (\mathrm{i}ct)^2$. Minkowski later observed that it was more natural to keep working over the reals, just by changing the last sign in the formula for the squared norm.

The idea is to define an "inner product" on $\mathbb{R}^{n+1}$, called the Minkowski (inner) product, (with $n = 3$ being the case for usual space–time) as

$$(\boldsymbol{v}, \boldsymbol{w}) := \boldsymbol{v}^{\mathsf{T}} \boldsymbol{\eta} \boldsymbol{w}$$

with[17]

$$\boldsymbol{\eta} = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix},$$

which is called the Minkowski metric.

Conventionally one uses indices from 0 to n (instead of 1 to $n+1$), so a vector $\boldsymbol{v}$ in the Minkowski space $(\mathbb{R}^{n+1}, (\ ,\ ))$ is denoted by

$$\boldsymbol{v} = \begin{pmatrix} v^0 \\ v^1 \\ \vdots \\ v^n \end{pmatrix},$$

and one denotes the standard basis by $(\boldsymbol{e}_0, \dots, \boldsymbol{e}_n)$. The components $v^1, \dots, v^n$ are thought of as space components, whereas v^0 is thought of as c times the time component. The Minkowski product then reads explicitly

$$(\boldsymbol{v}, \boldsymbol{w}) = -v^0 w^0 + \sum_{i=1}^{n} v^i w^i.$$

We have the "squared norm" $(\boldsymbol{v}, \boldsymbol{v}) = -c^2 t^2 + \sum_{i=1}^{n} (x^i)^2$, where we have set $v^0 = ct$ and $v^i = x^i$ for $i > 0$.

The Minkowski product is bilinear and symmetric but <u>not</u> positive definite. For example $(\boldsymbol{e}_0, \boldsymbol{e}_0) = -1$. On the other hand $(\boldsymbol{e}_i, \boldsymbol{e}_i) = 1$ for $i = 1, \dots, n$, and $(\boldsymbol{e}_i, \boldsymbol{e}_j) = 0$ for all $i \neq j$, $i, j = 0, \dots, n$. The standard basis is then orthogonal. It may also be considered orthonormal if we accept that one vector has "squared norm" -1 instead of $+1$.

By analogy with the basis vectors, one says that a vector $\boldsymbol{v}$ is time-like if $(\boldsymbol{v}, \boldsymbol{v}) < 0$ and space-like if $(\boldsymbol{v}, \boldsymbol{v}) > 0$. Note that there are also nonzero vectors $\boldsymbol{v}$ satisfying $(\boldsymbol{v}, \boldsymbol{v}) = 0$, e.g., $\boldsymbol{v} = \boldsymbol{e}_0 + \boldsymbol{e}_1$. Such vectors are called null or light-like.

---

[17] An equally spread convention defines

$$\boldsymbol{\eta} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & \ddots & \\ & & & -1 \end{pmatrix}.$$

3.5.5.1. *Lorentz transformations.* Much of the theory of inner products extends to the case of the Minkowski product. We will focus here only on the topic of orthogonal matrices, namely matrices $\boldsymbol{A}$ such that

$$(\boldsymbol{A}\boldsymbol{v}, \boldsymbol{A}\boldsymbol{w}) = (\boldsymbol{v}, \boldsymbol{w})$$

for all $\boldsymbol{v}, \boldsymbol{w}$. From the definition of the Minkowski product, we then have that $\boldsymbol{A}$ is orthogonal iff

$$\boldsymbol{A}^\mathsf{T} \boldsymbol{\eta} \boldsymbol{A} = \boldsymbol{\eta}.$$

Orthogonal matrices on the Minkowski space $(\mathbb{R}^{n+1}, (\ ,\ ))$ are also called Lorentz transformations. They form a group, called the Lorentz group, which is denoted by $\mathrm{O}(1, n)$.

Note that $\boldsymbol{A} \in \mathrm{O}(1, n)$ implies $(\det \boldsymbol{A})^2 = 1$. Lorentz transformations with determinant equal to $+1$ are called proper and form a subgroup denoted by $\mathrm{SO}(1, n)$.

If we denote by $\boldsymbol{a}_i$, $i = 0, \dots, n$, the columns of a Lorentz transformation $\boldsymbol{A}$, then we get

$$(\boldsymbol{a}_i, \boldsymbol{a}_j) = \begin{cases} -1 & \text{if } i = j = 0, \\ 1 & \text{if } i = j > 0, \\ 0 & \text{if } i \neq j. \end{cases}$$

That is, the columns of $\boldsymbol{A}$ are orthogonal to each other, the last n columns are normalized and space-like, and the first column is normalized and time-like. Writing

$$\boldsymbol{a}_0 = \begin{pmatrix} a_0^0 \\ a_0^1 \\ \vdots \\ a_0^n \end{pmatrix},$$

we then get

$$-(a_0^0)^2 + \sum_{i=1}^n (a_0^i)^2 = -1,$$

which shows that either $a_0^0 \geq 1$ or $a_0^0 \leq -1$. A Lorentz transformation $\boldsymbol{A}$ with $a_0^0 \geq 1$ is called orthochronous.

Orthochronous Lorentz transformations form a subgroup of $\mathrm{O}(1, n)$, which is denoted by $\mathrm{O}_+(1, n)$. The intersection of $\mathrm{O}_+(1, n)$ and $\mathrm{SO}(1, n)$ is the group $\mathrm{SO}_+(1, n)$ of proper, orthochronous Lorentz transformations:

$$\mathrm{SO}_+(1, n) = \{\boldsymbol{A} \in \mathrm{Mat}_{(n+1)\times(n+1)}(\mathbb{R}) \mid \boldsymbol{A}^\mathsf{T} \boldsymbol{\eta} \boldsymbol{A} = \boldsymbol{\eta},\ \det \boldsymbol{A} = 1,\ a_0^0 \geq 1\}.$$

3.5.5.2. *The group* $\mathrm{SO}_+(1,1)$. In the case of one space dimension, we essentially repeat the analysis of Section 3.5.3.1, with minor, but important, variations.

Let $\boldsymbol{A} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SO}_+(1,1)$. Its inverse, by (1.8), is $\boldsymbol{A} = \left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$. The defining equation

$$\boldsymbol{A}^\mathsf{T}\boldsymbol{\eta} = \boldsymbol{\eta}\boldsymbol{A}^{-1}$$

reads explicitly

$$\begin{pmatrix} -a & c \\ -b & d \end{pmatrix} = \begin{pmatrix} -d & b \\ -c & a \end{pmatrix}.$$

Therefore, we get

$$d = a \quad \text{and} \quad b = c,$$

so $\boldsymbol{A} = \left(\begin{smallmatrix} a & c \\ c & a \end{smallmatrix}\right)$. Since $1 = \det \boldsymbol{A} = a^2 - c^2$ and $a \geq 1$, we may parametrize the entries by $a = \cosh\tau$ and $c = \sinh\tau$. We have proved the

PROPOSITION 3.106. *A matrix in* $\mathrm{SO}_+(1,1)$ *has the form*

$$\boldsymbol{L}(\tau) := \begin{pmatrix} \cosh\tau & \sinh\tau \\ \sinh\tau & \cosh\tau \end{pmatrix},$$

for a uniquely determined $\tau \in \mathbb{R}$.[18]

Unlike nontrivial rotations, Lorentz transformations leave certain lines invariant. Actually, $\boldsymbol{L}(\tau)$ is diagonalizable for every τ. In fact, for $\tau \neq 0$, the characteristic polynomial

$$P_{\boldsymbol{L}(\tau)} = \lambda^2 - 2\lambda\cosh\tau + 1$$

has the two distinct roots

$$\lambda_\pm = \cosh\tau \pm \sqrt{\cosh^2\tau - 1} = \cosh\tau \pm \sinh\tau = \mathrm{e}^{\pm\tau}.$$

As corresponding eigenvectors we may take

$$\boldsymbol{v}_+ = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \boldsymbol{v}_- = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Note that both $\boldsymbol{v}_+$ and $\boldsymbol{v}_-$ are light-like.

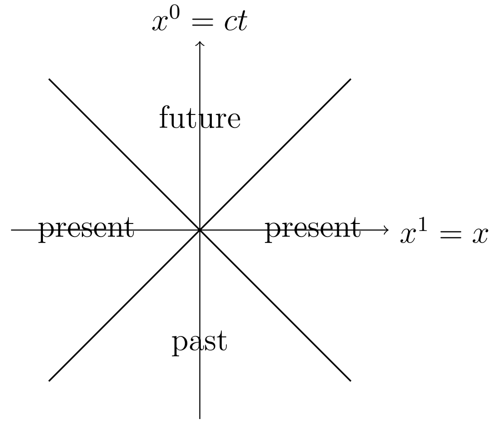
As a consequence, the diagonal lines $L_+ = \mathbb{R}\boldsymbol{v}_+$ and $L_- = \mathbb{R}\boldsymbol{v}_-$ are invariant under every Lorentz transformation. Moreover, each of the four connected regions in $\mathbb{R}^2 \setminus (L_+ \cup L_-)$ is also invariant.

---

[18] In physics, one uses the parameters $\gamma = \cosh\tau$ and $\beta = \tanh\tau$, so we have

$$\boldsymbol{L} = \begin{pmatrix} \gamma & \beta\gamma \\ \beta\gamma & \gamma \end{pmatrix},$$

with the relation $\gamma = \frac{1}{\sqrt{1-\beta^2}}$. Physically $\boldsymbol{L}$ with $\beta = \frac{v}{c}$ describes the transformation between two observers moving at relative velocity v with respect to each other.



Following the physical interpretation, these four regions are called future, past, present to the left, present to the right. They can be characterized as follows: a nonzero vector $\boldsymbol{v} = \binom{v^0}{v^1}$ is in the future if $(\boldsymbol{v}, \boldsymbol{v}) < 0$ and $v^0 > 0$, in the past if $(\boldsymbol{v}, \boldsymbol{v}) < 0$ and $v^0 < 0$, and in the present if $(\boldsymbol{v}, \boldsymbol{v}) > 0$ (to the left if $v^1 < 0$ and to the right if $v^1 > 0$).[19]

By the usual diagonalization procedure, we may write

$$\boldsymbol{L}(\tau) = \boldsymbol{S}^{-1} \begin{pmatrix} e^{\tau} & 0 \\ 0 & e^{-\tau} \end{pmatrix} \boldsymbol{S}$$

with $\boldsymbol{S} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. This may also be easily checked explicitly by observing that $\boldsymbol{S}^{-1} = \frac{1}{2}\boldsymbol{S}$.

Finally, note that $\begin{pmatrix} e^{\tau} & 0 \\ 0 & e^{-\tau} \end{pmatrix} = e^{\begin{pmatrix} \tau & 0 \\ 0 & -\tau \end{pmatrix}}$. Therefore, $\boldsymbol{L}(\tau) = e^{\boldsymbol{S}^{-1}\begin{pmatrix} \tau & 0 \\ 0 & -\tau \end{pmatrix}\boldsymbol{S}}$. Since $\boldsymbol{S}^{-1} \begin{pmatrix} \tau & 0 \\ 0 & -\tau \end{pmatrix} \boldsymbol{S} = \begin{pmatrix} 0 & \tau \\ \tau & 0 \end{pmatrix}$, we finally have

$$\boldsymbol{L}(\tau) = e^{\begin{pmatrix} 0 & \tau \\ \tau & 0 \end{pmatrix}}.$$

Exercises for Chapter 3

3.1. Let $V = \mathrm{Mat}_{n\times n}(\mathbb{R})$ be the vector space of $n \times n$ real matrices. Show that

$$(\boldsymbol{A}, \boldsymbol{B}) := \mathrm{tr}(\boldsymbol{A}^{\mathsf{T}}\boldsymbol{B})$$

is an inner product on V.

3.2. The goal of this exercise is to show that the following two conditions on $\boldsymbol{g} = \begin{pmatrix} \alpha & \beta \\ \beta & \delta \end{pmatrix}$ are equivalent:

(A) $\boldsymbol{g}$ is positive definite (i.e., $(\boldsymbol{v}, \boldsymbol{w}) := \boldsymbol{v}^{\mathsf{T}}\boldsymbol{g}\boldsymbol{w}$ is an inner product).

[19] In higher dimensions, one gets a similar structure with the difference that the present becomes a connected region, whereas future and past are still disconnected by the light cone $\{\boldsymbol{v} \mid (\boldsymbol{v}, \boldsymbol{v}) = 0\}$.

(B) $\alpha > 0$ and $\det \boldsymbol{g} > 0$.

Throughout the exercise we write $\boldsymbol{x} = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$.

(a) Compute $(\boldsymbol{x}, \boldsymbol{x})$.

(b) Assuming (A) show the following statements:

(i) If $(\boldsymbol{x}, \boldsymbol{x}) > 0$ for $y = 0$ and every $x \neq 0$, then $\alpha > 0$.

(ii) Now assume $y \neq 0$ and $\alpha > 0$. Define $f(z) := \frac{(\boldsymbol{x},\boldsymbol{x})}{y^2}$, with $z = \frac{x}{y}$. Considering the minimum of f for $z \in \mathbb{R}$, show that $\det \boldsymbol{g} > 0$.

(c) Assuming (B), show the following statements:

(i) $(\boldsymbol{x}, \boldsymbol{x}) > 0$ for $y = 0$ and every $x \neq 0$.

(ii) Now assume $y \neq 0$. Show that

$$(\boldsymbol{x}, \boldsymbol{x}) > \frac{(\alpha x + \beta y)^2}{\alpha}.$$

Conclude that $\boldsymbol{g}$ is positive definite.

3.3. In each of the following cases, compute the angle between the vectors v and w.

(a) $V = \mathbb{R}^3$ with the dot product, $v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $w = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$.

(b) $V = \mathbb{R}^2$, inner product defined by $\boldsymbol{g} = \left(\begin{smallmatrix} 1 & 2 \\ 2 & 5 \end{smallmatrix}\right)$, $v = \left(\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right)$, $w = \left(\begin{smallmatrix} 3 \\ 4 \end{smallmatrix}\right)$.

(c) $V = C^0([0,1])$, $(f,g) = \int_0^1 fg \, \mathrm{d}x$, v the function $1 - x$, w the function x^2.

3.4. Let V be the space of real polynomial functions (you may regard it as a subspace of $C^0(\mathbb{R})$). Show that

$$(f,g) := \int_{-\infty}^{\infty} \mathrm{e}^{-x^2} fg \, \mathrm{d}x$$

is an inner product.

3.5. Using the Cauchy–Schwarz inequality, find the maximum and the maximum point(s) of the function $x + 2y + 3z$ on the unit sphere

$$S^2 = \{(x,y,z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}.$$

Hint: Note that using the dot product the function can be expressed as $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \boldsymbol{v}$ with $\boldsymbol{v} \in S^2$.

3.6. Recall that a sequence $a = (a_1, a_2, \dots)$ of real numbers is called square summable if the series $\sum_{i=1}^{\infty} (a_i)^2$ converges. Let $\ell^2(\mathbb{R})$ denote the set of square summable real sequences.

(a) Show that the sum of two square summable real sequences is again square summable: i.e., given $(a_1, a_2, \dots), (b_1, b_2, \dots) \in$

$\ell^2(\mathbb{R})$, one has $\sum_{i=1}^{\infty}(a_i+b_i)^2<\infty$.
Hint: Use the triangle inequality on $\mathbb{R}^N$ to find an estimate of $\sum_{i=1}^{N}(a_i+b_i)^2$.

(b) Show that for every $(a_1,a_2,\dots),(b_1,b_2,\dots)\in\ell^2(\mathbb{R})$, the series $\sum_{i=1}^{\infty}|a_ib_i|$ converges.
Hint: Use the Cauchy–Schwarz inequality on $\mathbb{R}^N$ to find an estimate of $\sum_{i=1}^{N}|a_ib_i|$.

(c) Use the above to show that $\ell^2(\mathbb{R})$ is a real vector space with the addition rule $(a_1,a_2,\dots)+(b_1,b_2,\dots)=(a_1+b_1,a_2+b_2,\dots)$ and that

$$((a_1,a_2,\dots),(b_1,b_2,\dots)):=\sum_{i=1}^{\infty}a_ib_i$$

is an inner product.

3.7. In this exercise we work on $V=\mathbb{R}^3$ with the dot product.

(a) Let $\boldsymbol{w}=\begin{pmatrix}1\\1\\1\end{pmatrix}$. For each of the following $\boldsymbol{v}\in\mathbb{R}^3$ find the orthogonal decomposition $\boldsymbol{v}=\boldsymbol{v}_\parallel+\boldsymbol{v}_\perp$ with $\boldsymbol{v}_\parallel$ proportional to $\boldsymbol{w}$ and $\boldsymbol{v}_\perp$ orthogonal to it.

(i) $\boldsymbol{v}=\begin{pmatrix}1\\0\\0\end{pmatrix}$.

(ii) $\boldsymbol{v}=\begin{pmatrix}3\\3\\3\end{pmatrix}$.

(iii) $\boldsymbol{v}=\begin{pmatrix}3\\-4\\1\end{pmatrix}$.

(b) Let $\boldsymbol{w}_1=\begin{pmatrix}1\\1\\1\end{pmatrix}$, $\boldsymbol{w}_2=\begin{pmatrix}1\\0\\-1\end{pmatrix}$, $\boldsymbol{w}_3=\begin{pmatrix}2\\-4\\2\end{pmatrix}$.

(i) Show that $(\boldsymbol{w}_1,\boldsymbol{w}_2,\boldsymbol{w}_3)$ is an orthogonal basis.

(ii) Compute the corresponding orthonormal basis $(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3)$ obtained by normalizing the vectors.

(iii) Expand $\begin{pmatrix}1\\2\\3\end{pmatrix}$ and $\begin{pmatrix}4\\5\\6\end{pmatrix}$ in this orthonormal basis.

3.8. In $V=C^0([0,1])$—the space of continuous functions on the interval $[0,1]$ with inner product $(f,g)=\int_0^1 fg\,\mathrm{d}x$—let $g(x)=x^2$. For each of the following $f\in V$ find the orthogonal decomposition $f=f_\parallel+f_\perp$ with $f_\parallel$ proportional to g and $f_\perp$ orthogonal to it.

(a) $f(x)=x$.

(b) $f(x)=x^2(1-2x^3)$.

(c) $f(x)=\mathrm{e}^x$.

3.9. Consider the function

$$f(x) = \frac{1}{2}\left[\left(x - \frac{\pi}{2}\right)^2 - \left(\frac{\pi}{2}\right)^2\right]$$

in

$$V := \{\phi \in C^0([0,\pi]) \mid \phi(0) = \phi(\pi) = 0\}.$$

Compute the coefficients $b_k = (e_k, f)$ with respect to the orthonormal system

$$e_k(x) = \sqrt{\frac{2}{\pi}} \sin(kx),$$

$k \in \mathbb{N}_{>0}$.
Hint: Use $e_k = -\frac{e_k''}{k^2}$ and integration by parts.

3.10.
(a) Apply the Gram–Schmidt process to the following vectors:
(i) $\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$ in $\mathbb{R}^3$.
(ii) $\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \boldsymbol{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \boldsymbol{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$ in $\mathbb{R}^4$.
(b) Find an orthonormal basis for the positive definite matrix $\boldsymbol{g} = \left(\begin{smallmatrix} 1 & 2 \\ 2 & 5 \end{smallmatrix}\right)$, and write $\boldsymbol{g} = \boldsymbol{E}^\mathsf{T}\boldsymbol{E}$ for some matrix $\boldsymbol{E}$.

3.11. Apply the Gram–Schmidt process to the three vectors $(1, x, x^2)$ in the space V of polynomial functions on $\mathbb{R}$ with inner product

$$(f, g) := \int_{-\infty}^{\infty} \mathrm{e}^{-x^2} f g \, \mathrm{d}x.$$

Hint: You can use the formulae

$$I(\alpha) := \int_{-\infty}^{\infty} \mathrm{e}^{-\alpha x^2} \mathrm{d}x = \sqrt{\frac{\pi}{\alpha}},$$

$$\left(\frac{\mathrm{d}}{\mathrm{d}\alpha}\right)^k I(\alpha) = (-1)^k \int_{-\infty}^{\infty} \mathrm{e}^{-\alpha x^2} x^{2k} \, \mathrm{d}x,$$

which hold for every real $\alpha > 0$ and every positive integer k.

3.12. The goal of this exercise is to prove the following statement:
If all the leading principal minors of a real symmetric matrix $\boldsymbol{g}$ are positive, then $\boldsymbol{g}$ is positive definite.
We prove it by induction on the size of the matrix $\boldsymbol{g}$.
(a) Show that the statement is true if $\boldsymbol{g}$ is a 1×1 matrix.

(b) Assume that the statement holds for $n \times n$ matrices and let $\boldsymbol{g}$ be an $(n+1) \times (n+1)$ symmetric matrix satisfying the condition in the statement. Write

$$\boldsymbol{g} = \begin{pmatrix} \boldsymbol{h} & \boldsymbol{b} \\ \boldsymbol{b}^\mathsf{T} & a \end{pmatrix}$$

with $\boldsymbol{h}$ an $n \times n$ matrix, $\boldsymbol{b}$ an n-column vector and a a scalar.

(i) Show that $\boldsymbol{h}$ is positive definite, so there is an invertible matrix $\boldsymbol{E}$ such that $\boldsymbol{h} = \boldsymbol{E}^\mathsf{T}\boldsymbol{E}$.
Hint: Use the induction hypothesis.

