Differential Geometry - Dynamical Systems
Monographs # 2

Constantin Udrişte

Ioana Boca

LINEAR ALGEBRA

Geometry Balkan Press

Bucharest, Romania

Linear Algebra
Monographs # 2

Differential Geometry - Dynamical Systems * Monographs
Editor-in-Chief Prof.Dr. Constantin Udrişte
Politehnica University of Bucharest

Analytic and Differential Geometry
Constantin Udrişte, Ioana Boca. - Bucharest:
Differential Geometry - Dynamical Systems * Monographs, 2000

Includes bibliographical references.

Preface

This textbook covers the standard linear algebra material taught at the University Politehnica of Bucharest, and is designed for a 1-semester course.

The prerequisites are high–school algebra and geometry.

Chapters 1–4 are intended to introduce first year students to the basic notions of vector space, linear transformation, eigenvectors and eigenvalues, bilinear and quadratic forms, and to the usual linear algebra techniques.

The linear algebra language is used in Chapters 5, 6, 7 to present some notions and results on vectors, straight lines and planes, transformations of coordinate systems.

We end with some exam samples. Each sample involves facts from two or more chapters.

The topics treated in this book and the presentation of the material are similar to those in several of the first author's previous works [19]–[25]. Parts of some linear algebra sections follow [1]. The selection of topics and problems relies on the teaching experience of the authors at the University Politehnica of Bucharest, including lectures and seminars taught in English at the Department of Engineering Sciences.

The publication of this volume was supported by MEN Grant #21815, 28.09.98, CNCSU-31; this support provided the oportunity to include the present textbook in the University Lectures Series published by the Editorial House of Balkan Society of Geometers.

We wish to thank our colleagues for helpful discussions on the problems and topics treated in this book and on our teaching activities. Any further suggestions will be greatly appreciated.

The authors

July 12, 2000

Contents

Chapter 1

Vector Spaces

1 Vector Spaces

The vector space structure is one of the most important algebraic structures.
The basic models for (real) vector spaces are the spaces of n–dimensional row or column matrices:

$$M_{1,n}(\mathbf{R}) = \left\{ v = [a_1, \ldots, a_n] \ ; \ a_j \in \mathbf{R}, \ j = \overline{1, n} \right\}$$

$$M_{n,1}(\mathbf{R}) = \left\{ v = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \ ; \ a_j \in \mathbf{R}, \ j = \overline{1, n} \right\} .$$

We will identify $\mathbf{R}^n$ with either one of $M_{1,n}(\mathbf{R})$ or $M_{n,1}(\mathbf{R})$. A row (column) matrix is also called a *row (column) vector*.

The definition of matrix multiplication makes the use of column vectors more convenient for us. We will also write a column vector in the form ${}^t[a_1, \ldots, a_n]$ or ${}^t(a_1, \ldots, a_n)$, as the transpose of a row vector, in order to save space.

An abstract vector space is endowed with two operations: addition of vectors and multiplication by scalars, where the scalars will be the elements of a field. For the examples above, these are just the usual matrix addition and multiplication of a matrix by a real number:

$$\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{bmatrix}$$

$$k \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} ka_1 \\ \vdots \\ ka_n \end{bmatrix} , \quad \text{where } k \in \mathbf{R} .$$

Let us first recall the definition of a field.

DEFINITION 1.1 A *field* $\mathbf{K}$ is a set endowed with two laws of composition:

$$\begin{aligned} \mathbf{K} \times \mathbf{K} \longrightarrow \mathbf{K}, \quad (a,b) \longrightarrow a+b \\ \mathbf{K} \times \mathbf{K} \longrightarrow \mathbf{K}, \quad (a,b) \longrightarrow ab \ , \end{aligned}$$

called *addition* and respectively *multiplication*, satisfying the following axioms :

i) $(\mathbf{K},+)$ is an abelian group. Its identity element is denoted by 0.

ii) $(\mathbf{K}^\star,\cdot)$ is an abelian group, where $\mathbf{K}^\star = \mathbf{K} \setminus \{0\}$. Its identity element is denoted by 1.

iii) Distributive law: $(a+b)c = ac+bc, \quad \forall\, a,b,c \in \mathbf{K}$.