(ii) Show that

$$a > \|\boldsymbol{F}\boldsymbol{b}\|_\mathrm{E}^2,$$

where $\boldsymbol{F} = \boldsymbol{E}^{\mathsf{T},-1}$ and $\|\boldsymbol{v}\|_\mathrm{E} := \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}$ denotes the euclidean norm on $\mathbb{R}^n$.
Hint: Use the condition in the statement and the identity

$$\det \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} = \det \boldsymbol{A} \det(\boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}),$$

which holds for every block matrix with $\boldsymbol{A}$ invertible.[20]

(iii) For a fixed n-column vector $\boldsymbol{w}$ consider the function

$$f(x) := \begin{pmatrix} \boldsymbol{w}^\mathsf{T} & x \end{pmatrix} \boldsymbol{g} \begin{pmatrix} \boldsymbol{w} \\ x \end{pmatrix}.$$

(A) Show that

$$f(x) = ax^2 + 2\,\boldsymbol{b} \cdot \boldsymbol{w}\, x + \|\boldsymbol{E}\boldsymbol{w}\|_\mathrm{E}^2.$$

(B) Show that the minimum value of f is

$$f_\mathrm{min} = \|\boldsymbol{E}\boldsymbol{w}\|_\mathrm{E}^2 - \frac{(\boldsymbol{b} \cdot \boldsymbol{w})^2}{a}.$$

(C) Assuming $\boldsymbol{b} \cdot \boldsymbol{w} \neq 0$, show that

$$f_\mathrm{min} > \frac{\|\boldsymbol{F}\boldsymbol{b}\|_\mathrm{E}^2 \|\boldsymbol{E}\boldsymbol{w}\|_\mathrm{E}^2 - (\boldsymbol{b} \cdot \boldsymbol{w})^2}{\|\boldsymbol{F}\boldsymbol{b}\|_\mathrm{E}^2}.$$

Hint: Use point 12(b)ii.

---

[20] This identity may be proved by writing

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{C} & \boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B} \end{pmatrix} \begin{pmatrix} \boldsymbol{1} & \boldsymbol{A}^{-1}\boldsymbol{B} \\ \boldsymbol{0} & \boldsymbol{1} \end{pmatrix}$$

and computing the determinant.

(D) Show that

$$\|\boldsymbol{F}\boldsymbol{b}\|_{\mathrm{E}}^2 \|\boldsymbol{E}\boldsymbol{w}\|_{\mathrm{E}}^2 \geq (\boldsymbol{b}\cdot\boldsymbol{w})^2.$$

Hint: Use the Cauchy–Schwarz inequality for the dot product.

(iv) Conclude that $\boldsymbol{g}$ is positive definite.

3.13.

(a) Let $W \subset \mathbb{R}^4$ be the subspace generated by

$$\boldsymbol{w}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \quad \text{and} \quad \boldsymbol{w}_2 = \begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \end{pmatrix}$$

(i) Find an orthonormal basis of W.
(ii) Use it to compute the projection operators P_W and $P_{W^\perp}$.
(iii) Apply P_W and $P_{W^\perp}$ to the vector $\boldsymbol{v} = (0,1,2,1)^{\mathsf{T}}$.

(b) Let V be the vector space of all polynomials of degrees two or less restricted to the interval $[0,1] \in \mathbb{R}$, endowed with the inner product $(f,g) = \int_0^1 f(x)g(x)\mathrm{d}x$.
(i) Find an orthonormal basis for the subspace V' spanned by $f(x) = x-1$ and $g(x) = x + x^2$.
(ii) Find the projection operators $P_{V'}$ and $P_{V'^\perp}$.
(iii) Apply the projection operators to the vector $h(x) = 1$.

3.14. Which of the following matrices are orthogonal?

$$\boldsymbol{A} = \begin{pmatrix} \sqrt{3} & 1 \\ 1 & -\sqrt{3} \end{pmatrix}, \quad \boldsymbol{B} = \frac{1}{7}\begin{pmatrix} 3 & -2 & -6 \\ -2 & 6 & -3 \\ -6 & -3 & -2 \end{pmatrix}, \quad \boldsymbol{C} = \begin{pmatrix} \frac{3}{\sqrt{10}} & -\frac{1}{\sqrt{10}} & 0 \\ -\frac{1}{\sqrt{35}} & -\frac{3}{\sqrt{35}} & \sqrt{\frac{5}{7}} \\ \frac{1}{\sqrt{14}} & \frac{3}{\sqrt{14}} & \sqrt{\frac{2}{7}} \end{pmatrix}.$$

3.15. Let $V = C_c^0(\mathbb{R})$ be the space of compactly supported functions on $\mathbb{R}$ with inner product $(f,g) = \int_{-\infty}^{+\infty} fg\,\mathrm{d}x$.

(a) For which values of $a, b \in \mathbb{R}$ is the following endomorphism of V orthogonal?

$$(Ff)(x) = bf(ax).$$

(b) Show that

$$(\tilde{F}f)(x) = \sqrt{\cosh(x)}\,f(\sinh(x))$$

is an orthogonal endomorphism of V.

3.16. In this exercise we investigate some properties of the cross product in $\mathbb{R}^3$. We do not recall its explicit definition but define it implicitly as follows: for every $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^3$, the cross product $\boldsymbol{a} \times \boldsymbol{b}$ is the uniquely determined vector satisfying

$$\boldsymbol{a} \times \boldsymbol{b} \cdot \boldsymbol{c} = \det(\boldsymbol{a}\ \boldsymbol{b}\ \boldsymbol{c}),$$

for every $\boldsymbol{c} \in \mathbb{R}^3$. Here $(\boldsymbol{a}\ \boldsymbol{b}\ \boldsymbol{c})$ denotes the matrix with columns $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$. Using this characterization of the cross product, prove the following statements.

(a) $\forall \boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^3$

$$\boldsymbol{a} \times \boldsymbol{b} = -\boldsymbol{b} \times \boldsymbol{a}.$$

(b) $\forall \boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^3$, $\boldsymbol{a} \neq \boldsymbol{0}$,

$$\boldsymbol{a} \times \boldsymbol{b} = \boldsymbol{a} \times \boldsymbol{b}_\perp,$$

with the orthogonal decomposition $\boldsymbol{b} = \boldsymbol{b}_\parallel + \boldsymbol{b}_\perp$ with respect to $\boldsymbol{a}$.

(c) $\forall \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^3$

$$\boldsymbol{a} \times \boldsymbol{b} \cdot \boldsymbol{c} = \boldsymbol{c} \times \boldsymbol{a} \cdot \boldsymbol{b} = \boldsymbol{b} \times \boldsymbol{c} \cdot \boldsymbol{a}.$$

(d) $\forall \boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^3$

$$\boldsymbol{a} \times \boldsymbol{b} \perp \boldsymbol{a} \quad \text{and} \quad \boldsymbol{a} \times \boldsymbol{b} \perp \boldsymbol{b}.$$

(e) Assume $\boldsymbol{a} \perp \boldsymbol{b}$. Then

$$\|\boldsymbol{a} \times \boldsymbol{b}\| = \|\boldsymbol{a}\| \|\boldsymbol{b}\|.$$

Hint: Observe that, for $\boldsymbol{a}$ and $\boldsymbol{b}$ different from zero,

$$\|\boldsymbol{a} \times \boldsymbol{b}\|^2 = \det(\boldsymbol{a}\ \boldsymbol{b}\ \boldsymbol{a} \times \boldsymbol{b}) = \|\boldsymbol{a}\| \|\boldsymbol{b}\| \|\boldsymbol{a} \times \boldsymbol{b}\| \det \boldsymbol{S},$$

where $\boldsymbol{S}$ is an orthogonal matrix. Why is $\boldsymbol{a} \times \boldsymbol{b} \neq \boldsymbol{0}$?

(f) $\forall \boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^3$, with $\|\boldsymbol{a}\| = 1$,

$$\|\boldsymbol{a} \times \boldsymbol{b}\| = \|\boldsymbol{b}_\perp\|.$$

(g) For $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^3$ with $\|\boldsymbol{a}\| = 1$ define $\boldsymbol{v} = (\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{a}$. Use all the above to show the following statements.
  (i) $\boldsymbol{v} \in \operatorname{Span}\{\boldsymbol{a}, \boldsymbol{b}\}$.
  (ii) $\boldsymbol{v} = \lambda \boldsymbol{b}_\perp$ for some $\lambda \in \mathbb{R}$.
  (iii) $\boldsymbol{v} \cdot \boldsymbol{b}_\perp = \|\boldsymbol{b}_\perp\|^2$.

Conclude that

$$(\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{a} = \boldsymbol{b}_\perp.$$

Therefore,

$$\boldsymbol{b} = (\boldsymbol{a} \cdot \boldsymbol{b})\, \boldsymbol{a} + (\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{a}$$

3.17. Let $\boldsymbol{a} \in \mathbb{R}^3$ with $\|\boldsymbol{a}\| = 1$. For every $\boldsymbol{b} \in \mathbb{R}^3$, we use the orthogonal decomposition $\boldsymbol{b} = \boldsymbol{b}_\parallel + \boldsymbol{b}_\perp$ with respect to $\boldsymbol{a}$.

(a) Assuming $\boldsymbol{b}_\perp \neq 0$, show that $(\boldsymbol{a}, \boldsymbol{b}_\perp, \boldsymbol{a} \times \boldsymbol{b})$ is an orthogonal basis of $\mathbb{R}^3$.

(b) Show that

$$\boldsymbol{R}_{\boldsymbol{a}}(\theta)\boldsymbol{b} = \boldsymbol{b}_\parallel + \cos\theta\, \boldsymbol{b}_\perp + \sin\theta\, \boldsymbol{a} \times \boldsymbol{b}_\perp,$$

where $\boldsymbol{R}_{\boldsymbol{a}}(\theta) \in \mathrm{SO}(3)$ is the counterclockwise rotation by the angle θ around $\boldsymbol{a}$.

Hint: Normalize the orthogonal basis of the previous point and use points (16d) and (16f) of exercise 3.16.

3.18. Recall that $\mathfrak{so}(3)$ is the vector space of 3×3 real skew-symmetric matrices. Show the following statements.

(a) For all $\boldsymbol{A}, \boldsymbol{B} \in \mathfrak{so}(3)$ and for all $\boldsymbol{S} \in \mathrm{SO}(3)$, we have

(i) $[\boldsymbol{A}, \boldsymbol{B}] := \boldsymbol{AB} - \boldsymbol{BA} \in \mathfrak{so}(3)$.

(ii) $\boldsymbol{SAS}^{-1} \in \mathfrak{so}(3)$.

(iii) $\boldsymbol{S}[\boldsymbol{A}, \boldsymbol{B}]\boldsymbol{S}^{-1} = [\boldsymbol{SAS}^{-1}, \boldsymbol{SBS}^{-1}]$.

(b) The following matrices are a basis of $\mathfrak{so}(3)$:

$$R_1 := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad R_2 := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad R_3 := \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

(c) Writing

$$\boldsymbol{x} \cdot \boldsymbol{R} := \sum_{i=1}^{3} x^i R_i,$$

show that

$$[\boldsymbol{x} \cdot \boldsymbol{R}, \widetilde{\boldsymbol{x}} \cdot \boldsymbol{R}] = \boldsymbol{x} \times \widetilde{\boldsymbol{x}} \cdot \boldsymbol{R}.$$

Hint: Use the explicit formula

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \times \begin{pmatrix} \widetilde{x} \\ \widetilde{y} \\ \widetilde{z} \end{pmatrix} = \begin{pmatrix} y\widetilde{z} - z\widetilde{y} \\ z\widetilde{x} - x\widetilde{z} \\ x\widetilde{y} - y\widetilde{z} \end{pmatrix}.$$

3.19. Let

$$\boldsymbol{O} = \frac{1}{3}\begin{pmatrix} 2 & -1 & 2 \\ 2 & 2 & -1 \\ -1 & 2 & 2 \end{pmatrix}.$$

(a) Check that $\boldsymbol{O} \in \mathrm{SO}(3)$.

(b) Find its principal axis.
Hint: Find an eigenvector of $\boldsymbol{O}$. Note that you should already know an eigenvalue, so there is no need to solve the characteristic equation.

(c) Find its rotation angle.
Hint: Choose a unit vector $\boldsymbol{w}$ orthogonal to the principal axis and compute $\boldsymbol{O}\boldsymbol{w}$.

CHAPTER 4

Hermitian products

In this chapter we extend, and adapt, the notion of inner product to complex vector spaces. Among the reasons to do so, one is that working over complex numbers gives more tools also to solve problems over the reals. Another is that this framework is needed for quantum mechanics.

4.1. The standard hermitian product on $\mathbb{C}^n$

The dot product $\boldsymbol{z} \cdot \boldsymbol{w} = \boldsymbol{z}^\mathsf{T} \boldsymbol{w}$ may in principle be extended to complex vectors. The problem is that $\boldsymbol{z} \cdot \boldsymbol{z} = \sum_i (z^i)^2$ is not generally a real number (even less generally a nonnegative real number), so it cannot be used to define a norm.

Recall that for a complex number z, one may indeed define a norm using the absolute value

$$|z| = \sqrt{\bar{z}z}.$$

If one writes $z = a + \mathrm{i}b$, with a and b real, then $|z| = \sqrt{a^2 + b^2}$, so the absolute value of a complex number is the same as the euclidean norm of the corresponding real vector under the identification (isomorphism of real vector spaces) $\mathbb{C} = \mathbb{R}^2$.

Guided by this, we want to define the norm of a vector $\boldsymbol{z}$ in $\mathbb{C}^n$ as $\|\boldsymbol{z}\| = \sqrt{\sum_{i=1}^n \bar{z}^i z^i}$. Again, note that by writing $z^i = a^i + \mathrm{i}b^i$, with a^i and b^i real, we have $\|\boldsymbol{z}\| = \sqrt{\sum_{i=1}^n ((a^i)^2 + (b^i)^2)}$, so this norm is the same as the euclidean norm under the identification (isomorphism of real vector spaces) $\mathbb{C}^n = \mathbb{R}^{2n}$.

Following the case of the inner product on a real space, we want to get the norm as the square root of the inner product of a vector with itself. In order to do so, we have to adapt the definition of the inner product to the following:[1]

$$\langle \boldsymbol{z}, \boldsymbol{w} \rangle := \sum_{i=1}^{n} \bar{z}^i w^i = \bar{\boldsymbol{z}}^\mathsf{T} \boldsymbol{w},$$

where $\bar{\boldsymbol{z}}$ denotes the vector whose components are the complex conjugates of the components of $\boldsymbol{z}$. This is called the standard hermitian product on $\mathbb{C}^n$ (after Charles Hermite).

Note that the standard hermitian product is linear in the second argument. With respect to the first argument, we have instead

$$\langle \lambda_1 \boldsymbol{z}_1 + \lambda_2 \boldsymbol{z}_2, \boldsymbol{w} \rangle = \bar{\lambda}_1 \langle \boldsymbol{z}_1, \boldsymbol{w} \rangle + \bar{\lambda}_2 \langle \boldsymbol{z}_2, \boldsymbol{w} \rangle .$$

One says that $\langle\ ,\ \rangle$ is antilinear in the first argument.

Also note that $\langle\ ,\ \rangle$ is not symmetric. Instead it satisfies

$$\langle \boldsymbol{z}, \boldsymbol{w} \rangle = \overline{\langle \boldsymbol{w}, \boldsymbol{z} \rangle}.$$

One says that $\langle\ ,\ \rangle$ is conjugate symmetric or hermitian symmmetric.

Finally, we have that $\langle \boldsymbol{z}, \boldsymbol{z} \rangle$ is a nonnegative real number, so we can take its root, which gives back the norm we wanted to consider:

$$\|\boldsymbol{z}\| = \sqrt{\langle \boldsymbol{z}, \boldsymbol{z} \rangle}.$$

4.2. Hermitian spaces

Motivated by the example of the standard hermitian product and its properties, we introduce the following

Definition 4.1 (Hermitian forms). A map $\langle\ ,\ \rangle : V \times V \to \mathbb{C}$, where V is a complex vector space, is called a hermitian form if it satisfies the following two properties.

Linearity in the second argument: For all $v, w_1, w_2 \in V$ and all $\lambda_1, \lambda_2 \in \mathbb{C}$, we have

$$\langle v, \lambda_1 w_1 + \lambda_2 w_2 \rangle = \lambda_1 \langle v, w_1 \rangle + \lambda_2 \langle v, w_2 \rangle .$$

Hermitian symmetry: For all $v, w \in V$, we have[2]

$$\langle v, w \rangle = \overline{\langle w, v \rangle}.$$

[1] This is the most common convention in physics. In most math texts, one uses instead

$$\langle \boldsymbol{z}, \boldsymbol{w} \rangle := \sum_{i=1}^{n} z^i \bar{w}^i = \boldsymbol{z}^\mathsf{T} \bar{\boldsymbol{w}},$$

which produces anyway the same norm $\|\boldsymbol{z}\| = \sqrt{\langle \boldsymbol{z}, \boldsymbol{z} \rangle}$.

[2] Note that complex conjugation is applied to the image of $\langle\ ,\ \rangle$. In the case of the standard hermitian product, we also have $\overline{\langle \boldsymbol{w}, \boldsymbol{z} \rangle} = \langle \bar{\boldsymbol{w}}, \bar{\boldsymbol{z}} \rangle$, but for a general complex vector space V, the complex conjugation "$\bar{v}$" of a vector v is not defined.

LEMMA 4.2. *A hermitian form* $\langle\ ,\ \rangle$ *on* V *also satisfies the following properties.*

Antilinearity in the first argument: *For all* $v_1, v_2, w \in V$ *and all* $\lambda_1, \lambda_2 \in \mathbb{C}$*, we have*

$$\langle \lambda_1 v_1 + \lambda_2 v_2, w \rangle = \bar{\lambda}_1 \langle v_1, w \rangle + \bar{\lambda}_2 \langle v_2, w \rangle .$$

Reality: *For all* $v \in V$*, we have* $\langle v, v \rangle \in \mathbb{R}$*.*

PROOF. This is a very easy computation left to the reader. □

REMARK 4.3. We can summarize the antilinearity in the first argument and the linearity in the second, by saying that $\langle\ ,\ \rangle$ is sesquilinear.

Sesquilinearity: For all $v, v_1, v_2, w, w_1, w_2 \in V$ and all $\lambda_1, \lambda_2 \in \mathbb{C}$, we have[3]

$$\langle \lambda_1 v_1 + \lambda_2 v_2, w \rangle = \bar{\lambda}_1 \langle v_1, w \rangle + \bar{\lambda}_2 \langle v_2, w \rangle ,$$
$$\langle v, \lambda_1 w_1 + \lambda_2 w_2 \rangle = \lambda_1 \langle v, w_1 \rangle + \lambda_2 \langle v, w_2 \rangle .$$

Note that the reality property allows asking whether $\langle v, v \rangle$ is positive or negative or zero.

DEFINITION 4.4. A hermitian form $\langle\ ,\ \rangle$ on V is called positive definite if $(v, v) > 0$ for all $v \in V \setminus \{0\}$.

We now come the fundamental

DEFINITION 4.5 (Hermitian products). A hermitian product is a positive-definite hermitian form. A complex vector space V together with a hermitian product $\langle\ ,\ \rangle$ is called a hermitian product space.[4]

REMARK 4.6. A finite-dimensional hermitian product space is also called a finite-dimensional Hilbert space.

We may now adapt several examples from Section 3.2.

EXAMPLE 4.7 (The standard hermitian product). The standard hermitian product $\bar{\boldsymbol{v}}^{\mathsf{T}} \boldsymbol{w}$ on $\mathbb{C}^n$ is an example of hermitian product. We will see (Theorem 4.65) that, upon choosing an appropriate basis, a hermitian product on a finite-dimensional vector space can always be brought to this form.

[3] This is the most common convention in physics. In most math texts, a sesquilinear form is instead linear in the first argument and antilinear in the second. In both cases, a hermitian form may be defined as a hermitian-symmetric sesquilinear form.

[4] Terminology at this point starts to differ wildly among different authors. You may check, e.g., https://en.wikipedia.org/wiki/Inner_product_space to get some coherent version.

EXAMPLE 4.8 (Subspaces). If W is a subspace of a hermitian product space V, we may restrict the hermitian product to elements of W. This makes W itself into a hermitian product space.

EXAMPLE 4.9. On $\mathbb{C}^n$ we may also define

$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle := \sum_{i=1}^{n} \lambda_i \bar{v}^i w^i$$

for a given choice of complex numbers $\lambda_1, \ldots, \lambda_n$. This is clearly sesquilinear, and it is hermitian symmetric iff all λ_is are real. One can easily verify that it is positive definite iff $\lambda_i > 0$ for all $i = 1, \ldots, n$.

EXAMPLE 4.10. More generally, on $\mathbb{C}^n$ we may define

$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle := \bar{\boldsymbol{v}}^{\mathsf{T}} \boldsymbol{g} \boldsymbol{w} = \bar{v}^i g_{ij} w^j, \tag{4.1}$$

where $\boldsymbol{g}$ is a given $n \times n$ complex matrix, and we have used Einstein's convention in the last term. (The previous example is the case when $\boldsymbol{g}$ is diagonal.) Sesquilinearity is clear. Hermitian symmetry is satisfied iff $\boldsymbol{g}$ is self adjoint, i.e.,

$$\bar{\boldsymbol{g}}^{\mathsf{T}} = \boldsymbol{g},$$

where $\bar{\boldsymbol{g}}$ denotes the matrix whose entries are the complex conjugates of the entries of $\boldsymbol{g}$. A self-adjoint matrix $\boldsymbol{g}$ is called positive definite if the corresponding hermitian form is positive definite, i.e., if $\bar{\boldsymbol{v}}^{\mathsf{T}} \boldsymbol{g} \boldsymbol{v} > 0$ for every nonzero vector $\boldsymbol{v}$.

The above example motivates the following

DEFINITION 4.11. The adjoint (a.k.a. the hermitian conjugate or the hermitian transpose) of a complex matrix $\boldsymbol{A}$ is the matrix[5]

$$\boldsymbol{A}^{\dagger} := \bar{\boldsymbol{A}}^{\mathsf{T}},$$

where the symbol † is pronounced "dagger." A complex square matrix $\boldsymbol{A}$ is called self adjoint (or hermitian) if it satisfies

$$\boldsymbol{A}^{\dagger} = \boldsymbol{A}.$$

Using this terminology, we may say that (4.1) defines a hermitian form iff $\boldsymbol{g}$ is self-adjoint.

The adjunction has several properties that parallel those of transposition, as described in the following Lemma, whose proof we leave as an exercise.

[5]In the math literature, the adjoint matrix of $\boldsymbol{A}$ is also denoted by $\boldsymbol{A}^*$. The $*$ notation has some advantage in handwriting, for it avoids any possible confusion between the dagger symbol † and the transposition symbol T. Note however that, when dealing with complex matrices, in most cases you can bet one is using hermitian conjugation and not transposition.

LEMMA 4.12. *The adjunction has the following properties (for all matrices* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *and for every complex number* λ*):*

(1) $(\boldsymbol{A}^\dagger)^\dagger = \boldsymbol{A}$.
(2) $(\boldsymbol{A}+\boldsymbol{B})^\dagger = \boldsymbol{A}^\dagger + \boldsymbol{B}^\dagger$.
(3) $(\lambda\boldsymbol{A})^\dagger = \bar\lambda\boldsymbol{A}^\dagger$.
(4) $(\boldsymbol{A}\boldsymbol{B})^\dagger = \boldsymbol{B}^\dagger\boldsymbol{A}^\dagger$.
(5) *If* $\boldsymbol{A}$ *is invertible, then so is* $\boldsymbol{A}^\dagger$*, and we have* $(\boldsymbol{A}^\dagger)^{-1} = (\boldsymbol{A}^{-1})^\dagger$.

REMARK 4.13 (Representing matrix). If we have a hermitian product $\langle\ ,\ \rangle$ on a finite-dimensional space V with a basis $\mathcal{B} = (e_1, \dots, e_n)$, we define the representing matrix $\boldsymbol{g}$ with entries

$$g_{ij} := \langle e_i, e_j\rangle\,.$$

Hermitian symmetry of $\langle\ ,\ \rangle$ implies $g_{ji} = \bar g_{ij}$, so $\boldsymbol{g}$ is self-adjoint. If we expand $v = v^i e_i$ and $w = w^i e_i$, then by sesquilinearity we get

$$\langle v, w\rangle = \bar v^i g_{ij} w^j,$$

that is, formula (4.1). Upon using the isomorphism $\phi_{\mathcal{B}}\colon \mathbb{C}^n \to V$ of Remark 1.52, we therefore get on $\mathbb{C}^n$ the inner product of Example 3.7.

Examples 3.10 and 3.11, where the inner product was defined in terms of an integral, may be readily generalized to the complex case. We only need to spend a few words about complex-valued functions and their integrals.

Namely, if f is a complex-valued function, then one can write $f = u + \mathrm{i}v$ with uniquely determined real-valued functions u and v, called the real and imaginary part of f, respectively. One says that f is continuous/differentiable/smooth/... if u and v are so. The integral of f is defined as

$$\int_a^b f\,\mathrm{d}x := \int_a^b u\,\mathrm{d}x + \mathrm{i}\int_a^b v\,\mathrm{d}x. \tag{4.2}$$

EXAMPLE 4.14 (Continuous functions on a compact interval). Denote by $V = C^0([a,b],\mathbb{C})$ the vector space of complex-valued functions on the interval $[a,b]$. Then

$$\langle f, g\rangle := \int_a^b \bar f g\,\mathrm{d}x \tag{4.3}$$

is a hermitian product on V. Sesquilinearity and hermitian symmetry are obvious. Note that

$$\langle f, f\rangle = \int_a^b |f|^2\,\mathrm{d}x,$$

where $|\ |$ denotes the absolute value of complex numbers (for $f = u + \mathrm{i}v$, we have $|f| = \sqrt{u^2 + v^2}$). As $|f|^2$ is now a real-valued function, we may proceed exactly as in Example 3.9 to show positivity.

EXAMPLE 4.15 (Compactly supported continuous functions). Let $V = C^0_c(\mathbb{R}, \mathbb{C})$ be the vector space of complex-valued functions on $\mathbb{R}$ with compact support: i.e., $f = u + \mathrm{i}v$ belongs to $C^0_c(\mathbb{R}, \mathbb{C})$ iff u and v are continuous function with compact support—i.e., u and v belong to the $C^0_c(\mathbb{R})$. We define

$$\langle f, g\rangle := \int_{-\infty}^{\infty} \bar{f} g \, \mathrm{d}x.$$

This can be proved to be a hermitian product as in the previous example.

We also have the following immediate generalization of Example 3.11.

EXAMPLE 4.16. A sequence $a = (a_1, a_2, \dots)$ of complex numbers is called finite if only finitely many a_is are different from zero (equivalently, if there is an N such that $a_i = 0$ for all $i > N$). We denote by $\mathbb{C}^\infty$ the vector space of all finite real sequences, with vector space operations

$$\lambda(a_1, a_2, \dots) = (\lambda a_1, \lambda a_2, \dots)$$
$$(a_1, a_2, \dots) + (b_1, b_2, \dots) = (a_1 + b_1, a_2 + b_2, \dots).$$

It is a hermitian product space with

$$\langle a, b\rangle := \sum_{i=1}^{\infty} \bar{a}_i b_i,$$

where the right hand side clearly converges because it is a finite sum.

4.2.1. Nondegeneracy and Dirac's notation. As in the case of the inner product, the positivity condition of a hermitian product $\langle\ ,\ \rangle$ on V implies in particular the nondegeneracy conditions

$$\langle v, w\rangle = 0\ \forall w \iff v = 0 \quad \text{and} \quad \langle v, w\rangle = 0\ \forall v \iff w = 0.$$

These conditions imply that we have the following injective maps (bijective if V is finite-dimensional):

(1) An antilinear map

$$\begin{array}{ccc} V & \to & V^* \\ v & \mapsto & \langle v,\ \rangle \end{array}$$

(2) A linear map

$$\begin{array}{rcl} V & \to & \overline{V}^* \\ w & \mapsto & \langle\ , w\rangle \end{array}$$

where $\overline{V}^*$ denotes the space of antilinear maps $V \to \mathbb{C}$.

At this point, it is convenient to introduce a notation due to P. A. M. Dirac, which is commonly used in quantum mechanics.

In this notation, known as the bra–ket notation, the hermitian product, called the bracket, of two vectors v and w is denoted by $\langle\, v \,|\, w\,\rangle$, with a vertical bar instead of a comma. It has to be thought of as the juxtaposition of the bra $\langle v\,|$ and the ket $|\,w\rangle$.

In this notation, kets are just another way of denoting vectors (more precisely, we should think of $w \mapsto |\,w\rangle$ as the identity map on V with vectors written in two different notations). As remarked above, we may also regard kets as antilinear maps $V \to \mathbb{C}$.

Similarly, bras are just another way of denoting covectors (more precisely, we should think of $v \mapsto \langle v\,|$ as the antilinear injective map $V \to V^*$ introduced above).

The bra–ket notation helps remembering what is linear and what is antilinear (just by analogy to the properties of the bracket with respect to its arguments).

Dirac's bra–ket notation becomes even more useful when dealing with orthonormal bases, as we will see in Section 4.4.

4.2.2. The adjoint of an operator. The notion of adjoint may be extended to operators.

Definition 4.17. Let F be an endomorphism of an inner product space $(V, \langle\ ,\ \rangle)$. Its adjoint operator is an endomorphism $F^\dagger$ of V such that for every $v, w \in V$ we have

$$\langle v, Fw\rangle = \left\langle F^\dagger v, w\right\rangle.$$

Remark 4.18. The existence of $F^\dagger$ is not guaranteed if V is infinite-dimensional.[6] However, if $F^\dagger$ exists, then it is uniquely determined. In fact, let $\widetilde{F}^\dagger$ be also an endomorphism satisfying $\langle v, Fw\rangle = \left\langle \widetilde{F}^\dagger v, w\right\rangle$ for every $v, w \in V$. Then we have $\left\langle (\widetilde{F}^\dagger - F^\dagger)v, w\right\rangle = 0$ for every $v, w \in V$. By the nondegeneracy of the hermitian product, we then get $(\widetilde{F}^\dagger - F^\dagger)v =$ for every $v \in V$, i.e., $\widetilde{F}^\dagger = F^\dagger$.

[6]One can easily check that $F^\dagger$ exists iff for every v the linear form $\alpha_v\colon w \mapsto \langle v, Fw\rangle$ is in the image of the antilinear map $z \mapsto \langle z,\ \rangle$ introduced in Section 4.2.1.

EXAMPLE 4.19. In the case of the standard hermitian product on $\mathbb{C}^n$, the adjoint of the endomorphism defined by an $n \times n$ complex matrix $\boldsymbol{A}$ is the endomorphism defined by the adjoint of $\boldsymbol{A}$.

REMARK 4.20. The adjoint of an operator shares the same properties of the adjoint of a matrix as expressed in Lemma 4.12, which follow from the uniqueness of the adjoint:

(1) $(F^\dagger)^\dagger = F$.
(2) $(F+G)^\dagger = F^\dagger + G^\dagger$.
(3) $(\lambda F)^\dagger = \bar{\lambda} F^\dagger$.
(4) $(FG)^\dagger = G^\dagger F^\dagger$.

On a finite-dimensional space one also has that, if F is invertible, then so is $F^\dagger$, and we have $(F^\dagger)^{-1} = (F^{-1})^\dagger$.

4.3. The norm

As in the case of the inner product we can define a norm starting from a hermitian product. Namely, we set

$$\|v\| := \sqrt{\langle v, v\rangle}.$$

We immediately get the first two properties of a norm:

(N.1) $\|v\| \geq 0$ for all $v \in V$, and $\|v\| = 0$ iff $v = 0$.
(N.2) $\|\alpha v\| = |\alpha|\|v\|$ for all $v \in V$ and for all $\alpha \in \mathbb{C}$.[7]

Again we will get the triangle inequality from the

THEOREM 4.21 (Cauchy–Schwarz inequality). *Let $(V, \langle\ ,\ \rangle)$ be a hermitian product space, and let $\|\ \|$ denote the induced norm. Then all $v, w \in V$ satisfy the Cauchy–Schwarz inequality*

$$|\langle v, w\rangle| \leq \|v\|\|w\| \tag{4.4}$$

with equality saturated iff v and w are linearly dependent.

PROOF. The proof runs exactly along the lines of the proof to Theorem 3.16. We describe only the case $w \neq 0$ to outline the, minor, differences. Namely, we consider again the function

$$f(\lambda) := \|v + \lambda w\|^2$$

where now λ is a complex variable. We have

$$f(\lambda) = \langle v + \lambda w, v + \lambda w\rangle = \|v\|^2 + \bar{\lambda}\langle w, v\rangle + \lambda \langle v, w\rangle + \lambda^2\|w\|^2.$$

If we write $\lambda = a + \mathrm{i}b$, we may view f as a real-valued function of the real variables a and b. We have:

(1) $f(a + \mathrm{i}b) \geq 0$ for all a, b.

[7]Here $|\alpha| = \sqrt{\bar{\alpha}\alpha}$ is the absolute value of the complex number α.

(2) $\operatorname{Hess} f(a+\mathrm{i}b) = \begin{pmatrix} 2\|w\|^2 & 0 \\ 0 & 2\|w\|^2 \end{pmatrix}$ for all a, b.

By computing $\frac{\partial f}{\partial a}$ and $\frac{\partial f}{\partial b}$, we see that f has a unique critical point, which is a minimum because the Hessian is positive definite. By explicit computation we get

$$f_{\min} = \|v\|^2 - \frac{|\langle v, w\rangle|^2}{\|w\|^2}.$$

We leave all the remaining details, as well as the proof of the second part of the theorem, to the reader. □

We then get, with the simple details of the proof left to the readers, the following

PROPOSITION 4.22 (The triangle inequality). *Let $(V, \langle\ ,\ \rangle)$ be a hermitian product space, and let $\|\ \|$ denote the induced norm. Then all $v, w \in V$ satisfy the* ***triangle inequality***

$$\|v + w\| \leq \|v\| + \|w\|. \tag{4.5}$$

As an immediate consequence we have the

THEOREM 4.23 (Properties of the norm). *Let $(V, \langle\ ,\ \rangle)$ be a hermitian product space. Then the induced norm $\|\ \|$ satisfies the following three properties*

(N.1) $\|v\| \geq 0$ *for all* $v \in V$, *and* $\|v\| = 0$ *iff* $v = 0$.
(N.2) $\|\alpha v\| = |\alpha|\|v\|$ *for all* $v \in V$ *and for all* $\alpha \in \mathbb{C}$.
(N.3) $\|v + w\| \leq \|v\| + \|w\|$ *for all* $v, w \in V$.

DIGRESSION 4.24 (Complex normed spaces). A **norm** on a complex vector space V is a function $\|\ \|\colon V \to \mathbb{R}$ satisfying properties (N.1), (N.2), and (N.3). A complex vector space endowed with a norm is called a complex **normed space**. The above theorem shows that a hermitian product space is automatically a complex normed space as well. On the other hand, there are norms that are not defined in terms of a hermitian product.

REMARK 4.25 (The other triangle inequality). Like in the case of real normed spaces, see Remark 3.20, also in a complex normed space we have the other triangle inequality,

$$|\|v\| - \|w\|| \leq \|v - w\| \tag{4.6}$$

for all v and w, which is proved exactly like in the real case.

4.3.1. Square-integrable continuous functions. We now want to introduce the space of complex-valued square-integrable continuous functions, which is important for quantum mechanics. We will rely on the discussion of Section 3.3.1.

The first point is to define the integral of a complex-valued continuous function on $\mathbb{R}$. We do this by extending (4.2). Namely, we use the following

DEFINITION 4.26. A complex-valued continuos function $f\colon \mathbb{R} \to \mathbb{C}$ is called integrable if both its real and imaginary parts u and v are integrable according to Definition 3.23. In this case we set

$$\int_{-\infty}^{\infty} f \,\mathrm{d}x := \int_{-\infty}^{\infty} u \,\mathrm{d}x + \mathrm{i} \int_{-\infty}^{\infty} v \,\mathrm{d}x.$$

Next we introduce the main concept.

DEFINITION 4.27. A complex-valued continuos function $f\colon \mathbb{R} \to \mathbb{C}$ is called square integrable if the improper integral $\int_{-\infty}^{\infty} |f|^2 \,\mathrm{d}x$ is finite.

Note that $|f|^2$ is real valued, so the definition reduces to what we discussed in Section 3.3.1.

If we write $f = u + \mathrm{i}v$, with u, v real valued, then $|f|^2 = u^2 + v^2$. Therefore, we get the

LEMMA 4.28. *A complex-valued continuos function is square integrable iff its real and imaginary parts are so.*

This also implies the following

LEMMA 4.29. *If f and g are complex-valued square-integrable continuous functions on $\mathbb{R}$, then so is $f + g$.*

PROOF. Write $f = u + \mathrm{i}v$ and $g = \widetilde{u} + \mathrm{i}\widetilde{v}$. By the previous lemma, we know that $u, v, \widetilde{u}, \widetilde{v}$ are real-valued square-integrable continuos functions. This implies, by Lemma 3.27, that $u+\widetilde{u}$ and $v+\widetilde{v}$ are also square integrable. Therefore, $f + g = (u + \widetilde{u}) + \mathrm{i}(v + \widetilde{v})$ is square integrable (again by Lemma 4.28). □

Since multipliying a complex-valued square-integrable continuous function by a complex constant clearly yields again a square-integrable continuous function, the lemma has the following

COROLLARY 4.30. *The set $L^{2,0}(\mathbb{R}, \mathbb{C})$ of complex-valued square-integrable continuous functions on $\mathbb{R}$ is a subspace of the complex vector space $C^0(\mathbb{R}, \mathbb{C})$ of complex-valued continuous functions on $\mathbb{R}$.*

Next, we have the

LEMMA 4.31. *If f and g are complex-valued square-integrable continuous functions on $\mathbb{R}$, then the product $\bar{f}g$ is integrable.*

PROOF. Write $f = u + \mathrm{i}v$ and $g = \widetilde{u} + \mathrm{i}\widetilde{v}$. Again, by Lemma 4.28, we know that $u, v, \widetilde{u}, \widetilde{v}$ are real-valued square-integrable continuos functions. We have

$$\bar{f}g = u\widetilde{u} + v\widetilde{v} + \mathrm{i}(u\widetilde{v} - v\widetilde{u}).$$

By Lemma 3.29, we have that the real-valued continuous functions $u\widetilde{u}$, $v\widetilde{v}$, $u\widetilde{v}$ and $v\widetilde{u}$ are absolutely integrable. Therefore, the sums $u\widetilde{u} + v\widetilde{v}$ and $u\widetilde{v} - v\widetilde{u}$ are also absolutely integrable, and hence integrable. Therefore, by Definition 4.26, $\bar{f}g$ is integrable. □

We can summarize these results in the

THEOREM 4.32. *On the complex vector space $L^{2,0}(\mathbb{R}, \mathbb{C})$ of complex-valued square-integrable continuous functions on $\mathbb{R}$, we have the hermitian product*

$$\langle f, g\rangle := \int_{-\infty}^{+\infty} \bar{f}g \, \mathrm{d}x.$$

PROOF. The only thing still to check is positivity. We leave it as an exercise. □

4.3.2. Square-summable sequences.

Analogously to the space of complex-valued square-integrable functions, we may study the space of square-summable complex sequences.

DEFINITION 4.33. A sequence $a = (a_1, a_2, \dots)$ of complex numbers is called square summable if the series $\sum_{i=1}^{\infty} |a_i|^2$ converges.

One can prove (we leave the details to the reader) that the sum of two square-summable complex sequences is again square-summable and that, if a and b are square summable, then $\sum_{i=1}^{\infty} \bar{a}_i b_i$ converges.

The hermitian product space of square-summable complex sequences, with hermitian product

$$\langle a, b\rangle := \sum_{i=1}^{\infty} \bar{a}_i b_i,$$

is denoted by ℓ^2 (or, more precisely, by $\ell^2_{\mathbb{C}}$ to stress that we are considering complex sequences).

4.4. Orthogonality

We briefly review the rather straightforward generalization of the concepts presented in Section 3.4.

Let $(V, \langle\ ,\ \rangle)$ be a hermitian product space. Two vectors v and w are called orthogonal if $\langle v, w\rangle = 0$. In this case one writes $v \perp w$. Note that this condition is symmetric (i.e., $v \perp w$ iff $w \perp v$).

An orthogonal system is again defined as a collection $(e_i)_{i\in S}$ of nonzero vectors in V such that $e_i \perp e_j$ for all $i \neq j$ in S. Again we have, with exactly the same proof as in the inner product case, the

Lemma 4.34. *An orhogonal system is linearly independent.*

If the orthogonal system generates the space, we call it an orthogonal basis.

If the vectors e_i are normalized (i.e., $\|e_i\|^2 = 1$), then we speak of an orthonormal system or an orthonormal basis, respectively.

Example 4.35. The standard basis $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ of $\mathbb{C}^n$ is an orthonormal basis for the standard hermitian product.

Remark 4.36. Note that the representing matrix of a hermitian product on a finite-dimensional space in an orthonormal basis is the identity matrix. We will see (Theorem 4.41) that every finite-dimensional space admits an orthonormal basis, so eventually we always go back to the case of the standard hermitian product.

If $v = \sum_i v^i e_i$ is in the span of an orthonormal system $(e_i)_{i\in S}$, we can get the coefficients of the expansion by the formula

$$v^i = \langle e_i, v\rangle\,. \tag{4.7}$$

Unlike in the case of the inner product, we now have to be careful on the order of the arguments, as we have

$$\bar{v}^i = \langle v, e_i\rangle\,.$$

We can also rewrite the expansion as

$$v = \sum_i \langle e_i, v\rangle\, e_i.$$

Moreover, note that, if $w = \sum_i w^i e_i$, then $\langle v, w\rangle = \sum_{ij} \bar{v}^i w^j \langle e_i, e_j\rangle$, so

$$\langle v, w\rangle = \sum_i \bar{v}^i w^i = \sum_i \langle v, e_i\rangle \langle e_i, w\rangle\,,$$

and, in particular,

$$\|v\|^2 = \sum_i |v^i|^2. \tag{4.8}$$

Remark 4.37 (Dirac's notation). The above formulae have a nice rewriting in terms of Dirac's bra–ket notation introduced in Section 4.2.1. The expansion formula for a vector v in the orhonormal basis (e_i), e.g., reads

$$|\,v\rangle = \sum_i |\,e_i\rangle\,\langle\, e_i\,|\,v\,\rangle\,, \tag{4.9}$$

and the bracket of two vectors is now

$$\langle\, v\,|\,w\,\rangle = \sum_i \langle\, v\,|\,e_i\,\rangle\,\langle\, e_i\,|\,w\,\rangle\,.$$

The nice mnemonic rule stemming from these formulae is that the expression $|\,e_i\rangle\langle e_i\,|$, summed over i, may be freely inserted or omittted in other expressions.[8] Another practical advantage of the bra–ket notation is that it allows for the shorthand notation $|\,i\rangle$ for $|\,e_i\rangle$. With it the above formulae simply read

$$|\,v\rangle = \sum_i |\,i\rangle\,\langle\, i\,|\,v\,\rangle \qquad \text{and} \qquad \langle\, v\,|\,w\,\rangle = \sum_i \langle\, v\,|\,i\,\rangle\,\langle\, i\,|\,w\,\rangle\,.$$

Remark 4.38 (Bessel's inequality). As in the case of the inner product, see Remark 3.37, we may apply (4.7) also to the case when the orthonormal system is not a basis. In particular, with the same proof as in the inner product case, we get again Bessel's inequality

$$\sum_{i=1}^{\infty} |v_i|^2 \leq \|v\|^2, \tag{4.10}$$

with

$$v^i = \langle e_i, v\rangle$$

where $(e_i)_{i\in\mathbb{N}_{>0}}$ is an orthonormal system.