Examples of fields

(*a*) $\mathbf{K} = \mathbf{R}$, the field of real numbers;

(*b*) $\mathbf{K} = \mathbf{C}$, the field of complex numbers;

(*c*) $\mathbf{K} = \mathbf{Q}$, the field of rational numbers;

(*d*) $\mathbf{K} = \mathbf{Q}[\sqrt{2}] = \{a + b\sqrt{2}\ ;\ a,b \in \mathbf{Q}\}$.

DEFINITION 1.2 A *vector space* $\mathbf{V}$ over a field $\mathbf{K}$ (or a $\mathbf{K}$–vector space) is a set endowed with two laws of composition:

(*a*) *addition* : $\mathbf{V} \times \mathbf{V} \longrightarrow \mathbf{V}, \quad (v,w) \longrightarrow v+w$

(*b*) *multiplication by scalars* : $\mathbf{K} \times \mathbf{V} \longrightarrow \mathbf{V}, \quad (k,v) \longrightarrow kv$

satisfying the following axioms:

(*i*) $(\mathbf{V},+)$ is an abelian group

(*ii*) multiplication by scalars is associative with respect to multiplication in $\mathbf{K}$:

$$k(lv) = (kl)v, \quad \forall k,l \in \mathbf{K},\ \forall v \in \mathbf{V}$$

(*iii*) the element $1 \in \mathbf{K}$ acts as identity:

$$1v = v, \quad \forall v \in \mathbf{V}$$

(*iv*) the distributive laws hold:

$$\begin{aligned} &k(v+w) = kv+kw, \\ &(k+l)v = kv+lv, \qquad \forall k,l \in \mathbf{K},\ \forall v,w \in \mathbf{V}. \end{aligned}$$

The elements of $\mathbf{K}$ are usually called *scalars* and the elements of $\mathbf{V}$ *vectors*. A vector space over $\mathbf{C}$ is called a *complex vector space*; a vector space over $\mathbf{R}$ is called a *real vector space*. When $\mathbf{K}$ is not specified, we understand $\mathbf{K} = \mathbf{R}$ or $\mathbf{K} = \mathbf{C}$.

Examples of vector spaces

(i) $\mathbf{K}$ as a vector space over itself, with addition and multiplication in $\mathbf{K}$.

(ii) $\mathbf{K}^n$ is a $\mathbf{K}$–vector space.

(iii) $M_{m,n}(\mathbf{K})$ with usual matrix addition and multiplication by scalars is a $\mathbf{K}$–vector space.

(iv) The set of space vectors $\mathbf{V}_3$ is a real vector space with addition given by the parallelogram law and scalar multiplication given as follows:

if $k \in \mathbf{R}$, $\overline{v} \in \mathbf{V}_3$, then $k\overline{v} \in \mathbf{V}_3$ and:

– the length of $k\overline{v}$ is the length of $\overline{v}$ multiplied by $\mid k \mid$;

– the direction of $k\overline{v}$ is the direction of $\overline{v}$;

– the sense of $k\overline{v}$ is the sense of $\overline{v}$ if $k > 0$, and the opposite sense of $\overline{v}$ if $k < 0$.

(v) If $\mathbf{S}$ is a nonempty set and $\mathbf{W}$ is a $\mathbf{K}$–vector space, then $\mathbf{V} = \{f \mid f : \mathbf{S} \longrightarrow \mathbf{W}\}$ becomes a $\mathbf{K}$–vector space with:

$$(f+g)(x) = f(x) + g(x), \quad (kf)(x) = kf(x), \qquad \text{for all } k \in \mathbf{K},\ f, g \in \mathbf{V}.$$

(vi) The solution set of an algebraic linear homogeneous system with n unknowns and coefficients in $\mathbf{K}$ is a $\mathbf{K}$–vector space with the operations induced from $\mathbf{K}^n$.

(vii) The solution set of an ordinary linear homogeneous differential equation is a real vector space with addition of functions and multiplication of functions by scalars.

($viii$) The set $\mathbf{K}[X]$ of all polynomials with coefficients in $\mathbf{K}$ is a $\mathbf{K}$–vector space.

THEOREM 1.3 *Let* $\mathbf{V}$ *be a* $\mathbf{K}$*–vector space. Then* $\mathbf{V}$ *has the following properties:*

(i) $0_K\, v = 0_V$, $\forall v \in \mathbf{V}$

(ii) $k\, 0_V = 0_V$, $\forall k \in \mathbf{K}$

(iii) $(-1)\, v = -v$, $\forall v \in \mathbf{V}$.

Proof. To see (i) we use the distributive law to write

$$0_K\, v + 0_K\, v = (0_K + 0_K)\, v = 0_K\, v + 0_V\ .$$

Now $0_K\, v$ cancels out (in the group $(\mathbf{V}, +)$), so we obtain $0_K\, v = 0_V$. Similarly,

$$k\, 0_V + k\, 0_V = k(0_V + 0_V) = k\, 0_V$$

implies $k\, 0_V = 0_V$.

For (iii), $v + (-1)\, v = (1 + (-1))\ v = 0_K\, v = 0_V$. Hence $(-1)\, v$ is the additive inverse of v. QED

COROLLARY 1.4 *Let* $\mathbf{V}$ *be a* $\mathbf{K}$*–vector space. Then:*

(i) $-(k\,v) = (-k)\,v = k\,(-v)$, *for all* $k \in \mathbf{K},\ v \in \mathbf{V}$

(ii) $(k-l)\,v = k\,v - l\,v$, *for all* $k \in \mathbf{K},\ v \in \mathbf{V}$

(iii) $k\,v = 0_V \iff v = 0_V$ *or* $k = 0_K$

(iv) $k\,v = l\,v,\ \ v \neq 0_V \implies k = l$

(v) $k\,v = k\,w,\ \ k \neq 0_K \implies v = w$

We leave the proof of the corollary as an exercise; all properties follow easily from Theorem 1.3 and the definition of a vector space.

We will usually write "0" instead of 0_K or 0_V, since in general it is clear whether we refer to the number or to the vector zero.

2 Vector Subspaces

Throughout this paragraph $\mathbf{V}$ denotes a $\mathbf{K}$–vector space and $\mathbf{S}$ a nonempty subset of $\mathbf{V}$.

DEFINITION 2.1 A *vector subspace* $\mathbf{W}$ of $\mathbf{V}$ is a nonempty subset with the following properties:

(i) If $u, v \in \mathbf{W}$, then $u + w \in \mathbf{W}$

(ii) If $u \in \mathbf{W}$, $k \in \mathbf{K}$, then $k\,u \in \mathbf{W}$.

REMARKS 2.2 (i) If $\mathbf{W}$ is a vector subspace, then $0_V \in \mathbf{W}$.

(ii) $\mathbf{W}$ is a vector subspace if and only if $\mathbf{W}$ is a vector space with the operations induced from $\mathbf{V}$.

(iii) $\mathbf{W}$ is a vector subspace if and only if

$$k, l \in \mathbf{K},\ \ u, w \in \mathbf{V} \implies k\,u + l\,w \in \mathbf{W}\ .$$

Examples of vector subspaces

(i) $\{0_V\}$ and $\mathbf{V}$ are vector subspaces of $\mathbf{V}$. Any other vector subspace is called a *proper vector subspace.*

(ii) The straight lines through the origin are proper vector subspaces of $\mathbf{R}^2$.

(iii) The straight lines and the planes through the origin are proper vector subspaces of $\mathbf{R}^3$.