Example 4.39 (Fourier series). Consider the complex vector space

$$V := \{\phi \in C^0([0,L],\mathbb{C}) \mid \phi(0) = \phi(L)\}$$

[8]The mathematical reason is that $(|\,e_i\rangle)$ denote an orthonormal basis of V, whereas $(\langle e_i\,|)$ denotes the corresponding dual basis of V^*. The term $|\,e_i\rangle\langle e_i\,|$ is then a dual vector (i.e., a linear map $V \to \mathbb{C}$) times a vector: this corresponds to a linear map $V \to V$, $|\,v\rangle \mapsto |\,e_i\rangle\,\langle\, e_i\,|\,v\,\rangle$. By (4.9), we have that the sum over i produces the identity operator:

$$\sum_i |\,e_i\rangle\langle e_i\,| = \mathrm{Id}\,.$$

of complex-valued continuous periodic functions on the interval $[0, L]$ with hermitian product

$$\langle f, g\rangle = \int_0^L \bar{f} g \,\mathrm{d}x.$$

Consider

$$e_k(x) := \frac{1}{\sqrt{L}}\,\mathrm{e}^{\frac{2\pi \mathrm{i} k x}{L}}$$

which belongs to V for every $k \in \mathbb{Z}$. We have

$$\langle e_k, e_k\rangle = \frac{1}{L}\int_0^L 1 \,\mathrm{d}x = 1,$$

for every k, and

$$\langle e_k, e_l\rangle = \frac{1}{L}\int_0^L \mathrm{e}^{\frac{2\pi \mathrm{i}(l-k)x}{L}} \,\mathrm{d}x = \frac{1}{2\pi \mathrm{i}(l-k)}\,\mathrm{e}^{\frac{2\pi \mathrm{i} k x}{L}}\Big|_0^L = 0,$$

for every $k \neq l$ in $\mathbb{Z}$. Therefore, $(e_k)_{k\in\mathbb{Z}}$ is an orthonormal system. Using Dirac's notation, and writing $|\,k\rangle$ instead of $|\,e_k\rangle$, we get the Fourier coefficients

$$f_k = \langle\, k \,|\, f\,\rangle = \frac{1}{\sqrt{L}}\int_0^L \mathrm{e}^{-\frac{2\pi \mathrm{i} k x}{L}} f \,\mathrm{d}x \tag{4.11}$$

for every $f \in V$. As $\mathbb{Z}$ is a countable set (i.e., it is isomorphic to $\mathbb{N}$), we still have Bessel's inequality, which now reads

$$\sum_{k\in\mathbb{Z}} |f_k|^2 \leq \|f\|^2. \tag{4.12}$$

In this particular case one can show—but this is beyond the scope of these notes—that Bessel's inequality is actually saturated:

$$\sum_{k\in\mathbb{Z}} (f_k)^2 = \|f\|^2. \tag{4.13}$$

This equality is known as Parseval's identity. Note that, if f is in the span of the e_ks, we then have

$$f(x) = \frac{1}{\sqrt{L}}\sum_k f_k\,\mathrm{e}^{\frac{2\pi \mathrm{i} k x}{L}}. \tag{4.14}$$

However, even if $f \in V$ is not in the span of the e_ks, it turns out that the series (4.14) actually converges, in an appropriate sense, to the original function f. This is an example of a Fourier series. Finally, note that working with complex-valued functions has several advantages, one being the easier way to show that (e_k) is orthonormal as compared to the case of sine or cosine series. This theory may also be used for real-valued functions, just regarded as a special case of complex-valued ones. The only observation regarding the Fourier coefficients is that f is real valued iff $f_{-k} = \bar{f}_k$ for every k.

4.4.1. The orthogonal projection. The definition and properties of the orthogonal projection go exactly as in the case of the inner product.

If w is a nonzero vector, we can uniquely decompose any vector v as $v = v_\| + v_\perp$ with $v_\| \in \mathbb{C}w$ and $v_\perp \perp w$ by

$$v_\| = \langle w, v\rangle \frac{w}{\|w\|^2} \quad \text{and} \quad v_\perp = v - v_\|.$$

Again we define the projections

$$P_w v = \langle w, v\rangle \frac{w}{\|w\|^2} \quad \text{and} \quad P'_w = \mathrm{Id} - P_w.$$

Their properties are again the following:

$$P_w^2 = P_w, \quad P_w'^2 = P'_w,$$

$$\operatorname{im} P_w = \mathbb{C}w, \quad \operatorname{im} P'_w = w^\perp = \{v \in V \mid v \perp w\},$$

and, for every $\lambda \in \mathbb{C} \setminus \{0\}$,

$$P_{\lambda w} = P_w, \quad P'_{\lambda w} = P'_w.$$

In particular, if w is normalized (i.e., $\|w\| = 1$), then we get the simpler formula

$$P_w v = \langle w, v\rangle\, w,$$

or, in Dirac's notation,

$$P_w v = |\, w\rangle \langle\, w \,|\, v\, \rangle\,.$$

4.4.2. The Gram–Schmidt process. The Gram–Schmidt process works in exactly the same way as in the case of the inner product. By the same proof, we get the following

PROPOSITION 4.40 (Gram–Schmidt process). *Let $(v_1, \dots, v_k)$ be a linearly independent system in a hermitian product space V. Then there is an orthonormal system $(e_1, \dots, e_k)$ with the same span. This system* *is determined by the following process:*

$$\begin{aligned}
\widetilde{v}_1 &:= v_1 & e_1 &:= \frac{\widetilde{v}_1}{\|\widetilde{v}_1\|}\\
\widetilde{v}_2 &:= v_2 - \langle e_1, v_2\rangle\, e_1 & e_2 &:= \frac{\widetilde{v}_2}{\|\widetilde{v}_2\|}\\
\widetilde{v}_3 &:= v_3 - \langle e_1, v_3\rangle\, e_1 - \langle e_2, v_3\rangle\, e_2 & e_3 &:= \frac{\widetilde{v}_3}{\|\widetilde{v}_3\|}\\
&\ \vdots & &\ \vdots\\
\widetilde{v}_k &:= v_k - \sum_{i=1}^{k-1} \langle e_i, v_k\rangle\, e_i & e_k &:= \frac{\widetilde{v}_k}{\|\widetilde{v}_k\|}
\end{aligned}$$

Again, if $(v_1, \dots, v_n)$ is a basis of V, then the Gram–Schmidt process yields an orthonormal basis $(e_1, \dots, e_n)$. Therefore, we have the

THEOREM 4.41 (Orthonormal bases). *A finite-dimensional or countably infinite-dimensional hermitian product space has an orthonormal basis.*

If $V = \mathbb{C}^n$ with hermitian product $\langle \boldsymbol{v}, \boldsymbol{w}\rangle = \bar{\boldsymbol{v}}^\mathsf{T} \boldsymbol{g}\boldsymbol{w}$, where $\boldsymbol{g}$ is a positive definite self-adjoint matrix, the elements of an orthonormal basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ satisfy

$$\bar{\boldsymbol{v}}_i^\mathsf{T} \boldsymbol{g}\boldsymbol{v}_j = \delta_{ij}.$$

Therefore, the invertible matrix $\boldsymbol{F}$ whose columns are these basis vectors satisfies

$$\boldsymbol{F}^\dagger \boldsymbol{g}\boldsymbol{F} = \begin{pmatrix} \bar{\boldsymbol{v}}_1^\mathsf{T} \\ \vdots \\ \bar{\boldsymbol{v}}_n^\mathsf{T} \end{pmatrix} \boldsymbol{g} \left(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\right) = \begin{pmatrix} \bar{\boldsymbol{v}}_1^\mathsf{T} \\ \vdots \\ \bar{\boldsymbol{v}}_n^\mathsf{T} \end{pmatrix} \left(\boldsymbol{g}\boldsymbol{v}_1, \dots, \boldsymbol{g}\boldsymbol{v}_n\right) = \mathbf{1},$$

so $\boldsymbol{g} = \boldsymbol{F}^{\dagger,-1}\boldsymbol{F}^{-1}$. By setting $\boldsymbol{E} = \boldsymbol{F}^{-1}$, we get the factorization $\boldsymbol{g} = \boldsymbol{E}^\dagger \boldsymbol{E}$. Note on the other hand that, if we have a matrix $\boldsymbol{g}$ of this form, then $\boldsymbol{g}$ is self adjoint and positive definite, for $\bar{\boldsymbol{v}}^\mathsf{T} \boldsymbol{g}\boldsymbol{v} = \overline{\boldsymbol{E}\boldsymbol{v}}^\mathsf{T} \boldsymbol{E}\boldsymbol{v}$. We then have the

COROLLARY 4.42. *A self-adjoint matrix $\boldsymbol{g}$ is positive definite iff it is of the form $\boldsymbol{g} = \boldsymbol{E}^\dagger \boldsymbol{E}$ with $\boldsymbol{E}$ an invertible matrix.*

REMARK 4.43. The matrix $\boldsymbol{E}$ in Corollary 4.42 is not uniquely determined (as we can choose different orthonormal bases). In particular, suppose that $\boldsymbol{E}'$ is also an invertible matrix with $\boldsymbol{g} = \boldsymbol{E}'^\dagger \boldsymbol{E}'$. Then we have $\boldsymbol{E}'^\dagger \boldsymbol{E}' = \boldsymbol{E}^\dagger \boldsymbol{E}$, or, equivalently, $\boldsymbol{E}'\boldsymbol{E}^{-1} = \boldsymbol{E}'^{\dagger,-1}\boldsymbol{E}^\dagger$. Since

$\boldsymbol{E}'^{\dagger,-1}\boldsymbol{E}^{\dagger} = (\boldsymbol{E}'\boldsymbol{E}^{-1})^{\dagger,-1}$, we get that the invertible matrix $\boldsymbol{U} := \boldsymbol{E}'\boldsymbol{E}^{-1}$ satisfies

$$\boldsymbol{U}^{\dagger} = \boldsymbol{U}^{-1}.$$

A matrix with this property is called unitary (see more on this in Section 4.5). Note that $\boldsymbol{E}' = \boldsymbol{U}\boldsymbol{E}$. In conclusion, any two matrices occurring in a factorization of the same positive-definite matrix are related by a unitary matrix.

Note that $\det \boldsymbol{E}^{\dagger}\boldsymbol{E} = |\det \boldsymbol{E}|^2$, which is positive if $\boldsymbol{E}$ is invertible, so we have the

LEMMA 4.44. *A positive-definite self-adjoint matrix necessarily has positive determinant.*

Exactly as in the case of inner products, and with the same proof, we then get the following

COROLLARY 4.45. *If $\boldsymbol{g}$ is a positive definite matrix, then all its leading principal minors are necessarily positive.*

The converse to Corollary 4.45 also holds. (For the proof, very similar to that to Lemma 3.53, see exercise 4.5.)

LEMMA 4.46. *If all the leading principal minors of a self-adjoint matrix $\boldsymbol{g}$ are positive, then $\boldsymbol{g}$ is positive definite.*

We can summarize the results of Corollary 4.45 and of Lemma 4.46 as the following

THEOREM 4.47 (Sylvester's criterion). *A self-adjoint matrix is positive definite iff all its leading principal minors are positive.*

4.4.3. Orthogonal complements. The orthogonal subspace associated to a subspace W of a hermitian product space $(V, \langle\ ,\ \rangle)$ is defined exactly as in the the case of the inner product:

$$W^{\perp} := \{v \in V \mid v \perp w\ \forall w \in W\}.$$

The orthogonal space has exactly the same properties, with the same proofs, as in Proposition 3.58:

(1) $\{0\}^{\perp} = V$.
(2) $V^{\perp} = \{0\}$.
(3) $W \cap W^{\perp} = \{0\}$.
(4) $W \subseteq Z \implies Z^{\perp} \subseteq W^{\perp}$.
(5) $W \subseteq W^{\perp\perp}$.
(6) $W^{\perp\perp\perp} = W^{\perp}$.

Moreover, again by the same proof as in the case of Proposition 3.59, we have the following

PROPOSITION 4.48. *If W is a finite-dimensional subspace of a hermitian product space V, then $W^\perp$ is a complement of W, called the* orthogonal complement:

$$V = W \bigoplus W^\perp.$$

In particular, this implies the following

THEOREM 4.49. *Let V be a finite-dimensional hermitian product space. Then, for every subspace W we have the* orthogonal complement *$W^\perp$. In particular,*

$$\dim W + \dim W^\perp = \dim V,$$

and

$$W^{\perp\perp} = W.$$

Again we say that two subspaces W_1 and W_2 of a hermitian product space V are orthogonal if every vector in W_1 is orthogonal to every vector in W_2. In this case we write

$$W_1 \perp W_2.$$

DEFINITION 4.50. Let $(W_i)_{i\in S}$ be a collection of subspaces of a hermitian product space V. The collection is called orthogonal if

$$W_i \perp W_j \text{ for all } i \neq j.$$

PROPOSITION 4.51. *If $(W_i)_{i\in S}$ is an orthogonal collection of subspaces, then the sum of the W_is is direct.*

This is proved exactly as in the case of Proposition 3.64.

DEFINITION 4.52. If $(W_i)_{i\in S}$ is an orthogonal collection of subspaces of V and their sum is the whole of V, then

$$V = \bigoplus_{i\in S} W_i$$

is called an orthogonal decomposition of V.

REMARK 4.53. Suppose we have an orthogonal decomposition $V = \bigoplus_{i\in S} W_i$. Let P_i denote the projection to the W_i-component. As in Remark 3.67, we may prove that

$$\langle P_i v, v' \rangle = \langle v, P_i v' \rangle$$

for all $v, v' \in V$. This is an example of a self-adjoint operator (more on this in Section 4.5.3).

4.5. Unitary operators

An endomorphism F of a hermitian product space V is called a unitary operator if

$$\langle Fv, Fw\rangle = \langle v, w\rangle$$

for every $v, w \in V$.

EXAMPLE 4.54 (Unitary matrices). In the case of the standard hermitian product on $\mathbb{C}^n$, the endomorphism defined by an $n\times n$ complex matrix $\boldsymbol{A}$ is unitary iff $\bar{\boldsymbol{v}}^{\mathsf{T}}\boldsymbol{w} = \overline{\boldsymbol{A}\boldsymbol{v}}^{\mathsf{T}}\boldsymbol{A}\boldsymbol{w} = \bar{\boldsymbol{v}}^{\mathsf{T}}\boldsymbol{A}^\dagger\boldsymbol{A}\boldsymbol{w}$ for all $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{C}^n$. Taking $\boldsymbol{v} = \boldsymbol{e}_i$ and $\boldsymbol{w} = \boldsymbol{e}_j$, the condition implies $(\boldsymbol{A}^\dagger\boldsymbol{A})_{ij} = \delta_{ij}$ for all $i, j = 1, \dots, n$. Therefore, we see that the endomorphism defined by $\boldsymbol{A}$ is unitary iff

$$\boldsymbol{A}^\dagger\boldsymbol{A} = \mathbf{1}.$$

A matrix satisfying this identity is called a unitary matrix. Note that the condition implies that $\boldsymbol{A}$ is invertible and also that $\boldsymbol{A}^{-1}$ is itself a unitary matrix. Moreover, if $\boldsymbol{A}$ and $\boldsymbol{B}$ are unitary, then so is their product $\boldsymbol{AB}$.

REMARK 4.55. In particular, we have that an endomorphism of a finite-dimensional hermitian product space is unitary iff its representing matrix in any orthonormal basis is unitary.

REMARK 4.56. Let $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ be the columns of an $n\times n$ matrix $\boldsymbol{A}$. Then $(\bar{\boldsymbol{v}}_1^{\mathsf{T}}, \dots, \bar{\boldsymbol{v}}_n^{\mathsf{T}})$ are the rows of $\boldsymbol{A}^\dagger$. We then see that

$$\boldsymbol{A} \text{ is unitary } \iff (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n) \text{ is an orthonormal system.}$$

Here is a useful characterization of unitary operators:

THEOREM 4.57. *Let F be an endomomorphism of a hermitian product space V. Then the following are equivalent:*

(1) F is unitary.
(2) F preserves all norms.

PROOF. If F is unitary, then in particular $\|Fv\|^2 = \langle Fv, Fv\rangle = \langle v, v\rangle = \|v\|^2$ for every v, so F preserves all norms.

The reversed implication is obtained by making use of the polarization identity

$$\langle v, w\rangle = \frac{1}{4}(\|v+w\|^2 - \|v-w\|^2 - \mathrm{i}|v+\mathrm{i}w\|^2 + \mathrm{i}\|v-\mathrm{i}w\|^2), \quad \forall v, w,$$

whose proof is left as an exercise. □

Condition (2) implies that a unitary operator is injective. This has the following immediate corollary.

COROLLARY 4.58. *A unitary operator on a finite-dimensional hermitian product space is invertible.*

One can easily prove that the composition of two unitary operators is a unitary operator, that the inverse of a unitary operator is a unitary operator, and that the identity map is a unitary operator. As a consequence of this and of Corollary 4.58, we have the

PROPOSITION 4.59. *The set $\mathrm{U}(V)$ of unitary operators on a finite-dimensional hermitian product space V is a group, called the* unitary group *of V.*

REMARK 4.60. In the case of $\mathbb{C}^n$ with the standard hermitian product, we write $\mathrm{U}(n)$ for the corresponding group of unitary matrices

$$\mathrm{U}(n) = \{\boldsymbol{A} \in \mathrm{Mat}_{n\times n}(\mathbb{C}) \mid \boldsymbol{A}^\dagger \boldsymbol{A} = \mathbf{1}\},$$

called the unitary group.

REMARK 4.61. The characterization of unitary matrices of Example 4.54 extends to operators. Namely, assume V to be finite-dimensional (or, more generally, F to be invertible). Then F is unitary iff

$$F^\dagger F = \mathrm{Id},$$

where we have used the notion of adjoint operator introduced in Section 4.2.2. In fact, if F is unitary, then so is its inverse. Therefore,

$$\langle v, Fw\rangle = \left\langle F^{-1}v, F^{-1}Fw\right\rangle = \left\langle F^{-1}v, w\right\rangle,$$

which shows $F^\dagger = F^{-1}$. On the other hand, if $F^\dagger$ is the inverse of F, then

$$\langle Fv, Fw\rangle = \left\langle F^\dagger Fv, w\right\rangle = \left\langle F^{-1}Fv, w\right\rangle = \langle v, w\rangle,$$

so F is unitary.

4.5.1. Isometries. The notion of unitary operator may be generalized to linear maps between different spaces. Let $(V, \langle\ ,\ \rangle_V)$ and $(W, \langle\ ,\ \rangle_W)$ be hermitian product spaces. A linear map $F\colon V \to W$ is called an isometry if

$$\langle Fv_1, Fv_2\rangle_W = \langle v_1, v_2\rangle_V$$

for all $v_1, v_2 \in V$. Following verbatim the proof of Theorem 4.57, we see that F is an isometry iff it preserves all norms. We also see that an isometry is always injective. If V and W are finite-dimensional, we then have $\dim V \leq \dim W$, and F is an isomorphisms iff $\dim V = \dim W$.

EXAMPLE 4.62. The inclusion map of a subspace, with the restriction of the hermitian product as in Example 4.8, is an isometry.

EXAMPLE 4.63 (Fourier series and spaces of sequences). Consider the space

$$V := \{\phi \in C^0([0,L],\mathbb{C}) \mid \phi(0) = \phi(L)\}$$

of Example 4.39 with the orthonormal basis provided there. To each function $f \in V$ we may assign the complex sequence

$$(\dots, f_{-1}, f_0, f_1, f_2, \dots)$$

with $f_k := \langle e_k, f\rangle$ as in (4.11). For a general $f \in V$, infinitely many f_ks may be different from zero. Still we have Bessel's inequality (4.12) which shows that $(\dots, f_{-1}, f_0, f_1, f_2, \dots)$ is a square-summable sequence (see Section 4.3.2).[9] Therefore, we have a linear map

$$\begin{array}{rccc} F\colon & \{\phi \in C^0([0,L],\mathbb{C}) \mid \phi(0) = \phi(L)\} & \to & \ell^2_{\mathbb{C}} \\ & f & \mapsto & \left(\frac{1}{\sqrt{L}}\int_0^L \mathrm{e}^{-\frac{2\pi \mathrm{i} k x}{L}} f \, \mathrm{d}x\right)_{k\in\mathbb{Z}} \end{array}$$

Thanks to Parseval's identity (4.13), F is actually an isometry. It is not surjective though.[10]

EXAMPLE 4.64. If $(e_1, \dots, e_n)$ is an orthonormal basis of V, then the linear map $F\colon V \to \mathbb{C}^n$ that assigns to a vector v the column vector with components its coefficients $v^i = \langle e_i, v\rangle$ is a bijective isometry (in the notations of Remark 1.52, $F = \Phi_{\mathcal{B}}^{-1}$ with $\mathcal{B} = (e_1, \dots, e_n)$).

From this observation and from Theorem 4.41, we get the

THEOREM 4.65. *Every n-dimensional hermitian product space possesses a bijective isometry with $\mathbb{C}^n$ endowed with the standard hermitian product.*

4.5.2. The unitary groups. In this section we analyze the group $\mathrm{U}(n)$ of unitary matrices, introduced in Remark 4.60, in particular for $n = 1$ and $n = 2$.

Note that, for every complex square matrix, $\det \boldsymbol{A}^\dagger = \overline{\det \boldsymbol{A}}$. Therefore, the condition $\boldsymbol{A}^\dagger \boldsymbol{A} = \mathbf{1}$ for a unitary matrix implies $|\det \boldsymbol{A}|^2 = 1$. This means that $\det \boldsymbol{A}$ is of the form $\mathrm{e}^{\mathrm{i}\theta}$ for some $\theta \in \mathbb{R}$.

Unitary matrices with determinant 1 form a subgroup of $\mathrm{U}(n)$ called the special unitary group:

$$\mathrm{SU}(n) := \{\boldsymbol{A} \in \mathrm{Mat}_{n\times n}(\mathbb{C}) \mid \boldsymbol{A}^\dagger \boldsymbol{A} = \mathbf{1} \text{ and } \det \boldsymbol{A} = 1\}.$$

[9] To agree with the notation there, we have to relabel the indices via an isomorphism $\mathbb{N} \to \mathbb{Z}$.

[10] We may enlarge V to contain noncontinuous square-integrable functions. All the above works. By the injectivity of the isometry F, we then get square-summable sequences of coefficients beyond those coming from continuous functions.

REMARK 4.66. The groups U(1), SU(2), and SU(3) are particularly important in physics: SU(2) is related to spin in quantum mechanics, $\mathrm{U}(1) \times \mathrm{SU}(2)$ to the electroweak interaction, and SU(3) to the strong interaction.