(iv) The solution set of an algebraic linear homogeneous system with n unknowns is a vector subspace of $\mathbf{K}^n$.

(v) The set of odd functions and the set of even functions are vector subspaces of the space of real functions defined on the interval $(-a, a)$, $a \in \mathbf{R}_+^\star$.

DEFINITION 2.3 (i) Let $v_1, \ldots, v_p \in \mathbf{V}$. A *linear combination* of the vectors $v_1, \ldots, v_p$ is a vector of the form

$$w = k_1 v_1 + \ldots + k_p v_p \ , \quad k_j \in \mathbf{K} \ .$$

(ii) Let $\mathbf{S}$ be a nonempty subset (possibly infinite) of $\mathbf{V}$. A vector of the form

$$w = k_1 v_1 + \ldots + k_p v_p \ , \quad p \in \mathbf{N}^\star, \ k_j \in \mathbf{K}, \ v_j \in \mathbf{S}$$

is called a *finite linear combination of vectors of* $\mathbf{S}$.

The set of all finite linear combinations of vectors of $\mathbf{S}$ is denoted by *Span* $\mathbf{S}$ or $L(\mathbf{S})$. By definition, $Span\, \phi = \{0\}$.

THEOREM 2.4 Span $\mathbf{S}$ *is a vector subspace of* $\mathbf{V}$.

Proof. Let $u = \sum_{i=1}^{p} k_i u_i \in \text{Span}\, \mathbf{S}$ and $w = \sum_{j=1}^{q} l_j w_j \in \text{Span}\, \mathbf{S}$, with $k_i, l_j \in \mathbf{K}$, $u_i, w_j \in \mathbf{S}$.

Then $ku + lv = \sum_{i=1}^{p} (kk_i)\, u_i + \sum_{j=1}^{q} (ll_j)\, w_j$ is also a finite linear combination with elements of $\mathbf{S}$. It follows that *Span* $\mathbf{S}$ is a vector subspace, by Remark 2.2(iii). QED

Span $\mathbf{S}$ is called the *subspace spanned* by $\mathbf{S}$, or the *linear covering* of $\mathbf{S}$.

PROPOSITION 2.5 *If* $\mathbf{S}$ *is a subset of* $\mathbf{V}$ *and* $\mathbf{W}$ *is a vector subspace of* $\mathbf{V}$ *such that* $\mathbf{S} \subset \mathbf{W}$*, then* Span $\mathbf{S} \subset \mathbf{W}$.

This is obvious, for any finite linear combination of vectors from $\mathbf{S}$ is also a finite linear combination of vectors from $\mathbf{W}$, and $\mathbf{W}$ is closed under addition and scalar multiplication.

THEOREM 2.6 *If* $\mathbf{W}_1$, $\mathbf{W}_2$ *are two vector subspaces of* $\mathbf{V}$*, then:*

(i) $\mathbf{W}_1 + \mathbf{W}_2 = \{v = v_1 + v_2 \mid v_1 \in \mathbf{W}_1,\ v_2 \in \mathbf{W}_2\}$ *is a vector subspace, called the* `sum` *of the vector subspaces* $\mathbf{W}_1$ *and* $\mathbf{W}_2$.

(ii) $\mathbf{W}_1 \cap \mathbf{W}_2$ *is a vector subspace of* V*. More generally, if* $\{\mathbf{W}_i\}_{i \in I}$ *is a family of vector subspaces, then* $\bigcap_{i \in I} \mathbf{W}_i$ *is a vector subspace.*

(iii) $\mathbf{W}_1 \cup \mathbf{W}_2$ *is a vector subspace of* $\mathbf{V}$ *if and only if* $\mathbf{W}_1 \subseteq \mathbf{W}_2$ *or* $\mathbf{W}_2 \subseteq \mathbf{W}_1$.

Proof. (i) Let $u,\ w \in \mathbf{W}_1 + \mathbf{W}_2$, $u = u_1 + u_2$, $w = w_1 + w_2$, with $u_1, w_1 \in \mathbf{W}_1$, $u_2, w_2 \in \mathbf{W}_2$.