REMARK 4.67. If $\boldsymbol{A}$ is an $n \times n$ unitary matrix with determinant $\mathrm{e}^{\mathrm{i}\theta}$, then $\boldsymbol{B} := \mathrm{e}^{-\frac{\mathrm{i}\theta}{n}}\boldsymbol{A}$ is also unitary, but now with determinant equal to 1. Therefore, every $n \times n$ unitary matrix $\boldsymbol{A}$ can be written as $\lambda\boldsymbol{B}$ with $|\lambda| = 1$ and $\boldsymbol{B} \in \mathrm{SU}(n)$.

4.5.2.1. *The group* U(1). A matrix $\boldsymbol{A} \in \mathrm{U}(1)$ is of the form $\boldsymbol{A} = (\lambda)$ with $|\lambda| = 1$. Therefore, we have the

PROPOSITION 4.68. *A matrix in* U(1) *has the form* $(\mathrm{e}^{\mathrm{i}\theta})$. *The angle* θ *is uniquely determined if we take it in the interval* $\theta \in [0, 2\pi)$. *Geometrically,* U(1) *is the unit circle* S^1 *in the complex plane:*

$$\mathrm{U}(1) = S^1 = \{\lambda \in \mathbb{C} \mid |\lambda| = 1\}.$$

REMARK 4.69. If $z = \rho\mathrm{e}^{\mathrm{i}\alpha} \in \mathbb{C}$, then $\mathrm{e}^{\mathrm{i}\theta}z = \rho\mathrm{e}^{\mathrm{i}(\alpha+\theta)}$, so the group U(1) may be interpreted as the group of rotations on the complex plane centered at the origin.

As we observed in Remark 3.88, also the group SO(2) acts by rotations on the plane. Actually, U(1) and SO(2) are essentially the same group, as follows from the normal form of matrices in SO(2) given in Proposition 3.87. Namely, we have the following

PROPOSITION 4.70. *There is a group isomorphism*

$$\begin{array}{ccc} \mathrm{U}(1) & \to & \mathrm{SO}(2) \\ \mathrm{e}^{\mathrm{i}\theta} & \mapsto & \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \end{array}$$

We leave the simple details of the proof to the reader.

4.5.2.2. *The group* SU(2). We have seen that, geometrically, the group U(1) is the same as the unit circle in the plane. Our goal is to show the higher-dimensional analogue for SU(2).

PROPOSITION 4.71. *A matrix* $\boldsymbol{A}$ *in* SU(2) *has the form*

$$\boldsymbol{A} = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}$$

with $|\alpha|^2 + |\beta|^2 = 1$. *Therefore, geometrically, the group* SU(2) *is the same as the three-dimensional unit sphere* S^3 *in* $\mathbb{R}^4$.

Proof. Consider a 2×2 complex matrix $\boldsymbol{A} = \left(\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right)$ with determinant 1. By (1.8), its inverse is then $\left(\begin{smallmatrix}\delta & -\beta\\ -\gamma & \alpha\end{smallmatrix}\right)$. Equating it to its adjoint $\left(\begin{smallmatrix}\bar\alpha & \bar\gamma\\ \bar\beta & \bar\delta\end{smallmatrix}\right)$ yields $\bar\delta = \alpha$ and $\gamma = -\bar\beta$.

As a consequence, our matrix $\boldsymbol{A}$ has the form $\left(\begin{smallmatrix}\alpha & \beta\\ -\bar\beta & \bar\alpha\end{smallmatrix}\right)$, and its determinant, which we have to equate to 1, is $|\alpha|^2+|\beta|^2$. We have thus proved the first statement.

Next, if we identify $\mathbb{C}^2$ with $\mathbb{R}^4$ by taking the real and imaginary parts of α and β,[11]

$$\alpha = x^0 + \mathrm{i}x^3 \quad \text{and} \quad \beta = x^2 + \mathrm{i}x^1,$$

we see that the equation $|\alpha|^2 + |\beta|^2 = 1$, for $(\alpha,\beta)\in\mathbb{C}^2$, defines

$$S^3 = \{(x^0,x^1,x^2,x^3)\in\mathbb{R}^4 \mid (x^0)^2+(x^1)^2+(x^2)^2+(x^3)^2 = 1\},$$

which is the unit sphere. □

4.5.3. Self-adjoint and anti-self-adjoint operators. We conclude with a related concept.

Definition 4.72. An endomorphism F of a hermitian product space V is called self adjoint if

$$\langle Fv, w\rangle = \langle v, Fw\rangle$$

for every $v, w \in V$.

Remark 4.73. By the notion of adjoint operator introduced in Section 4.2.2, we immediately have that F is self adjoint iff its adjoint operator exists and

$$F^\dagger = F.$$

Example 4.74. In the case of the standard hermitian product on $\mathbb{C}^n$, the endomorphism defined by an $n\times n$ matrix $\boldsymbol{A}$ is self adjoint iff the matrix $\boldsymbol{A}$ is self adjoint (see Definition 4.11), i.e., $\boldsymbol{A}^\dagger = \boldsymbol{A}$. In particular, we have that an endomorphism of a finite-dimensional hermitian product space is self adjoint iff its representing matrix in some orthonormal basis is self adjoint.

Example 4.75. If we have an orthogonal decomposition

$$V = \bigoplus_{i\in S} W_i,$$

then the projection P_i to the W_i-component is self adjoint, as shown in Remark 4.53.

[11]The reason for naming the real and imaginary parts this way is to conform with a notation that we will introduce in Remark 4.85.

Self-adjoint operators are at the core of quantum mechanics, where they are used to describe physical observables.

DEFINITION 4.76. An endomorphism F of a hermitian product space V is called anti-self-adjoint or antihermitian if

$$\langle Fv, w\rangle = -\langle v, Fw\rangle$$

for every $v, w \in V$.

REMARK 4.77. By the notion of adjoint operator introduced in Section 4.2.2, we immediately have that F is anti-self-adjoint iff its adjoint operator exists and

$$F^\dagger = -F.$$

EXAMPLE 4.78. In the case of the standard hermitian product on $\mathbb{C}^n$, the endomorphism defined by an $n \times n$ matrix $\boldsymbol{A}$ is anti-self-adjoint iff the matrix $\boldsymbol{A}$ is anti-self-adjoint, i.e., $\boldsymbol{A}^\dagger = -\boldsymbol{A}$. In particular, we have that an endomorphism of a finite-dimensional inner product space is anti-self-adjoint iff its representing matrix in some orthonormal basis is anti-self-adjoint.

REMARK 4.79. Self-adjoint and anti-self-adjoint operators are closely related. Namely, suppose F is self adjoint. Then iF is anti-self-adjoint. In fact

$$\langle \mathrm{i}Fv, w\rangle = -\mathrm{i}\,\langle Fv, w\rangle = -\mathrm{i}\,\langle v, Fw\rangle = -\langle v, \mathrm{i}Fw\rangle\,.$$

Similarly, if F is anti-self-adjoint, then iF is self adjoint.

Anti-self-adjoint operators are also closely related to unitary operators. One relation is the following. Suppose $\boldsymbol{U}(t)$ is a differentiable map $\mathbb{R} \to \mathrm{U}(n)$. Then, differentiating the identity $\boldsymbol{U}^\dagger\boldsymbol{U} = \mathbf{1}$, we get $\dot{\boldsymbol{U}}^\dagger\boldsymbol{U} + \boldsymbol{U}^\dagger\dot{\boldsymbol{U}} = \mathbf{0}$. Therefore, the matrix $\boldsymbol{A} := \dot{\boldsymbol{U}}\boldsymbol{U}^{-1}$ is anti-self-adjoint.

Another relation is the following. Let $\boldsymbol{A}$ be anti-self-adjoint. Define $\boldsymbol{U}(t) := \mathrm{e}^{\boldsymbol{A}t}$. Then we have $\boldsymbol{U}^\dagger = \boldsymbol{U}^{-1}$, so $\boldsymbol{U}(t)$ is unitary for every t. Therefore, $\mathrm{e}^{\boldsymbol{A}t}$ is a differentiable map $\mathbb{R} \to \mathrm{U}(n)$. By Proposition 2.11, $\det \boldsymbol{U}(t) = \mathrm{e}^{t \operatorname{tr} \boldsymbol{A}}$ for every t. As a consequence, $\mathrm{e}^{\boldsymbol{A}t}$ is a differentiable map $\mathbb{R} \to \mathrm{SU}(n)$ iff the trace of $\boldsymbol{A}$ vanishes.

REMARK 4.80. The real vector space[12] of anti-self-adjoint $n \times n$ matrices is denoted by $\mathfrak{u}(n)$; its subspace of traceless matrices is denoted

[12] A linear combination with real coefficients of anti-self-adjoint matrices is still anti-self-adjoint. The same does not hold if we allow complex coefficients.

by $\mathfrak{su}(n)$:

$$\mathfrak{u}(n) := \{\boldsymbol{A} \in \mathrm{Mat}_{n\times n}(\mathbb{C}) \mid \boldsymbol{A}^\dagger = -\boldsymbol{A}\},$$
$$\mathfrak{su}(n) := \{\boldsymbol{A} \in \mathrm{Mat}_{n\times n}(\mathbb{C}) \mid \boldsymbol{A}^\dagger = -\boldsymbol{A},\ \operatorname{tr}\boldsymbol{A} = 0\}.$$

This notation helps remembering that we have the exponential maps

$$\begin{aligned} \exp\colon\ \mathfrak{u}(n) &\to \mathrm{U}(n) \\ \boldsymbol{A} &\mapsto \mathrm{e}^{\boldsymbol{A}} \end{aligned}$$
$$\begin{aligned} \exp\colon\ \mathfrak{su}(n) &\to \mathrm{SU}(n) \\ \boldsymbol{A} &\mapsto \mathrm{e}^{\boldsymbol{A}} \end{aligned}$$

REMARK 4.81. We will see in Corollary 4.107 that for every n these exponential maps are surjective. In the next section we will only focus on the case of $\mathfrak{su}(2)$.

4.5.4. Digression: Pauli matrices. The Pauli matrices, introduced by W. PAULI to describe the spin of a particle in quantum mechanics, are a basis of the real vector space $\mathrm{i}\mathfrak{su}(2)$ of traceless self-adjoint 2×2 matrices (see Remark 4.79 to relate self-adjoint and anti-self-adjoint operators). They are the following three traceless self-adjoint 2×2 matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

PROPOSITION 4.82. *The Pauli matrices* $(\sigma_1, \sigma_2, \sigma_3)$ *form a basis of the real vector space* $\mathrm{i}\mathfrak{su}(2)$ *of traceless self-adjoint* 2×2 *matrices. Consequently,* $(\mathrm{i}\sigma_1, \mathrm{i}\sigma_2, \mathrm{i}\sigma_3)$ *is a basis of the real vector space* $\mathfrak{su}(2)$ *of traceless anti-self-adjoint* 2×2 *matrices.*

PROOF. Let $\boldsymbol{A} = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & -\alpha \end{smallmatrix}\right)$ be a traceless 2×2 complex matrix. Then $\boldsymbol{A}^\dagger = \left(\begin{smallmatrix} \bar\alpha & \bar\gamma \\ \bar\beta & -\bar\alpha \end{smallmatrix}\right)$. Therefore, $\boldsymbol{A}$ is self adjoint iff

$$\alpha = \bar\alpha \quad \text{and} \quad \beta = \bar\gamma.$$

The first equation says that α is real: we write $\alpha = x^3 \in \mathbb{R}$. If we denote by x^1 and x^2 the real and imaginary parts of γ—i.e., $\gamma = x^1 + \mathrm{i}x^2$—then the second equation says that $\beta = x^1 - \mathrm{i}x^2$. Therefore, $\boldsymbol{A}$ is self adjoint and traceless iff it is of the form

$$\boldsymbol{A} = \begin{pmatrix} x^3 & x^1 - \mathrm{i}x^2 \\ x^1 + \mathrm{i}x^2 & -x^3 \end{pmatrix} = x^1\sigma_1 + x^2\sigma_2 + x^3\sigma_3$$

with x^1, x^2, x^3 real. One also immediately checks that this decomposition is unique, so $(\sigma_1, \sigma_2, \sigma_3)$ is a basis of $\mathrm{i}\mathfrak{su}(2)$. The second statement is an immediate consequence of Remark 4.79. $\square$

REMARK 4.83. To get a basis of the real vector space $\mathrm{i}\mathfrak{u}(2)$ of self-adjoint 2×2 matrices we just have to add one more basis element, e.g., $\sigma_0 = \mathbf{1}$. We will not consider this here.

REMARK 4.84. The expansion of a traceless self-adjoint 2×2 matrix in the basis of Pauli matrices is usually denoted by

$$\boldsymbol{x} \cdot \boldsymbol{\sigma} := x^i \sigma_i = x^1\sigma_1 + x^2\sigma_2 + x^3\sigma_3. \tag{4.15}$$

REMARK 4.85. Matrices in SU(2) can also be written in terms of Pauli matrices. In fact, by Proposition 4.71 every $\boldsymbol{A} \in \mathrm{SU}(2)$ if of the form $\boldsymbol{A} = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}$ with $|\alpha|^2 + |\beta|^2 = 1$. Setting $\alpha = x^0 + \mathrm{i}x^3$ and $\beta = x^2 + \mathrm{i}x^1$, we have

$$\boldsymbol{A} = x^0\,\mathbf{1} + \mathrm{i}\,\boldsymbol{x} \cdot \boldsymbol{\sigma}$$

with $(x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2 = 1$. It is convenient to set $\boldsymbol{x} = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix}$. The condition then reads $(x^0)^2 + \|\boldsymbol{x}\|^2 = 1$, where $\|\ \|$ denotes the euclidean norm on $\mathbb{R}^3$. This means that there is an angle θ such that $x^0 = \cos\theta$ and $\|\boldsymbol{x}\| = \pm\sin\theta$. We can therefore write

$$\boldsymbol{A} = \cos\theta\,\mathbf{1} + \mathrm{i}\,\sin\theta\,\widehat{\boldsymbol{x}} \cdot \boldsymbol{\sigma}, \tag{4.16}$$

where $\widehat{\boldsymbol{x}}$ is a unit vector.

One immediatley checks that the square of each Pauli matrix is the identity matrix and that

$$\sigma_1\sigma_2 = \mathrm{i}\sigma_3 = -\sigma_2\sigma_1, \quad \sigma_2\sigma_3 = \mathrm{i}\sigma_1 = -\sigma_3\sigma_2, \quad \sigma_3\sigma_1 = \mathrm{i}\sigma_2 = -\sigma_1\sigma_3.$$

One may summarize all these identities as

$$\sigma_i\sigma_j = \delta_{ij}\mathbf{1} + \mathrm{i}\sum_{k=1}^{3} \epsilon_{ijk}\sigma_k,$$

where δ_{ij} is the Kronecker delta and ϵ_{ijk} is the Levi-Civita symbol defined as zero if one index is repeated and as the sign of the permutation $123 \mapsto ijk$ otherwise. Explicitly,

$$\begin{aligned} \epsilon_{123} = \epsilon_{231} = \epsilon_{312} &= \quad 1, \\ \epsilon_{132} = \epsilon_{213} = \epsilon_{321} &= -1, \\ \epsilon_{ijk} &= 0 \text{ otherwise.} \end{aligned}$$

Using the notation of (4.15), we then have, for every $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^3$,

$$(\boldsymbol{x} \cdot \boldsymbol{\sigma})(\boldsymbol{y} \cdot \boldsymbol{\sigma}) = \boldsymbol{x} \cdot \boldsymbol{y}\,\mathbf{1} + \mathrm{i}\,\boldsymbol{x} \times \boldsymbol{y} \cdot \boldsymbol{\sigma}, \tag{4.17}$$

where on the left hand side we use matrix multiplication and on the right hand side we use the dot and the cross product of vectors. This formula implies the following

LEMMA 4.86. *For every* $\boldsymbol{y} \in \mathbb{R}^3 \setminus \{0\}$, *we have*

$$\mathrm{e}^{\mathrm{i}\boldsymbol{y}\cdot\boldsymbol{\sigma}} = \cos\|\boldsymbol{y}\| + \mathrm{i}\,\frac{\sin\|\boldsymbol{y}\|}{\|\boldsymbol{y}\|}\,\boldsymbol{y}\cdot\boldsymbol{\sigma}.$$

PROOF. By (4.17) with $\boldsymbol{x} = \boldsymbol{y}$ we have $(\boldsymbol{y}\cdot\boldsymbol{\sigma})^2 = \|\boldsymbol{y}\|^2\mathbf{1}$. Therefore, setting $\boldsymbol{J} := \frac{\mathrm{i}\boldsymbol{y}\cdot\boldsymbol{\sigma}}{\|\boldsymbol{y}\|}$, we have $\boldsymbol{J}^2 = -\mathbf{1}$. This implies $\boldsymbol{J}^{2s} = (-1)^s\mathbf{1}$ and $\boldsymbol{J}^{2s+1} = (-1)^s\boldsymbol{J}$ for all $s \in \mathbb{N}$. We then get, for every $\alpha \in \mathbb{R}$,

$$\mathrm{e}^{\alpha\boldsymbol{J}} = \sum_{n=0}^{\infty}\frac{\alpha^n}{n!}\boldsymbol{J}^n = \sum_{s=0}^{\infty}(-1)^s\frac{\alpha^{2s}}{(2s)!}\,\mathbf{1} + \sum_{s=0}^{\infty}(-1)^s\frac{\alpha^{2s+1}}{(2s+1)!}\,\boldsymbol{J}.$$

Recognizing the power series for sine and cosine, we then have

$$\mathrm{e}^{\alpha\boldsymbol{J}} = \cos\alpha\,\mathbf{1} + \sin\alpha\,\boldsymbol{J}.$$

Finally,

$$\mathrm{e}^{\mathrm{i}\boldsymbol{y}\cdot\boldsymbol{\sigma}} = \mathrm{e}^{\|\boldsymbol{y}\|\boldsymbol{J}} = \cos\|\boldsymbol{y}\|\,\mathbf{1} + \sin\|\boldsymbol{y}\|\,\boldsymbol{J},$$

which proves the lemma. □

By (4.16) we then have the

COROLLARY 4.87. *The exponential map* $\mathfrak{su}(2) \to \mathrm{SU}(2)$ *is surjective.*

4.5.4.1. SU(2) *and space rotations.* There is a very strong relation between the group SU(2) and the group SO(3) of space rotations. This is at the core of the appearance of SU(2) in quantum mechanics to describe the spin of particles.

The central observation is the following. Let $\boldsymbol{A} \in \mathrm{SU}(2)$ and $\boldsymbol{B}$ be a 2×2 traceless hermitian matrix. Define

$$\Phi_{\boldsymbol{A}}\boldsymbol{B} := \boldsymbol{A}\boldsymbol{B}\boldsymbol{A}^\dagger.$$

It is readily verified that $\Phi_{\boldsymbol{B}}\boldsymbol{A}$ is also traceless and hermitian. That is, we have defined a linear map

$$\begin{array}{rccc}\Phi_{\boldsymbol{A}}\colon & \mathrm{i}\mathfrak{su}(2) & \to & \mathrm{i}\mathfrak{su}(2)\\ & \boldsymbol{B} & \mapsto & \boldsymbol{A}\boldsymbol{B}\boldsymbol{A}^\dagger\end{array}$$

For any two $\boldsymbol{A}, \boldsymbol{A}' \in \mathrm{SU}(2)$, we also clearly have

$$\Phi_{\boldsymbol{A}\boldsymbol{A}'}\boldsymbol{B} = \boldsymbol{A}\boldsymbol{A}'\boldsymbol{B}(\boldsymbol{A}\boldsymbol{A}')^\dagger = \Phi_{\boldsymbol{A}}(\Phi_{\boldsymbol{A}'}\boldsymbol{B}).$$

That is,

$$\Phi_{\boldsymbol{A}\boldsymbol{A}'} = \Phi_{\boldsymbol{A}}\Phi_{\boldsymbol{A}'},$$

where on the right hand side we use the composition of linear maps. Abstractly, we have the group homomorphism

$$\begin{array}{rccc} \Phi\colon & \mathrm{SU}(2) & \to & \mathrm{Aut}(\mathrm{i}\mathfrak{su}(2)) \\ & \boldsymbol{A} & \mapsto & \Phi_{\boldsymbol{A}} \end{array}$$

where $\mathrm{Aut}(\mathrm{i}\mathfrak{su}(2))$ denotes the group of automorphisms (i.e., invertible linear maps) of the real vector space $\mathrm{i}\mathfrak{su}(2)$. Using a basis—e.g., by the Pauli matrices—we may identify $\mathrm{i}\mathfrak{su}(2)$ with $\mathbb{R}^3$ and $\mathrm{Aut}(\mathrm{i}\mathfrak{su}(2))$ with the group of invertible 3×3 real matrices.

One can also readily prove that $\det(\Phi_{\boldsymbol{A}}\boldsymbol{B}) = \det \boldsymbol{B}$ for every $\boldsymbol{A} \in \mathrm{SU}(2)$. If we expand $\boldsymbol{B}$ in the basis of Pauli matrices,

$$\boldsymbol{B} = \boldsymbol{x} \cdot \boldsymbol{\sigma} = \begin{pmatrix} x^3 & x^1 - \mathrm{i}x^2 \\ x^1 + \mathrm{i}x^2 & -x^3 \end{pmatrix},$$

we see that $\det \boldsymbol{B} = (x^1)^2 + (x^2)^2 + (x^3)^2 = \|\boldsymbol{x}\|^2$. This shows that the representing matrix of $\Phi_{\boldsymbol{A}}$ in the basis of the Pauli matrices preserves the euclidean norm, so it is orthogonal (by Theorem 3.74). The composition Ψ of the group homomorphism Φ with the isomorphism $\phi^{-1}_{(\sigma_1,\sigma_2,\sigma_3)}\colon \mathrm{i}\mathfrak{su}(2) \to \mathbb{R}^3$ is therefore a group homomorphism $\Psi\colon \mathrm{SU}(2) \to \mathrm{O}(3)$. By definition we have

$$\Phi_{\boldsymbol{A}}(\boldsymbol{x} \cdot \boldsymbol{\sigma}) = (\Psi_{\boldsymbol{A}}\boldsymbol{x}) \cdot \boldsymbol{\sigma}. \tag{4.18}$$

Thanks to (4.17) we can explicitly write down Ψ and get the following

Theorem 4.88. *For*

$$\boldsymbol{A} = \cos\theta\,\mathbf{1} + \mathrm{i}\sin\theta\,\widehat{\boldsymbol{x}} \cdot \boldsymbol{\sigma},$$

where $\widehat{\boldsymbol{x}}$ *is a unit vector, we have*

$$\Psi_{\boldsymbol{A}} = \boldsymbol{R}_{\widehat{\boldsymbol{x}}}(-2\theta),$$

where $\boldsymbol{R}_{\boldsymbol{n}}(\alpha)$ *denotes the counterclockwise rotation by the angle* α *around the oriented principal axis generated by the vector* $\boldsymbol{n}$*. Consequently,* Ψ *defines a* <u>*surjective*</u> *group homomorphism*

$$\Psi\colon\ \mathrm{SU}(2) \to \mathrm{SO}(3).$$

Moreover, $\Psi_{\boldsymbol{A}} = \Psi_{\boldsymbol{A}'}$ *iff* $\boldsymbol{A} = \pm\boldsymbol{A}'$*, so the the preimage of each point in* $\mathrm{SO}(3)$ *consists of exactly two points in* $\mathrm{SU}(2)$*. (One says that* $\mathrm{SU}(2)$ *is a double cover of* $\mathrm{SO}(3)$*.)*

Proof. Take $\boldsymbol{A}$ as in the statement and $\boldsymbol{B} = \boldsymbol{y} \cdot \boldsymbol{\sigma}$. By (4.17), we get

$$\begin{aligned} \boldsymbol{A}\boldsymbol{B} &= \cos\theta\,\boldsymbol{y} \cdot \boldsymbol{\sigma} + \mathrm{i}\sin\theta\,\widehat{\boldsymbol{x}} \cdot \boldsymbol{y}\,\mathbf{1} - \sin\theta\,\widehat{\boldsymbol{x}} \times \boldsymbol{y} \cdot \boldsymbol{\sigma} \\ &= \mathrm{i}\sin\theta\,\widehat{\boldsymbol{x}} \cdot \boldsymbol{y}\,\mathbf{1} + (\cos\theta\,\boldsymbol{y} - \sin\theta\,\widehat{\boldsymbol{x}} \times \boldsymbol{y}) \cdot \boldsymbol{\sigma}. \end{aligned}$$

Since $\boldsymbol{A}^\dagger = \cos\theta\,\mathbf{1} - \mathrm{i}\sin\theta\,\widehat{\boldsymbol{x}}\cdot\boldsymbol{\sigma}$, we then get, after some simplifications,

$$\boldsymbol{A}\boldsymbol{B}\boldsymbol{A}^\dagger = (\cos^2\theta\,\boldsymbol{y} + \sin^2\theta\,(\widehat{\boldsymbol{x}}\cdot\boldsymbol{y}\,\widehat{\boldsymbol{x}} - (\widehat{\boldsymbol{x}}\times\boldsymbol{y})\times\widehat{\boldsymbol{x}}) - 2\sin\theta\cos\theta\,\widehat{\boldsymbol{x}}\times\boldsymbol{y})\cdot\boldsymbol{\sigma}.$$

We have

$$\widehat{\boldsymbol{x}}\cdot\boldsymbol{y}\,\widehat{\boldsymbol{x}} = \boldsymbol{y}_\parallel,$$

where we use the orthogonal decomposition $\boldsymbol{y} = \boldsymbol{y}_\parallel + \boldsymbol{y}_\perp$ with $\boldsymbol{y}_\parallel$ proportional to $\widehat{\boldsymbol{x}}$ and $\boldsymbol{y}_\perp$ orthogonal to it. A computation using the properties of the cross product (see exercise 3.16) shows

$$(\widehat{\boldsymbol{x}}\times\boldsymbol{y})\times\widehat{\boldsymbol{x}} = \boldsymbol{y}_\perp.$$

Therefore, using (4.18) and exercise 3.17,

$$\begin{aligned}\Psi_{\boldsymbol{A}}\boldsymbol{y} &= \cos^2\theta\,(\boldsymbol{y}_\parallel + \boldsymbol{y}_\perp) + \sin^2\theta\,(\boldsymbol{y}_\parallel - \boldsymbol{y}_\perp) - 2\sin\theta\cos\theta\,\widehat{\boldsymbol{x}}\times\boldsymbol{y}\\ &= \boldsymbol{y}_\parallel + \cos(2\theta)\,\boldsymbol{y}_\perp - \sin(2\theta)\,\widehat{\boldsymbol{x}}\times\boldsymbol{y}_\perp\\ &= \boldsymbol{R}_{\widehat{\boldsymbol{x}}}(-2\theta)\boldsymbol{y}.\end{aligned}$$

This proves the first part of the theorem.

The explicit formula also shows that the image of Ψ is the whole group $\mathrm{SO}(3)$. Moreover, if $\boldsymbol{A}' = \cos\theta'\,\mathbf{1} + \mathrm{i}\sin\theta'\,\widehat{\boldsymbol{x}}'\cdot\boldsymbol{\sigma}$ and $\Psi_{\boldsymbol{A}} = \Psi_{\boldsymbol{A}'}$, we have $\boldsymbol{R}_{\widehat{\boldsymbol{x}}}(-2\theta) = \boldsymbol{R}_{\widehat{\boldsymbol{x}}'}(-2\theta')$. This first implies that the two rotations have the same principal axis, so $\widehat{\boldsymbol{x}} = \pm\widehat{\boldsymbol{x}}'$. Actually, up to choosing θ' appropriately, we may assume $\widehat{\boldsymbol{x}} = \widehat{\boldsymbol{x}}'$. We then have the condition $2\theta = 2\theta' \mod 2\pi$, i.e., $\theta = \theta' \mod 2\pi$ or $\theta = \theta' + \pi \mod 2\pi$. In the first case, $\boldsymbol{A} = \boldsymbol{A}'$ and in the second $\boldsymbol{A} = -\boldsymbol{A}'$. □

Remark 4.89. As a final remark, note that by Proposition 4.71, we can identify $\mathrm{SU}(2)$ with the three-dimensional sphere S^3. Moreover, if $\boldsymbol{A}$ corresponds to a point $x \in S^3 \subset \mathbb{R}^4$, then $-\boldsymbol{A}$ corresponds to the antipodal point $-x$. The map Ψ therefore provides a surjective map $S^3 \to \mathrm{SO}(3)$ with the property that the preimage of each point in $\mathrm{SO}(3)$ consists of two antipodal points in S^3.

4.6. Diagonalization and normal form for some important classes of matrices

We will now apply hermitian products to diagonalize unitary and self-adjoint matrices. This will also lead to the diagonalization of real symmetric matrices and to a normal form for orthogonal and for real skew-symmetric matrices.

The three mentioned cases of real matrices can actually be reduced to the study of unitary or self-adjoint matrices.

Proposition 4.90. *Let $\boldsymbol{A}$ be an $n\times n$ real matrix, which we will regard as an $n\times n$ complex matrix. Then the following hold:*

(1) $\boldsymbol{A}$ *is orthogonal iff* $\boldsymbol{A}$ *is unitary.*
(2) $\boldsymbol{A}$ *is symmetric if iff* $\boldsymbol{A}$ *is self adjoint.*
(3) $\boldsymbol{A}$ *is skew-symmetric iff* $\mathrm{i}\boldsymbol{A}$ *is self-adjoint.*

Proof. The three statements simply follow from the fact that, $\boldsymbol{A}$ being real, we have $\overline{\boldsymbol{A}} = \boldsymbol{A}$, so $\boldsymbol{A}^\dagger = \boldsymbol{A}^\mathsf{T}$. For the last statement, we use $(\mathrm{i}\boldsymbol{A})^\dagger = -\mathrm{i}\boldsymbol{A}^\dagger = -\mathrm{i}\boldsymbol{A}^\mathsf{T}$. □

Finally, unitary and self-adjoint matrices are particular examples of the more general concept of normal matrices.

Definition 4.91 (Normal matrices). A complex matrix $\boldsymbol{A}$ is called normal if it commutes with its adjoint:

$$\boldsymbol{A}^\dagger \boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^\dagger.$$

Lemma 4.92. *Unitary and self-adjoint matrices are normal.*

Proof. This is readily verified and left as an exercise. □

We will therefore start with a discussion of normal matrices and their diagonalizability, then we will specialize this result on unitary and self-adoint matrices, and finally we will draw consequences for orthogonal, real symmetric and real skew-symmetric matrices.

4.6.1. Normal matrices and normal operators. We begin by extending the definition of normal matrices to operators.

Definition 4.93 (Normal operators). An endomorphism F of an inner product space $(V, \langle\ ,\ \rangle)$ is called normal if its adjoint, as defined in Section 4.2.2, exists and satisfies

$$F^\dagger F = F F^\dagger.$$

In the next two propositions we state some important properties of normal operators.

Proposition 4.94. *Let F be an endomorphism on V with adjoint $F^\dagger$. Then F is normal iff*

$$\langle Fv, Fw \rangle = \left\langle F^\dagger v, F^\dagger w \right\rangle$$

for all $v, w \in V$.

Proof. By definition of adjoint operator, we have

$$\langle Fv, Fw \rangle = \left\langle F^\dagger F v, w \right\rangle \qquad \text{and} \qquad \left\langle F^\dagger v, F^\dagger w \right\rangle = \left\langle F F^\dagger v, w \right\rangle.$$

Therefore, if F is normal, the stated equality follows immediately. If, on the other hand, the stated equality holds, then we get

$$\left\langle F^\dagger F v - F F^\dagger v, w \right\rangle = 0$$

for all v, w. By the nondegeneracy of $\langle\ ,\ \rangle$, we then have $F^\dagger F v - F F^\dagger v = 0$ for all v, i.e., $F^\dagger F = F F^\dagger$. □

PROPOSITION 4.95. *Let F be normal. Then the following hold:*

(1) v is an eigenvector of F with eigenvalue λ iff it is an eigenvector of $F^\dagger$ with eigenvalue $\overline{\lambda}$:

$$Fv = \lambda v \iff F^\dagger v = \overline{\lambda} v.$$

(2) $\operatorname{Eig}(F, \lambda) = \operatorname{Eig}(F^\dagger, \overline{\lambda})$ for every eigenvalue λ of F.

PROOF. By definition, $v \neq 0$ is an eigenvector of F with eigenvalue λ iff $Fv = \lambda v$. This occurs iff $\|(F - \lambda \operatorname{Id})v\| = 0$. However,

$$\begin{aligned}\|(F - \lambda \operatorname{Id})v\|^2 &= \langle (F - \lambda \operatorname{Id})v, (F - \lambda \operatorname{Id})v\rangle \\ &= \left\langle (F^\dagger - \overline{\lambda} \operatorname{Id})v, (F^\dagger - \overline{\lambda} \operatorname{Id})v\right\rangle = \|(F^\dagger - \overline{\lambda} \operatorname{Id})v\|^2,\end{aligned}$$

where we have used that F (and hence $F - \lambda \operatorname{Id}$) is normal. Therefore, the previous condition holds iff $\|(F^\dagger - \overline{\lambda} \operatorname{Id})v\| = 0$, i.e., iff $F^\dagger v = \overline{\lambda} v$, viz., iff $v \neq 0$ is an eigenvector of $F^\dagger$ with eigenvalue $\overline{\lambda}$. □

PROPOSITION 4.96. *Let v be an eigenvector of a normal operator F. Then $v^\perp$ is an F-invariant and $F^\dagger$-invariant subspace. Moreover, the restriction of F to $v^\perp$ is normal.*

PROOF. We have $Fv = \lambda v$. Let $w \in v^\perp$. Then, using Proposition 4.95,

$$\langle v, Fw\rangle = \left\langle F^\dagger v, w\right\rangle = \left\langle \overline{\lambda} v, w\right\rangle = \lambda \langle v, w\rangle = 0,$$

so $Fw \in v^\perp$. Similarly one proves that $v^\perp$ is also $F^\dagger$-invariant:

$$\left\langle v, F^\dagger w\right\rangle = \langle Fv, w\rangle = \langle \lambda v, w\rangle = \overline{\lambda} \langle v, w\rangle = 0,$$

Finally, if $w, w' \in v^\perp$, the identity $\langle Fw, Fw'\rangle = \left\langle F^\dagger w, F^\dagger w'\right\rangle$ shows that $F_{|_{v^\perp}}$ is normal (by Proposition 4.94). □

The above results lead to the following

THEOREM 4.97. *An operator F on a finite-dimensional hermitian product space V is normal iff there is an orthonormal basis of eigenvectors (in particular, F is diagonalizable by Theorem 2.41).*

PROOF. Suppose first that we have an orthonormal basis $(v_1, \ldots, v_n)$ of eigenvectors of F with corresponding eigenvalues $(\lambda_1, \ldots, \lambda_n)$. We then have

$$\left\langle F^\dagger v_i, v_j\right\rangle = \langle v_i, F v_j\rangle = \lambda_j \delta_{ij}.$$

Expanding $F^\dagger v_i = \sum_k \alpha_{ik} v_k$ yields $\overline{\alpha}_{ij} = \lambda_j \delta_{ij}$. Therefore, $F^\dagger v_i = \sum_k \overline{\lambda}_k \delta_{ik} v_k = \overline{\lambda}_i v_i$. As a consequence,

$$F^\dagger F v_i = F^\dagger(\lambda_i v_i) = \lambda_i F^\dagger v_i = |\lambda_i|^2 v_i.$$

On the other hand,

$$F F^\dagger v_i = F(\overline{\lambda}_i v_i) = \overline{\lambda}_i F v_i = |\lambda_i|^2 v_i,$$

so

$$F^\dagger F v_i = F F^\dagger v_i.$$

Since this holds for every basis element v_i, we conclude that $F^\dagger F = FF^\dagger$, so F is normal.[13]

Now suppose instead that F is normal. We prove that it admits an orthonormal basis of eigenvectors by induction on the dimension n of V. If $n = 1$, there is nothing to prove.

Assume we have proved the statement for spaces of dimension n, and let $\dim V = n + 1$. Let λ be an eigenvalue of F (which exists because, by the fundamental theorem of algebra, the characteristic polynomial has roots). Let v be an eigevector for λ. We can assume that $\|v\| = 1$ (otherwise we just rescale v by its norm). By Proposition 4.96, $v^\perp$ is F-invariant and F restricted to it is normal. Moreover, by Theorem 4.49, $\dim v^\perp = n$. Therefore, by the induction hypothesis, $v^\perp$ has an orthonormal basis of eigenvectors. This basis together with v is then an orthonormal basis of V. □

Remark 4.98. Note that the proof of the theorem also gives a recursive procedure to obtain an orthonormal basis of eigenvectors of a normal operator.

Remark 4.99. Note that Theorem 4.97 also implies that the eigenspaces corresponding to different eigenvalues of a normal operator F are orthogonal to each other:

$$\operatorname{Eig}(F,\lambda) \perp \operatorname{Eig}(F,\mu) \quad \text{if } \lambda \neq \mu.$$

In particular, this implies that any two eigenvectors corresponding to different eigenvalues are orthogonal to each other:

$$Fv = \lambda v,\ Fw = \mu w,\ \lambda \neq \mu \implies v \perp w. \tag{4.19}$$

---

[13] An equivalent proof is based on the observation that, in an orthonormal basis of eigenvectors, F is represented by a diagonal matrix $\boldsymbol{D} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ and $F^\dagger$ by $\boldsymbol{D}^\dagger = \begin{pmatrix} \overline{\lambda}_1 & & \\ & \ddots & \\ & & \overline{\lambda}_n \end{pmatrix}$. Therefore, $\boldsymbol{D}^\dagger \boldsymbol{D} = \boldsymbol{D}\boldsymbol{D}^\dagger$.

A second consequence is that the spectral decomposition of Section 2.3.1 is now orthogonal

$$V = \operatorname{Eig}(F, \lambda_1) \bigoplus^{\perp} \cdots \bigoplus^{\perp} \operatorname{Eig}(F, \lambda_k),$$

where $\lambda_1, \ldots, \lambda_k$ are the pairwise distinct eigenvalues of the normal operator F. Therefore, the projection operator P_i corresponding to the eigenspace $\operatorname{Eig}(F, \lambda_i)$ is self adjoint for every i (see Remark 4.53). In summary,

$$P_i^\dagger = P_i \ \forall i, \quad P_i^2 = P_i \ \forall i, \quad P_i P_j = P_j P_i = 0 \ \forall i \neq j, \quad \sum_{i=1}^{k} P_i = \mathrm{Id},$$

and

$$F = \sum_{i=1}^{k} \lambda_i P_i.$$

The theorem also has the following fundamental

Corollary 4.100. *An $n \times n$ complex matrix $\boldsymbol{A}$ is normal iff there is a unitary matrix $\boldsymbol{S}$ such that*

$$\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{D} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

where $(\lambda_1, \ldots, \lambda_n)$ are the eigenvalues of $\boldsymbol{A}$.

Proof. View $\boldsymbol{A}$ as the endomorphism of $\mathbb{C}^n$ defined by $\boldsymbol{v} \mapsto \boldsymbol{A}\boldsymbol{v}$.

If there is a unitary matrix $\boldsymbol{S}$ as in the statement, then its columns $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ are an orthonormal basis of $\mathbb{C}^n$. Moreover, the columns of $\boldsymbol{A}\boldsymbol{S}$ are $(\boldsymbol{A}\boldsymbol{v}_1, \ldots, \boldsymbol{A}\boldsymbol{v}_n)$, whereas the columns of $\boldsymbol{S}\boldsymbol{D}$, which by the statement is the same as $\boldsymbol{A}\boldsymbol{S}$, are $(\lambda_1\boldsymbol{v}_1, \ldots, \lambda_n\boldsymbol{v}_n)$. Therefore, the $\boldsymbol{v}_i$s are eigenvectors. Since we have an orthonormal basis of eigenvectors, then $\boldsymbol{A}$ is normal by Theorem 4.97.

On the other hand, if $\boldsymbol{A}$ is normal, then by Theorem 4.97 we have an orthonormal basis $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ of eigenvectors. If we let $\boldsymbol{S}$ be the matrix with columns the $\boldsymbol{v}_i$s, then $\boldsymbol{S}$ is unitary and satisfies the stated identity. □

In applications, especially to quantum mechanics, it is often important to diagonalize two different normal endomorphisms at the same time: we say that two <u>normal</u> endomorphisms F and G on a hermitian product space V are simultaneously diagonalizable if they possess a common <u>orthonormal</u> basis of eigenvectors. We have the following generalization of Proposition 2.47.

PROPOSITION 4.101 (Simultaneous diagonalization). *Two normal endomorphisms F and G on a hermitian product space V are simultaneously diagonalizable iff they commute, i.e., $FG = GF$.*

PROOF. See Exercise 4.7. □

In the case of matrices, the above proposition reads more explictly as follows.

COROLLARY 4.102. *Two normal matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ commute (i.e., $\boldsymbol{AB} = \boldsymbol{BA}$) iff there is a unitary matrix $\boldsymbol{S}$ such that $\boldsymbol{S}^{-1}\boldsymbol{AS} = \boldsymbol{D}$ and $\boldsymbol{S}^{-1}\boldsymbol{BS} = \boldsymbol{D}'$ where $\boldsymbol{D}$ and $\boldsymbol{D}'$ are diagonal matrices.*

4.6.2. Diagonalization of unitary matrices. Let $\boldsymbol{U}$ be a unitary matrix, hence normal. We can then apply Corollary 4.100. We first however make the following observation.

PROPOSITION 4.103. *The eigenvalues of a unitary matrix have absolute value* 1. *Therefore, they are of the form* $\mathrm{e}^{\mathrm{i}\theta}$ *for some* $\theta \in \mathbb{R}$.

PROOF. Let $\boldsymbol{v}$ be an eigenvector to the eigenvalue λ of the unitary matrix $\boldsymbol{U}$: $\boldsymbol{Uv} = \lambda\boldsymbol{v}$. Then

$$\|\boldsymbol{v}\|^2 = \langle \boldsymbol{v}, \boldsymbol{v}\rangle = \langle \boldsymbol{Uv}, \boldsymbol{Uv}\rangle = \langle \lambda\boldsymbol{v}, \lambda\boldsymbol{v}\rangle = |\lambda|^2\|\boldsymbol{v}\|^2,$$

so $|\lambda|^2 = 1$, since $\boldsymbol{v} \neq \boldsymbol{0}$. □

REMARK 4.104. By Remark 4.99 we have that any two eigenvectors of a unitary matrix corresponding to different eigenvalues are orthogonal to each other, see (4.19). This can also be checked directly. Namely, assume $\boldsymbol{Uv} = \lambda\boldsymbol{v}$ and $\boldsymbol{Uw} = \mu\boldsymbol{w}$. Then

$$\langle \boldsymbol{v}, \boldsymbol{w}\rangle = \langle \boldsymbol{Uv}, \boldsymbol{Uw}\rangle = \langle \lambda\boldsymbol{v}, \mu\boldsymbol{w}\rangle = \overline{\lambda}\mu \langle \boldsymbol{v}, \boldsymbol{w}\rangle = \frac{\mu}{\lambda}\langle \boldsymbol{v}, \boldsymbol{w}\rangle.$$

If $\lambda \neq \mu$, then this implies $\langle \boldsymbol{v}, \boldsymbol{w}\rangle = \boldsymbol{0}$.

As an immediate consequence of Corollary 4.100, we then get the following

THEOREM 4.105. *An $n \times n$ complex matrix $\boldsymbol{U}$ is unitary iff there is a unitary matrix $\boldsymbol{S}$ such that*

$$\boldsymbol{S}^{-1}\boldsymbol{US} = \begin{pmatrix} \mathrm{e}^{\mathrm{i}\theta_1} & & \\ & \ddots & \\ & & \mathrm{e}^{\mathrm{i}\theta_n} \end{pmatrix},$$

where $\theta_1, \dots, \theta_n$ are real numbers.