Let $k, l \in \mathbf{K}$. Then $ku + lw = (ku_1 + lw_1) + (ku_2 + lw_2)$ is in $\mathbf{W}_1 + \mathbf{W}_2$, as $ku_1 + lw_1 \in \mathbf{W}_1$, and $ku_2 + lw_2 \in \mathbf{W}_2$ (using Remark 2.2 (iii) for the subspaces $\mathbf{W}_1$, $\mathbf{W}_2$). It follows that $\mathbf{W}_1 + \mathbf{W}_2$ is a subspace.

(ii) Exercise !

(iii) The implication from the right to the left is obvious. For the other one, assume by contradiction that none of the inclusions holds. Then we can take $u_1 \in \mathbf{W}_1 \setminus \mathbf{W}_2$, $u_2 \in \mathbf{W}_2 \setminus \mathbf{W}_1$.

Now $u_1, u_2 \in \mathbf{W}_1 \cup \mathbf{W}_2$ implies $u_1 + u_2 \in \mathbf{W}_1 \cup \mathbf{W}_2$ since $\mathbf{W}_1 \cup \mathbf{W}_2$ is a vector subspace.

But either $u_1 + u_2 \in \mathbf{W}_1$ or $u_1 + u_2 \in \mathbf{W}_2$ contradicts $u_2 \notin \mathbf{W}_1$ or $u_1 \notin \mathbf{W}_2$ respectively. Therefore $\mathbf{W}_1 \subseteq \mathbf{W}_2$ or $\mathbf{W}_2 \subseteq \mathbf{W}_1$. QED

REMARKS 2.7 (i) Span $(\mathbf{W}_1 \cup \mathbf{W}_2) = \mathbf{W}_1 + \mathbf{W}_2$.

(ii) Span $(\mathbf{S}_1 \cup \mathbf{S}_2) = \operatorname{Span} \mathbf{S}_1 + \operatorname{Span} \mathbf{S}_2$.

These are straightforward from Proposition 2.5 and Theorem 2.6

PROPOSITION 2.8 *Let* $\mathbf{W}_1$, $\mathbf{W}_2$ *be vector subspaces of* $\mathbf{V}$*. Then, the decomposition*

$$v = v_1 + v_2\,, \quad v_1 \in \mathbf{W}_1\,,\ v_2 \in \mathbf{W}_2$$

is unique for each $v \in \mathbf{W}_1 + \mathbf{W}_2$ *if and only if* $\mathbf{W}_1 \cap \mathbf{W}_2 = \{0\}$.

Proof. Assume the uniqueness of the decomposition and let $v \in \mathbf{W}_1 \cap \mathbf{W}_2 \subseteq \mathbf{W}_1 + \mathbf{W}_2$. Then

$$v = v + 0\,, \quad v \in \mathbf{W}_1\,,\ 0 \in \mathbf{W}_2$$

and

$$v = 0 + v\,,\ 0 \in \mathbf{W}_1\,,\ v \in \mathbf{W}_2$$

represent the same decomposition, implying $v = 0$.

Conversely, asume $\mathbf{W}_1 \cap \mathbf{W}_2 = \{0\}$ and let $v \in \mathbf{W}_1 + \mathbf{W}_2$, $v = v_1 + v_2 = v_1' + v_2'$ with $v_1, v_1' \in \mathbf{W}_1$, $v_2, v_2' \in \mathbf{W}_2$. Then $v_1 - v_1' = v_2' - v_2 \in \mathbf{W}_1 \cap \mathbf{W}_2$, thus $v_1 - v_1' = 0 = v_2 - v_2'$. Consequently, $v_1 = v_1'$, $v_2 = v_2'$. QED

DEFINITION 2.9 If $\mathbf{W}_1$, $\mathbf{W}_2$ are vector subspaces with $\mathbf{W}_1 \cap \mathbf{W}_2 = \{0\}$, then $\mathbf{W}_1 + \mathbf{W}_2$ is called the *direct sum* of $\mathbf{W}_1$ and $\mathbf{W}_2$ and we write $\mathbf{W}_1 \oplus \mathbf{W}_2$; $\mathbf{W}_1$ and $\mathbf{W}_2$ are called *independent vector subspaces.*

If $\mathbf{V} = \mathbf{W}_1 \oplus \mathbf{W}_2$, then $\mathbf{W}_1$ and $\mathbf{W}_2$ are called *supplementary subspaces*; each of $\mathbf{W}_1$ and $\mathbf{W}_2$ is the *supplement* of the other one.

Examples

(i) Set

$$\begin{aligned} &\mathbf{V} = \{f \mid f : (-a, a) \longrightarrow \mathbf{R}\}\,, \quad a > 0\,,\\ &\mathbf{W}_1 = \{f \in \mathbf{V} \mid f \text{ is an odd function}\}\,,\\ &\mathbf{W}_2 = \{f \in \mathbf{V} \mid f \text{ is an even function}\}\,. \end{aligned}$$

Then $\mathbf{V} = \mathbf{W}_1 \oplus \mathbf{W}_2$.

(To see this, write $f(x) = \dfrac{f(x) - f(-x)}{2} + \dfrac{f(x) + f(-x)}{2}$, for any $f \in \mathbf{V}$ and any $x \in (-a, a)$.)

(*ii*) Set

$$\mathbf{D}_1 = \{(x,y) \mid 2x+y=0\}$$
$$\mathbf{D}_2 = \{(x,y) \mid x-y=0\}$$

Then $\mathbf{R}^2 = \mathbf{D}_1 \oplus \mathbf{D}_2$.

3 Linear Dependence. Linear Independence

Let $\mathbf{V}$ be a $\mathbf{K}$–vector space and $\mathbf{S} \subseteq \mathbf{V}$, $\mathbf{S} \neq \emptyset$.

DEFINITION 3.1 (*i*) Let $\mathbf{S} = \{v_1, \ldots, v_p\}$ be a finite subset of $\mathbf{V}$, where $v_i \neq v_j$ if $i \neq j$. The set $\mathbf{S}$ is called *linearly dependent* (the vectors $v_1, \ldots, v_p$ are called linearly dependent) if there exist $k_1, \ldots, k_p \in \mathbf{K}$ not all zero such that

$$\sum_{i=1}^{p} k_i v_i = 0 \ .$$

(*ii*) An infinite subset $\mathbf{S}$ of $\mathbf{V}$ is called *linearly dependent* if there exists a finite subset of $\mathbf{S}$ which is linearly dependent.

(*iii*) A set which is not linearly dependent is called *linearly independent*.

In order to study the linear dependence or independence of the vectors $v_1, \ldots, v_p$ we usually exploit the equality $\sum\limits_{i=1}^{p} k_i v_i = 0$. If this relation implies the existence of $(k_1, \ldots, k_p) \neq (0, \ldots, 0)$, then the vectors $v_1, \ldots, v_p$ are linearly dependent; if the relation implies $(k_1, \ldots, k_p) = (0, \ldots, 0)$, then the vectors are linearly independent.

REMARKS 3.2 (*i*) Let $\mathbf{S}$ be an arbitrary nonempty subset of $\mathbf{V}$. Then $\mathbf{S}$ is linearly dependent if and only if $\mathbf{S}$ contains a linearly dependent subset.

(*ii*) Let $\mathbf{S}$ be an arbitrary nonempty subset of $\mathbf{V}$. Then $\mathbf{S}$ is linearly independent if and only if all finite subsets of $\mathbf{S}$ are linearly independent.