PROOF. The only-if part follows directly from Corollary 4.100 and Proposition 4.103. The if part is just a computation: observe that the diagonal matrix $\boldsymbol{D}$ on the right hand side is unitary and that therefore $\boldsymbol{S}\boldsymbol{D}\boldsymbol{S}^{-1}$ is also unitary, since $\boldsymbol{S}$ is so. □

REMARK 4.106. Note that $\boldsymbol{U}$ in the theorem is special unitary iff $\theta_1 + \cdots + \theta_n = 2\pi k$ with $k \in \mathbb{Z}$. We can actually assume without loss of generality (just by changing θ_n to $\theta_n - 2\pi k$) that

$$\theta_1 + \cdots + \theta_n = 0.$$

We conclude with the following

COROLLARY 4.107. *The exponential maps*

$$\begin{array}{rccc} \exp\colon & \mathfrak{u}(n) & \to & \mathrm{U}(n) \\ & \boldsymbol{A} & \mapsto & \mathrm{e}^{\boldsymbol{A}} \end{array}$$

$$\begin{array}{rccc} \exp\colon & \mathfrak{su}(n) & \to & \mathrm{SU}(n) \\ & \boldsymbol{A} & \mapsto & \mathrm{e}^{\boldsymbol{A}} \end{array}$$

are surjective.

PROOF. Let $\boldsymbol{U} \in \mathrm{U}(n)$. By Theorem 4.105, there is a unitary matrix $\boldsymbol{S}$ such that $\boldsymbol{U} = \boldsymbol{S}\boldsymbol{D}\boldsymbol{S}^{-1}$ where

$$\boldsymbol{D} = \begin{pmatrix} \mathrm{e}^{\mathrm{i}\theta_1} & & \\ & \ddots & \\ & & \mathrm{e}^{\mathrm{i}\theta_n} \end{pmatrix} = \mathrm{e}^{\boldsymbol{B}} \qquad \text{with} \qquad \boldsymbol{B} = \begin{pmatrix} \mathrm{i}\theta_1 & & \\ & \ddots & \\ & & \mathrm{i}\theta_n \end{pmatrix}$$

and $\theta_1, \dots, \theta_n$ real. By (2.16) we then have $\boldsymbol{U} = \mathrm{e}^{\boldsymbol{A}}$ with $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{B}\boldsymbol{S}^{-1} = \boldsymbol{S}\boldsymbol{B}\boldsymbol{S}^{\dagger}$. Since $\boldsymbol{B}$ is clearly anti-self-adjoint, then so is $\boldsymbol{A}$. Therefore, $\boldsymbol{A} \in \mathfrak{u}(n)$.

If $\boldsymbol{U} \in \mathrm{SU}(n)$, we may assume, by Remark 4.106, that $\theta_1 + \cdots + \theta_n = 0$. Therefore, $\operatorname{tr} \boldsymbol{B} = 0$ and hence $\operatorname{tr} \boldsymbol{A} = 0$. Therefore, $\boldsymbol{A} \in \mathfrak{su}(n)$. □

4.6.3. Diagonalization of self-adjoint matrices. Let $\boldsymbol{A}$ be self adjoint, hence normal. We can then apply Corollary 4.100. We first however make the following observation.

PROPOSITION 4.108. *The eigenvalues of a self-adjoint matrix are real.*

PROOF. Let $\boldsymbol{v}$ be an eigenvector to the eigenvalue λ of the self-adjoint matrix $\boldsymbol{A}$: $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$. Then, on the one hand,

$$\langle \boldsymbol{v}, \boldsymbol{A}\boldsymbol{v} \rangle = \langle \boldsymbol{v}, \lambda\boldsymbol{v} \rangle = \lambda \langle \boldsymbol{v}, \boldsymbol{v} \rangle = \lambda \|\boldsymbol{v}\|^2,$$

and, on the other hand,

$$\langle \boldsymbol{A}\boldsymbol{v}, \boldsymbol{v} \rangle = \langle \lambda\boldsymbol{v}, \boldsymbol{v} \rangle = \overline{\lambda} \langle \boldsymbol{v}, \boldsymbol{v} \rangle = \overline{\lambda} \|\boldsymbol{v}\|^2.$$

Since $\langle \boldsymbol{v}, \boldsymbol{A}\boldsymbol{v}\rangle = \langle \boldsymbol{A}\boldsymbol{v}, \boldsymbol{v}\rangle$ for $\boldsymbol{A}$ self adjoint and $\boldsymbol{v} \neq \boldsymbol{0}$, we get $\lambda = \overline{\lambda}$. □

REMARK 4.109. The fact that the eigenvalues of a self-adjoint matrix are real is of fundamental importance for applications in quantum mechanics. We will see that it is also important in view of the diagonalization of real symmetric matrices.

REMARK 4.110. By Remark 4.99 we have that any two eigenvectors of a self-adjoint matrix corresponding to different eigenvalues are orthogonal to each other, see (4.19). This can also be checked directly. Namely, assume $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$ and $\boldsymbol{A}\boldsymbol{w} = \mu\boldsymbol{v}$. Then

$$\langle \boldsymbol{v}, \boldsymbol{A}\boldsymbol{w}\rangle = \mu \langle \boldsymbol{v}, \boldsymbol{w}\rangle \qquad \text{and} \qquad \langle \boldsymbol{A}\boldsymbol{v}, \boldsymbol{w}\rangle = \overline{\lambda} \langle \boldsymbol{v}, \boldsymbol{w}\rangle = \lambda \langle \boldsymbol{v}, \boldsymbol{w}\rangle ,$$

as λ must be real. Since $\langle \boldsymbol{v}, \boldsymbol{A}\boldsymbol{v}\rangle = \langle \boldsymbol{A}\boldsymbol{v}, \boldsymbol{v}\rangle$, for $\lambda \neq \mu$ this implies $\langle \boldsymbol{v}, \boldsymbol{w}\rangle = \boldsymbol{0}$.

As an immediate consequence of Corollary 4.100, we then get the following

THEOREM 4.111. *An $n \times n$ complex matrix $\boldsymbol{A}$ is self adjoint iff there is a unitary matrix $\boldsymbol{S}$ such*

$$\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

where $\lambda_1, \dots, \lambda_n$ are real numbers.

PROOF. The only-if part follows directly from Corollary 4.100 and Proposition 4.108. The if part is just a computation: observe that the diagonal matrix $\boldsymbol{D}$ on the right hand side is self adjoint and that therefore $\boldsymbol{S}\boldsymbol{D}\boldsymbol{S}^{-1}$ is also self adjoint, since $\boldsymbol{S}$ is unitary. □

4.6.4. Normal form of orthogonal matrices. An orthogonal matrix $\boldsymbol{O}$ when viewed as a complex matrix is unitary. Therefore, by Theorem 4.105, it is diagonalizable, as a complex matrix, with eigenvalues of the form $\mathrm{e}^{\mathrm{i}\theta}$. As some eigenvalues may not be real, in general an orthogonal matrix is not diagonalizable (over the reals). However, one can arrange the orthonormal basis of eigenvectors so as to prove the following normal form theorem.

THEOREM 4.112. *Let $\boldsymbol{O}$ be an $n \times n$ orthogonal matrix. Then there is an orthogonal matrix $\boldsymbol{S}$ such that $\boldsymbol{S}^{-1}\boldsymbol{O}\boldsymbol{S} = \boldsymbol{R}$ where $\boldsymbol{R}$ has one of the following block diagonal forms (we have to distinguish four cases).*

	$n=2r$	$n=2r+1$
$\det \boldsymbol{O}=1$	$\begin{pmatrix} \boldsymbol{R}(\theta_1) & & \\ & \ddots & \\ & & \boldsymbol{R}(\theta_r) \end{pmatrix}$	$\begin{pmatrix} 1 & & & \\ & \boldsymbol{R}(\theta_1) & & \\ & & \ddots & \\ & & & \boldsymbol{R}(\theta_r) \end{pmatrix}$
$\det \boldsymbol{O}=-1$	$\begin{pmatrix} 1 & & & & \\ & -1 & & & \\ & & \boldsymbol{R}(\theta_1) & & \\ & & & \ddots & \\ & & & & \boldsymbol{R}(\theta_{r-1}) \end{pmatrix}$	$\begin{pmatrix} -1 & & & \\ & \boldsymbol{R}(\theta_1) & & \\ & & \ddots & \\ & & & \boldsymbol{R}(\theta_r) \end{pmatrix}$

with

$$\boldsymbol{R}(\theta) := \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

For the proof we will need a generalization of Proposition 4.96.

LEMMA 4.113. *If W is an $\boldsymbol{O}$-invariant subspace, for an orthogonal matrix $\boldsymbol{O}$, then so is $W^\perp$.*

PROOF. First observe that $\boldsymbol{O}$ restricted to W is still injective and therefore, as a map $W \to W$, bijective. Therefore, W is also $\boldsymbol{O}^{-1}$-invariant. Then, for every $\boldsymbol{v} \in W^\perp$ and $\boldsymbol{w} \in W$, we have

$$\boldsymbol{w} \cdot (\boldsymbol{O}\boldsymbol{v}) = (\boldsymbol{O}^{-1}\boldsymbol{w}) \cdot \boldsymbol{v} = 0,$$

where in the first equality we have used the fact that also $\boldsymbol{O}^{-1}$ is orthogonal and in the second that $\boldsymbol{O}^{-1}\boldsymbol{w}$ is in W. This shows that $\boldsymbol{O}\boldsymbol{v} \in W^\perp$. □

PROOF OF THEOREM 4.112. If an eigenvalue is real, hence necessarily equal to ± 1, then we can choose the corresponding eigenvector to be real. By Lemma 4.113, its orthogonal space is $\boldsymbol{O}$-invariant. We can then proceed by induction until no real eigenvalues are left.

If an eigenvalue is not real, then its complex conjugate is also an eigenvalue, since the characteristic polynomial of $\boldsymbol{O}$ is real and therefore each nonreal root comes with its complex conjugate. If $\mathrm{e}^{\mathrm{i}\theta}$ is an eigenvalue, with θ different from 0 and π modulo 2π, then $\mathrm{e}^{-\mathrm{i}\theta}$ is a distinct eigenvalue. If $\boldsymbol{v}$ is an eigenvector for $\mathrm{e}^{\mathrm{i}\theta}$, taking the complex conjugation of $\boldsymbol{O}\boldsymbol{v} = \mathrm{e}^{\mathrm{i}\theta}\boldsymbol{v}$ yields $\boldsymbol{O}\overline{\boldsymbol{v}} = \mathrm{e}^{-\mathrm{i}\theta}\overline{\boldsymbol{v}}$, so $\overline{\boldsymbol{v}}$ is an eigenvector for $\mathrm{e}^{-\mathrm{i}\theta}$. By Remark 4.104, $\boldsymbol{v} \perp \overline{\boldsymbol{v}}$. Assuming $\|\boldsymbol{v}\| = 1$ (otherwise just divide $\boldsymbol{v}$ by its norm), we also have $\|\overline{\boldsymbol{v}}\| = 1$. It follows that the real vectors

$$\boldsymbol{a} := \frac{\boldsymbol{v} + \overline{\boldsymbol{v}}}{\sqrt{2}} \qquad \text{and} \qquad \boldsymbol{b} := \frac{\boldsymbol{v} - \overline{\boldsymbol{v}}}{\mathrm{i}\sqrt{2}}$$

are an orthonormal system[14] and therefore linearly independent. Moreover,

$$\boldsymbol{O}\boldsymbol{a} = \frac{\mathrm{e}^{\mathrm{i}\theta}\boldsymbol{v} + \mathrm{e}^{-\mathrm{i}\theta}\overline{\boldsymbol{v}}}{\sqrt{2}} = \cos\theta\,\frac{\boldsymbol{v}+\overline{\boldsymbol{v}}}{\sqrt{2}} + \mathrm{i}\sin\theta\,\frac{\boldsymbol{v}-\overline{\boldsymbol{v}}}{\sqrt{2}} = \cos\theta\,\boldsymbol{a} - \sin\theta\,\boldsymbol{b},$$

$$\boldsymbol{O}\boldsymbol{b} = \frac{\mathrm{e}^{\mathrm{i}\theta}\boldsymbol{v} - \mathrm{e}^{-\mathrm{i}\theta}\overline{\boldsymbol{v}}}{\mathrm{i}\sqrt{2}} = \cos\theta\,\frac{\boldsymbol{v}-\overline{\boldsymbol{v}}}{\mathrm{i}\sqrt{2}} + \mathrm{i}\sin\theta\,\frac{\boldsymbol{v}+\overline{\boldsymbol{v}}}{\mathrm{i}\sqrt{2}} = \cos\theta\,\boldsymbol{b} + \sin\theta\,\boldsymbol{a}.$$

Therefore, $\boldsymbol{O}$ restricted to $\mathrm{Span}\{\boldsymbol{a},\boldsymbol{b}\}$ is represented, in the basis $(\boldsymbol{b},\boldsymbol{a})$, by the matrix $\boldsymbol{R}(\theta)$.

We can keep grouping pairs of eigenvectors with conjugate eigenvalues, apply the above construction, restrict $\boldsymbol{O}$ to the orthogonal complement of their real span (which is $\boldsymbol{O}$-invariant by Lemma 4.113) and continue by induction. This way we get an orthonormal basis of $\mathbb{R}^n$ in which $\boldsymbol{O}$ is represented by a block diagonal matrix whose blocks are either 1×1 with entry ± 1 or 2×2 of the form $\boldsymbol{R}(\theta)$.

We can also group a pair of eigenvectors with eigenvalue $+1$. In this case, $\boldsymbol{O}$ restricted to their span is the 2×2 identity matrix, i.e., $\boldsymbol{R}(0)$. If instead we group a pair of eigenvectors with eigenvalue -1, then $\boldsymbol{O}$ restricted to their span is minus the 2×2 identity matrix, i.e., $\boldsymbol{R}(\pi)$. Therefore, we can rearrange the orthonormal basis (and get the orthogonal matrix $\boldsymbol{S}$ whose columns are the elements of this basis) so that $\boldsymbol{O}$ is represented as in the table. The four cases just correspond to the fact that the number of ± 1-eigenvectors can be even/odd. $\square$

We conclude with the following

COROLLARY 4.114. *The exponential map*

$$\begin{array}{rccc} \exp\colon & \mathfrak{so}(n) & \to & \mathrm{SO}(n) \\ & \boldsymbol{A} & \mapsto & \mathrm{e}^{\boldsymbol{A}} \end{array}$$

is surjective.

PROOF. Let $\boldsymbol{O} \in \mathrm{SO}(n)$. By Theorem 4.112, there is an orthogonal matrix $\boldsymbol{S}$ such that $\boldsymbol{O} = \boldsymbol{S}\boldsymbol{R}\boldsymbol{S}^{-1}$ with $\boldsymbol{R}$ as in the first row of the table. Note that $\boldsymbol{R} = \mathrm{e}^{\boldsymbol{\rho}}$ with $\boldsymbol{\rho}$ of the form

$n = 2r$	$n = 2r+1$
$\begin{pmatrix} \boldsymbol{\rho}(\theta_1) & & \\ & \ddots & \\ & & \boldsymbol{\rho}(\theta_r) \end{pmatrix}$	$\begin{pmatrix} 0 & & & \\ & \boldsymbol{\rho}(\theta_1) & & \\ & & \ddots & \\ & & & \boldsymbol{\rho}(\theta_r) \end{pmatrix}$

[14]Namely, $\langle \boldsymbol{a},\boldsymbol{a}\rangle = \langle \boldsymbol{b},\boldsymbol{b}\rangle = 1$ and $\langle \boldsymbol{a},\boldsymbol{b}\rangle = 0$. Since $\boldsymbol{a}$ and $\boldsymbol{b}$ are real, this is the same as $\boldsymbol{a}\cdot\boldsymbol{a} = \boldsymbol{b}\cdot\boldsymbol{b} = 1$ and $\boldsymbol{a}\cdot\boldsymbol{b} = 0$.

with

$$\boldsymbol{\rho}(\theta) := \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix}.$$

Therefore, $\boldsymbol{O} = \mathrm{e}^{\boldsymbol{A}}$ with $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{\rho}\boldsymbol{S}^{-1} = \boldsymbol{S}\boldsymbol{\rho}\boldsymbol{S}^{\mathsf{T}}$, which is skew-symmetric, since $\boldsymbol{\rho}$ is so. □

4.6.5. Diagonalization of real symmetric matrices and bilinear forms. A real symmetric matrix $\boldsymbol{A}$ when viewed as a complex matrix is self adjoint. Therefore, by Theorem 4.105, it is diagonalizable, as a complex matrix. By Proposition 4.108 its eigenvalues are however real. As a consequence, $\boldsymbol{A}$ is diagonalizable also as a real matrix:

THEOREM 4.115. *Let $\boldsymbol{A}$ be an $n \times n$ real symmetric matrix. Then there is an orthogonal matrix $\boldsymbol{S}$ such that*

$$\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

PROOF. The proof proceeds by induction as in the case of Theorem 4.97 by the following two remarks.

First, a real symmetric matrix has real eigenvalues, so it has at least an eigenvector. Second, if $\boldsymbol{v}$ is an eigenvector for some eigenvalue λ, $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$, then for every $\boldsymbol{w}$

$$\boldsymbol{v} \cdot (\boldsymbol{A}\boldsymbol{w}) = (\boldsymbol{A}^{\mathsf{T}}\boldsymbol{v}) \cdot \boldsymbol{w} = (\boldsymbol{A}\boldsymbol{v}) \cdot \boldsymbol{w} = \lambda\, \boldsymbol{v} \cdot \boldsymbol{w}.$$

Therefore, $\boldsymbol{w} \in \boldsymbol{v}^{\perp}$ implies $\boldsymbol{A}\boldsymbol{w} \in \boldsymbol{v}^{\perp}$. We can then proceed by induction to show that there is an orthonormal basis of eigenvectors.

The matrix $\boldsymbol{S}$ is obtained as the matrix whose columns are the elements of this basis. □

A symmetric matrix $\boldsymbol{A}$ is often used to define a symmetric bilinear form by

$$(\boldsymbol{v}, \boldsymbol{w}) := \boldsymbol{v}^{\mathsf{T}}\boldsymbol{A}\boldsymbol{w}.$$

From this point of view, one has to consider symmetric matrices up to congruency as in Remark 1.71. Namely, recall that $\boldsymbol{A}$ and $\boldsymbol{B}$ are congruent if there is an invertible matrix $\boldsymbol{T}$ such that $\boldsymbol{B} = \boldsymbol{T}^{\mathsf{T}}\boldsymbol{A}\boldsymbol{T}$. Since $\boldsymbol{S}$ in Theorem 4.115 is orthogonal, we have that a symmetric matrix $\boldsymbol{A}$ is congruent to the diagonal matrix with the eigenvalues on its diagonal. We can actually get an even more standard form.

THEOREM 4.116. *Let $\boldsymbol{A}$ be a real symmetric matrix. Then there is an invertible matrix $\boldsymbol{T}$ such that $\boldsymbol{T}^{\mathsf{T}}\boldsymbol{A}\boldsymbol{T}$ is a diagonal matrix whose diagonal entries are only from the set $\{0, 1, -1\}$. This diagonal matrix is called a normal form for $\boldsymbol{A}$.*

Proof. Let $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ be an orthonormal basis of eigenvectors of $\boldsymbol{A}$ (e.g., the columns of the matrix $\boldsymbol{S}$ of Theorem 4.115). We define

$$\widetilde{\boldsymbol{v}}_i := \begin{cases} \boldsymbol{v}_i & \text{if } \lambda_i = 0, \\ \frac{\boldsymbol{v}_i}{\sqrt{\lambda_i}} & \text{if } \lambda_i > 0, \\ \frac{\boldsymbol{v}_i}{\sqrt{-\lambda_i}} & \text{if } \lambda_i < 0. \end{cases}$$

Then $(\widetilde{\boldsymbol{v}}_1, \dots, \widetilde{\boldsymbol{v}}_n)$ is an orthogonal basis of eigenvectors. Moreover, $\widetilde{\boldsymbol{v}}_i^{\mathsf{T}} \widetilde{\boldsymbol{v}}_i$ is equal to 1 in the first case and to $\frac{1}{\lambda_i}$ in the second and third case. The matrix $\boldsymbol{T}$ is finally obtained as the matrix whose columns are the elements of this basis. □

We can rearrange the basis vectors (i.e., permute the columns of $\boldsymbol{T}$) in such a way that the first diagonal entries of the diagonal matrix are equal to 0, the second to -1 and last to $+1$. The number of the entries of each type is an invariant under congruency. To prove this, we first consider the following

Lemma 4.117. *Let $(\ ,\ )$ be a symmetric bilinear form on a real finite-dimensional vector space V. Set*

$$N := \{v \in V \mid (v, w) = 0 \ \forall w \in V\},$$

which is clearly a subspace (called the null subspace of the bilinear form). Let V_-, V_+ be subspaces of V such that following hold:

(1) $V = N \oplus V_- \oplus V_+$.
(2) For every v and w in different summands $(v, w) = 0$.
(3) $\forall v \in V_- \setminus \{0\}$, $(v, v) < 0$.
(4) $\forall v \in V_+ \setminus \{0\}$, $(v, v) > 0$.

Then, for every other decomposition N, W_-, W_+ of V with these properties, we have $\dim W_- = \dim V_-$ and $\dim W_+ = \dim V_+$.

In particular, the numbers $n_0 = \dim N$, $n_- = \dim V_-$, and $n_+ = \dim V_+$ only depend on the bilinear form and not on the decomposition. The triple (n_0, n_-, n_+) is called the signature of the symmetric bilinear form.

Example 4.118. The signature of an inner product on V, $\dim V = n$, is $(0, 0, n)$. The signature of the Minkowski product on $\mathbb{R}^{n+1}$ is $(0, 1, n)$.