(*iii*) Let $v_1, \ldots, v_n \in \mathbf{K}^m$. Denote by $A = [v_1, \ldots, v_n] \in M_{m,n}(\mathbf{K})$, the matrix whose columns are $v_1, \ldots, v_n$. Then $v_1, \ldots, v_n$ are linearly independent if and only if the system

$$A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = 0$$

admits only the trivial solution, and this is equivalent to $rank\, A = n$.

Examples

(*i*) $\{0\}$ is linearly dependent; if $0 \in \mathbf{S}$, then $\mathbf{S}$ is linearly dependent.

(*ii*) If $v \in \mathbf{V}$, $v \neq 0$, then $\{v\}$ is linearly independent.

(*iii*) A set $\{v_1, v_2\}$ of two vectors is linearly dependent if and only if either $v_1 = 0$ or else v_2 is a scalar multiple of v_1.

(*iv*) If $f_1(x) = e^x$; $f_2(x) = e^{-x}$; $f_3(x) = \operatorname{sh} x$, then $\{f_1, f_2, f_3\}$ is linearly dependent since $f_1 - f_2 - 2f_3 = 0$.

PROPOSITION 3.3 *Let* $\mathbf{L}$, $\mathbf{S}$ *be nonempty subsets of* $\mathbf{V}$, $v \in \mathbf{V} \setminus \mathbf{L}$ *and* $\mathbf{L}$ *linearly independent. Then:*

(*i*) $\mathbf{S}$ *is linearly dependent if and only if there exists an element* $w \in \mathbf{S}$ *such that* $w \in Span\,(\mathbf{S} \setminus \{w\})$.

(*ii*) $\mathbf{L} \cup \{v\}$ *is linearly independent if and only if* $v \notin Span\,\mathbf{L}$.

Proof. (*i*) Assume that $\mathbf{S}$ is linearly dependent. Then, there exist distinct elements $v_1, \ldots, v_p \in \mathbf{S}$ and $k_1, \ldots, k_p \in \mathbf{K}$ not all zero, such that

$$\sum_{i=1}^{p} k_i v_i = 0 \ .$$

We may assume without loss of generality that $k_1 \neq 0$. Then

$$v_1 = -\sum_{i=2}^{p} \frac{k_i}{k_1} v_i \in Span\,(\mathbf{S} - \{v_1\}) \ .$$

Conversely, if $w \in \mathbf{S}$ and $w \in Span\,(\mathbf{S} \setminus \{w\})$, then it can be written as $w = \sum_{j=1}^{q} c_j w_j$ for some $q \geq 1$, $c_j \in \mathbf{K}$, $w_j \in \mathbf{S} \setminus \{w\}$. It follows that

$$w + \sum_{j=1}^{q} (-c_j) w_j = 0 \ ,$$

therefore the set $\{w, w_1, \ldots, w_q\}$ is linearly dependent. Then $\mathbf{S}$ is linearly dependent since it contains a linearly dependent subset.

(*ii*) Now (*i*) applies for $\mathbf{S} = \mathbf{L} \cup \{v\}$, $\mathbf{S} \setminus \{v\} = \mathbf{L}$. Since L is linearly independent, it follows that $\mathbf{L} \cup \{v\}$ is linearly dependent if and only if $v \in Span\,\mathbf{L}$, or equivalently $\mathbf{L} \cup \{v\}$ is linearly independent if and only if $v \notin Span\,\mathbf{L}$. QED

PROPOSITION 3.4 *Let* $\mathbf{S} = \{v_1, \ldots, v_p\} \subset \mathbf{V}$ *be a linearly independent set and denote by* $Span\,\mathbf{S}$ *the vector subspace spanned by* $\mathbf{S}$. *Then any distinct* $p+1$ *elements of* $Span\,\mathbf{S}$ *are linearly dependent.*