Proof of the lemma. Let $\pi_\pm \colon V \to V_\pm$ be the projections to the correspnding summand; i.e., if v decomposes (uniquely) as $v = v_0 + v_- + v_+$, then $\pi_\pm v = v_\pm$.

Now consider subspaces W_-, W_+ also satisfying (1), (2), (3), and (4). We claim that ${\pi_-}_{|W_-}$ and ${\pi_+}_{|W_+}$ are injective. In fact, let $w \in W_-$.

If $\pi_- w = 0$, then $w = w_0 + w_+$ with $w_0 \in N$ and $w_+ \in V_+$. Therefore, $(w, w) = (w_+, w_+) \geq 0$. Since $(w, w) \leq 0$, we conclude $(w, w) = 0$, so $w = 0$. Similarly, for $w \in W_+$.

As a consequence, $\dim W_- \leq \dim V_-$ and $\dim W_+ \leq \dim V_+$. Since, by (1), $\dim W_- + \dim W_+ = \dim V_- + \dim V_+$, we conclude that both inequalities are saturated. □

The spans of the basis elements given by the columns of any $\boldsymbol{T}$ as in Theorem 4.116 corresponding to the diagonal entries 0, −1 and 1, respectively, yield a choice of decomposition $N \oplus V_+ \oplus V_-$. This first of all shows that such a decomposition exists. It also shows that the numbers of 0s, −1s, and 1s on the diagonal of the normal form do not depend on the choice of basis:

PROPOSITION 4.119 (Sylvester's law of inertia). *Let $\boldsymbol{A}$ be an $n \times n$ real symmetric matrix. Let $\boldsymbol{D}$ be a diagonal matrix congruent to $\boldsymbol{A}$ whose diagonal entries are in the set $\{0, 1, -1\}$. Let (d_0, d_-, d_+) be the number of diagonal entries of $\boldsymbol{D}$ equal to 0, −1 and 1, respectively. Then (d_0, d_-, d_+) is equal to the signature (n_0, n_-, n_+) of the symmetric bilinear form associated to $\boldsymbol{A}$.*

In particular, we have that for every $n \times n$ real symmetric matrix $\boldsymbol{A}$ there is an invertible matrix $\boldsymbol{T}$ such that

$$\boldsymbol{T}^\mathsf{T} \boldsymbol{A} \boldsymbol{T} = \begin{pmatrix} \mathbf{0}_{n_0} & & \\ & -\mathbf{1}_{n_-} & \\ & & \mathbf{1}_{n_+} \end{pmatrix}.$$

4.6.6. Digression: Normal form of real skew-symmetric bilinear forms. If $\boldsymbol{B}$ is an $n \times n$ real skew-symmetric matrix, then $\boldsymbol{A} := \mathrm{i}\boldsymbol{B}$ is self adjoint. Therefore, by Theorem 4.111, $\boldsymbol{A}$ is diagonalizable with an orthonormal basis of eigenvectors; moreover, by Proposition 4.108, its eigenvalues are real. As a consequence the eigenvalues of $\boldsymbol{B}$ are purely imaginary or zero.

To proceed, we will need a generalization of Proposition 4.96, similar to Lemma 4.113.

LEMMA 4.120. *If W is an $\boldsymbol{B}$-invariant subspace, for a real skew-symmetric matrix $\boldsymbol{B}$, then so is $W^\perp$.*

PROOF. For every $\boldsymbol{v} \in W^\perp$ and $\boldsymbol{w} \in W$, we have

$$\boldsymbol{w} \cdot (\boldsymbol{B}\boldsymbol{v}) = (\boldsymbol{B}^\mathsf{T}\boldsymbol{w}) \cdot \boldsymbol{v} = -(\boldsymbol{B}\boldsymbol{w}) \cdot \boldsymbol{v} = 0,$$

since $\boldsymbol{B}\boldsymbol{w} \in W$ by assumption. This shows that $\boldsymbol{B}\boldsymbol{v} \in W^\perp$. □

Let us restrict $\boldsymbol{B}$ to the orthogonal complement of its kernel, which is $\boldsymbol{B}$-invariant by Lemma 4.120. Its eigenvalues there will then be purely imaginary and different from zero. We proceed as in the case of orthogonal matrices. Since $\boldsymbol{B}$ is real, every nonreal eigenvalue comes with its complex conjugate. Therefore, if $\boldsymbol{v}$ is an eigenvector for the eigenvalue $\mathrm{i}\lambda$, $\lambda \in \mathbb{R}_{\neq 0}$, then $\overline{\boldsymbol{v}}$ is an eigenvector for the distinct eigenvalue $-\mathrm{i}\lambda$. By Remark 4.110, $\boldsymbol{v} \perp \overline{\boldsymbol{v}}$. Assuming $\|\boldsymbol{v}\| = 1$ (otherwise just divide $\boldsymbol{v}$ by its norm), we also have $\|\overline{\boldsymbol{v}}\| = 1$. It follows that the real vectors

$$\boldsymbol{a} := \frac{\boldsymbol{v} + \overline{\boldsymbol{v}}}{\sqrt{2}} \qquad \text{and} \qquad \boldsymbol{b} := \frac{\boldsymbol{v} - \overline{\boldsymbol{v}}}{\mathrm{i}\sqrt{2}}$$

are an orthonormal system and therefore linearly independent. Moreover,

$$\boldsymbol{B}\boldsymbol{a} = \mathrm{i}\lambda\frac{\boldsymbol{v} - \overline{\boldsymbol{v}}}{\sqrt{2}} = -\lambda\boldsymbol{b}, \qquad \boldsymbol{B}\boldsymbol{b} = \mathrm{i}\lambda\frac{\boldsymbol{v} + \overline{\boldsymbol{v}}}{i\sqrt{2}} = \lambda\boldsymbol{a}.$$

Therefore, $\boldsymbol{B}$ restricted to $\operatorname{Span}\{\boldsymbol{a}, \boldsymbol{b}\}$ is represented, in the basis $(\boldsymbol{a}, \boldsymbol{b})$, by the matrix $\left(\begin{smallmatrix} 0 & \lambda \\ -\lambda & 0 \end{smallmatrix}\right)$.

Going to the orthogonal space and proceeding by induction (thanks to Lemma 4.120), we then get an orthogonal matrix $\boldsymbol{S}$, whose columns are the elements of an orthogonal basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_r, \boldsymbol{a}_1, \boldsymbol{b}_1, \dots, \boldsymbol{a}_k, \boldsymbol{b}_k)$, with $r = \dim\ker \boldsymbol{B}$, such that

$$\boldsymbol{S}^{-1}\boldsymbol{B}\boldsymbol{S} = \begin{pmatrix} \boldsymbol{0}_r & & & & & \\ & 0 & \lambda_1 & & & \\ & -\lambda_1 & 0 & & & \\ & & & \ddots & & \\ & & & & 0 & \lambda_k \\ & & & & -\lambda_k & 0 \end{pmatrix}$$

with $\lambda_i \neq 0$ for $i = 1, \dots, k$. Without loss of generality (swap a_i with b_i otherwise), we can assume $\lambda_i > 0$ for every i.

From the point of view of the associated skew-symmetric bilinear form $(\boldsymbol{v}, \boldsymbol{w}) := \boldsymbol{v}^{\mathsf{T}}\boldsymbol{B}\boldsymbol{w}$, it is natural to consider $\boldsymbol{B}$ up to congruency. We can in this case choose a new basis by setting

$$\widetilde{\boldsymbol{v}}_i = \boldsymbol{v}_i, \qquad \widetilde{\boldsymbol{a}}_i = \frac{\boldsymbol{a}_i}{\sqrt{\lambda_i}}, \qquad \widetilde{\boldsymbol{b}}_i = \frac{\boldsymbol{b}_i}{\sqrt{\lambda_i}},$$

where we assume, as remarked above, $\lambda_i > 0$ for every i. We then set $\boldsymbol{T}$ to be the invertible matrix whose columns are the new basis vectors. We have then proved the following

THEOREM 4.121. *Let $\boldsymbol{B}$ be an $n \times n$ real skew-symmetric matrix. Then $n = r + 2k$, with $r = \dim\ker \boldsymbol{B}$ and some k, and there is an* *invertible matrix $\boldsymbol{T}$ such that*

$$\boldsymbol{T}^\top \boldsymbol{B}\boldsymbol{T} = \begin{pmatrix} \mathbf{0}_r & & & & & \\ & 0 & 1 & & & \\ & -1 & 0 & & & \\ & & & \ddots & & \\ & & & & 0 & 1 \\ & & & & -1 & 0 \end{pmatrix}$$

with k nondegenerate blocks.

In particular, we have the following

Corollary 4.122. *Let $\boldsymbol{B}$ be an $n\times n$ nondegenerate real skew-symmetric matrix. Then n is even.*

This case is particular important for mechanics in Hamilton's formulation, where half of the basis elements, say the $\boldsymbol{a}$s, correspond to the positions of the system and the other half, say the $\boldsymbol{b}$s, correspond to the momenta.

Exercises for Chapter 4

4.1. For which of the following matrices $\boldsymbol{g}$ is the bilinear form $\langle \boldsymbol{v}, \boldsymbol{w}\rangle := \overline{\boldsymbol{v}}^\top \boldsymbol{g}\boldsymbol{w}$ a hermitian product? Motivate your answer.
(a) $\boldsymbol{g} = \left(\begin{smallmatrix} 1 & \mathrm{i} \\ \mathrm{i} & 1 \end{smallmatrix}\right)$.
(b) $\boldsymbol{g} = \left(\begin{smallmatrix} 1 & \mathrm{i} \\ -\mathrm{i} & 1 \end{smallmatrix}\right)$.
(c) $\boldsymbol{g} = \left(\begin{smallmatrix} 1 & \mathrm{i} \\ -\mathrm{i} & 2 \end{smallmatrix}\right)$.

4.2. Let $V = \mathrm{Mat}_{n\times n}(\mathbb{C})$ be the complex vector space of $n \times n$ complex matrices. Show that

$$(\boldsymbol{A}, \boldsymbol{B}) := \mathrm{tr}(\boldsymbol{A}^\dagger \boldsymbol{B})$$

is a hermitian product on V.

4.3. Recall that an operator F on a hermitian product space is called self adjoint if it has an adjoint $F^\dagger$ and $F^\dagger = F$. Consider the linear operator $F := \mathrm{i}\frac{\mathrm{d}}{\mathrm{d}x}$ (i.e., $Ff = \mathrm{i}f'$) on the following complex vector spaces,[15]

$$\begin{aligned} V_1 &:= C^\infty([a,b],\mathbb{C}), \\ V_2 &:= \{f \in V_1 \mid f(a) = f(b)\}, \\ V_3 &:= \{f \in V_2 \mid f(a) = 0\}, \end{aligned}$$

[15] C^∞ denotes the space of infinitely differentiable functions.

with hermitian product

$$\langle f, g\rangle := \int_a^b \bar{f} g \,\mathrm{d}x.$$

(a) Show that F on V_2 and on V_3 is self adjoint.

(b) We now want to show that F on V_1 is not self adjoint because it does not have an adjoint.

(i) Assume by contradiction that $F^\dagger\colon V_1 \to V_1$ exists. Let $G := F^\dagger - F$. Show that $\forall g \in V_3$

$$\langle Gf, g\rangle = 0.$$

Hint: Observe first that $\langle Gf, g\rangle = \langle f, Fg\rangle - \langle Ff, g\rangle$.

(ii) Show that, if $h \in V_1$ satisfies $\langle h, g\rangle = 0$ $\forall g \in V_3$, then $h = 0$.

Hint: You may use the fact that, for every $x_0 \in \mathbb{R}$ and $\epsilon > 0$, it is possible to find an infinitely differentiable nonnegative function g that is equal to 1 in the interval $[x_0 - \epsilon, x_0 + \epsilon]$ and equal to 0 outside the interval $(x_0 - 2\epsilon, x_0 + 2\epsilon)$. See Figure 4.1 for an example.[16]

(iii) Conclude that G is the zero operator.

(iv) Show on the other hand that there are $f, g \in V_1$ such that $\langle Gf, g\rangle \neq 0$, so G cannot be the zero operator.

4.4. Consider the complex vector space $\mathbb{C}^\infty$ of finite complex sequences (i.e., sequences $(a_1, a_2, \dots)$ of complex numbers that have only finitely many nonzero terms) with the hermitian product

$$\langle a, b\rangle = \sum_{i=1}^{\infty} \overline{a}_i b_i.$$

Let $F\colon \mathbb{C}^\infty \to \mathbb{C}^\infty$ be the linear operator

$$F(b_1, b_2, \dots) = \left(\sum_{i=1}^{\infty} b_i, 0, 0, \dots\right).$$

Assume by contradiction that the adjoint $F^\dagger$ exists. Show that there is an $a \in \mathbb{C}^\infty$ such that

$$\left\langle F^\dagger a, e_i\right\rangle \neq 0 \ \forall i,$$

[16] The colored portions in the figure correspond to the following functions:

$$y_{\text{black}} = 0, \quad y_{\text{red}} = \frac{1}{1 + \mathrm{e}^{\left(\frac{1}{x} - \frac{1}{1-x}\right)}}, \quad y_{\text{yellow}} = 1, \quad y_{\text{blue}} = \frac{1}{1 + \mathrm{e}^{\left(\frac{1}{4-x} - \frac{1}{x-3}\right)}}.$$

Note that these functions join smoothly (i.e., all left and right derivatives of any order coincide) at the points $x = 0, 1, 3, 4$.

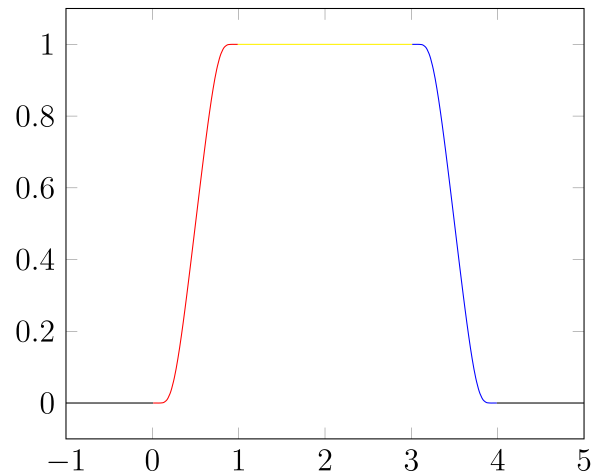


FIGURE 4.1. A function g for $x_0 = 2$ and $\epsilon = 1$

where e_i is the basis element of $\mathbb{C}^\infty$ given by the sequence whose ith term is equal to 1 and whose every other term is equal to 0. Conclude that this is a contradiction.[17]

4.5. The goal of this exercise is to prove the following statement:

If all the leading principal minors of a self-adjoint matrix $\boldsymbol{g}$ are positive, then $\boldsymbol{g}$ is positive definite.

We prove it by induction on the size of the matrix $\boldsymbol{g}$.

(a) Show that the statement is true if $\boldsymbol{g}$ is a 1×1 matrix.

(b) Assume that the statement holds for $n \times n$ matrices and let $\boldsymbol{g}$ be an $(n+1) \times (n+1)$ self-adjoint matrix satisfying the condition in the statement. Write

$$\boldsymbol{g} = \begin{pmatrix} \boldsymbol{h} & \boldsymbol{b} \\ \bar{\boldsymbol{b}}^\mathsf{T} & a \end{pmatrix}$$

with $\boldsymbol{h}$ a self-adjoint $n \times n$ matrix, $\boldsymbol{b}$ an n-column complex vector and a a real number.

(i) Show that $\boldsymbol{h}$ is positive definite, so there is an invertible matrix $\boldsymbol{E}$ such that $\boldsymbol{h} = \boldsymbol{E}^\dagger \boldsymbol{E}$.

Hint: Use the induction hypothesis.

[17] Another way to interpret this exercise is in terms of representing matrices. Using the given basis, we define the representing matrix $\boldsymbol{A}$ of an endomorphism F as the infinite matrix with entries A_i^j defined by $Fe_i = \sum_{j=1}^\infty A_i^j e_j$. Note that an infinite matrix may represent an endomorphism iff each of its columns has only finitely many nonzero entries. The transpose of such a matrix may violate this condition.

(ii) Show that

$$a > \|\boldsymbol{F}\boldsymbol{b}\|_{\mathrm{H}}^2,$$

where $\boldsymbol{F} := \boldsymbol{E}^{\dagger,-1}$ and $\|\boldsymbol{v}\|_{\mathrm{H}} := \sqrt{\bar{\boldsymbol{v}}^{\mathsf{T}}\boldsymbol{v}}$ denotes the standard hermitian norm on $\mathbb{C}^n$.
Hint: Use the identity on determinants of block matrices presented in exercise 3.12.

(iii) For a fixed n-column complex vector $\boldsymbol{w}$ consider the real-valued function

$$f(z) := \begin{pmatrix} \bar{\boldsymbol{w}}^{\mathsf{T}} & \bar{z} \end{pmatrix} \boldsymbol{g} \begin{pmatrix} \boldsymbol{w} \\ z \end{pmatrix}, \quad z \in \mathbb{C}.$$

(A) Setting $z = u + \mathrm{i}v$ and $\bar{\boldsymbol{b}}^{\mathsf{T}}\boldsymbol{w} = \alpha + \mathrm{i}\beta$—with u, v, α, and β real—show that

$$f(z) = a(u^2 + v^2) + 2(\alpha u + \beta v) + \|\boldsymbol{E}\boldsymbol{w}\|_{\mathrm{H}}^2.$$

(B) Show that the minimum value of f, as a function of $(u, v) \in \mathbb{R}^2$, is

$$f_{\min} = \|\boldsymbol{E}\boldsymbol{w}\|_{\mathrm{H}}^2 - \frac{|\bar{\boldsymbol{b}}^{\mathsf{T}}\boldsymbol{w}|^2}{a}.$$

(C) Assuming $\bar{\boldsymbol{b}}^{\mathsf{T}}\boldsymbol{w} \neq 0$, show that

$$f_{\min} > \frac{\|\boldsymbol{F}\boldsymbol{b}\|_{\mathrm{H}}^2\|\boldsymbol{E}\boldsymbol{w}\|_{\mathrm{H}}^2 - |\bar{\boldsymbol{b}}^{\mathsf{T}}\boldsymbol{w}|^2}{\|\boldsymbol{F}\boldsymbol{b}\|_{\mathrm{H}}^2}.$$

Hint: Use point 5(b)ii.

(D) Show that

$$\|\boldsymbol{F}\boldsymbol{b}\|_{\mathrm{H}}^2\|\boldsymbol{E}\boldsymbol{w}\|_{\mathrm{H}}^2 \geq |\bar{\boldsymbol{b}}^{\mathsf{T}}\boldsymbol{w}|^2.$$

Hint: Use the Cauchy–Schwarz inequality for the standard hermitian product.

(iv) Conclude that $\boldsymbol{g}$ is positive definite.

4.6. Consider the following matrices:

$$\boldsymbol{A} = \frac{1}{2}\begin{pmatrix} 1 & -i & -1+i \\ i & 1 & 1+i \\ 1+i & -1+i & 0 \end{pmatrix}, \quad \boldsymbol{B} = \frac{1}{2}\begin{pmatrix} 1 & -i & 1-i \\ i & 1 & 1-i \\ 1+i & -1-i & 0 \end{pmatrix},$$

$$\boldsymbol{C} = \begin{pmatrix} 5/3 & 2i/3 & -2i/3 \\ -2i/3 & 5/3 & -2/3 \\ 2i/3 & -2/3 & 5/3 \end{pmatrix}, \quad \boldsymbol{D} = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 1 & 1 \\ 1 & \frac{1}{2}(i-\sqrt{3}) & \frac{i}{2}(i+\sqrt{3}) \\ 1 & \frac{i}{2}(i+\sqrt{3}) & \frac{1}{2}(i-\sqrt{3}) \end{pmatrix}.$$

(a) Which of them is unitary and/or self adjoint?

(b) For each unitary and/or self-adjoint matrix in the list, find an orthonormal basis of eigenvectors.

4.7. The goal of this exercise is to show that two *normal* endomorphisms F and G on a finite-dimensional hermitian product space V are simultaneously diagonalizable—i.e., possess a common *orthonormal* basis of eigenvectors—iff they commute—i.e., $FG = GF$.[18]

(a) Assume that F and G have a common orthonormal basis of eigenvectors. Show that they commute.

(b) Now assume that F and G commute.

(i) Show that $F^\dagger$ and $G^\dagger$ commute.

(ii) Let λ be an eigenvalue of F. Show that $\mathrm{Eig}(F, \lambda)$ is a G-invariant and $G^\dagger$-invariant subspace and that the restriction of G to it is a normal operator.
Hint: Note that $\mathrm{Eig}(F, \lambda)$ is also an eigenspace for $F^\dagger$.

(iii) Conclude that it is possible to find a common eigenvector v of F and G.

(iv) Let v be a common eigenvector v of F and G. Show that $v^\perp$ is invariant under F, $F^\dagger$, G, $G^\dagger$, and that the restrictions of F and G to it are normal operators that commute with each other.

(v) Show that F and G possess a common orthonormal basis of eigenvectors.
Hint: Proceed by induction on the dimension of V.

4.8. Let

$$\boldsymbol{E} = \frac{1}{3}\begin{pmatrix} 4 & \frac{i}{2}(i+\sqrt{3}) & \frac{i}{2}(i-\sqrt{3}) \\ \frac{i}{2}(i-\sqrt{3}) & 4 & \frac{i}{2}(i+\sqrt{3}) \\ \frac{i}{2}(i+\sqrt{3}) & \frac{i}{2}(i-\sqrt{3}) & 4 \end{pmatrix} \quad \text{and} \quad \boldsymbol{F} = \frac{1}{3}\begin{pmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{pmatrix}.$$

(a) Show that $\boldsymbol{E}$ and $\boldsymbol{F}$ commute.

(b) Show $\boldsymbol{E}$ and $\boldsymbol{F}$ are simultaneously diagonalizable by finding a common orthonormal basis of eigenvectors.

[18] This result is of fundamental importance for quantum mechanics.

Index