Proof. Let $w_j = \sum_{i=1}^{p} a_{ij} v_i \in Span\,\mathbf{S}$, $j = 1, \ldots, p+1$ and $k_1, \ldots, k_{p+1} \in \mathbf{K}$ such that

$$\sum_{j=1}^{p+1} k_j w_j = 0 \ . \tag{3.1}$$

We replace the vectors w_j to obtain

$$\sum_{j=1}^{p+1} k_j w_j = \sum_{j=1}^{p+1} k_j \sum_{i=1}^{p} a_{ij} v_i = \sum_{i=1}^{p} \left(\sum_{j=1}^{p+1} k_j a_{ij} \right) v_i \ .$$

Then $\sum_{i=1}^{p}\left(\sum_{j=1}^{p+1} k_j a_{ij}\right) v_i = 0$ is equivalent to

$$\sum_{j=1}^{p+1} k_j a_{ij} = 0 \ , \quad i = 1, \ldots, p \ , \tag{3.2}$$

since $v_1, \ldots, v_p$ are linearly independent.

But (3.2) is a linear homogeneous system with p equations and $p+1$ unknowns $k_1, \ldots, k_{p+1}$. Such a system admits nontrivial solutions. Thus, there exist $k_1, \ldots, k_{p+1}$, not all zero, such that (3.1) holds, which means that $w_1, \ldots, w_{p+1}$ are linearly dependent. QED

4 Bases and Dimension

V is a **K** vector space. Let $\mathbf{B} \subset \mathbf{V}$ be a linearly independent set. Then *Span* **B** is a vector subspace of **V**. A natural question concerns the existence of **B** linearly independent such that *Span* $\mathbf{B} = \mathbf{V}$.

DEFINITION 4.1 A subset **B** of **V** which is linearly independent and also spans **V**, is called a *basis* of **V**.

It can be shown that any nonzero vector space admits a basis. We are going to use this general fact without proof. The results we are proving in this section concern only a certain class of vector spaces – the finitely generated ones.

DEFINITION 4.2 **V** is called *finitely generated* if there exists a finite set which spans **V**, or if $\mathbf{V} = \{0\}$.

Note that not all vector spaces are finitely generated. For example, the real vector space $\mathbf{R}_n[X]$ is finitely generated, while $\mathbf{R}[X]$ is not.

THEOREM 4.3 *Let* $\mathbf{V} \neq \{0\}$ *be a finitely generated vector space. Any finite set which spans* **V** *contains a basis.*

Proof. Let **S** be a finite set such that *Span* $\mathbf{S} = \mathbf{V}$ and set

$$\mathbf{S} = \{v_1, \ldots, v_p\} \ , \quad v_i \neq v_j \quad \text{if } i \neq j \ .$$

If **S** is linearly independent, then **S** is a basis. If **S** is not linearly independent, then there exist $k_1, \ldots, k_p \in \mathbf{K}$, not all zero, such that

$$k_1 v_1 + \ldots + k_p v_p = 0 \ .$$

Assume without loss of generality $k_p \neq 0$; then $v_p \in Span\{v_1, \ldots v_{p-1}\}$, which implies $Span\,\mathbf{S} = Span\{v_1, \ldots, v_{p-1}\} = \mathbf{V}$.

We repeat the above procedure for $\mathbf{S}_1 = \{v_1, \ldots, v_{p-1}\}$. Continuing this way, we eventually obtain a subset of **S** which is linearly dependent but still spans **V**. QED

COROLLARY 4.4 *Any nonzero finitely generated vector space admits a finite basis.*

Using Prop.3.4, the next consequences of the theorem are straightforward.

COROLLARY 4.5 *(i) Any linearly independent subset of a finitely generated vector space is finite.*
(ii) Any basis of a finitely generated vector space is finite.

THEOREM 4.6 *Let* $\mathbf{V}$ *be a finitely generated vector space,* $\mathbf{V} \neq \{0\}$. *Then any two bases have the same number of elements.*

Proof. Let $\mathbf{B}$, $\mathbf{B}'$ be two bases of $\mathbf{V}$. Assume that $\mathbf{B}$ has n elements and $\mathbf{B}'$ has n' elements. We have $\mathbf{V} = Span\,\mathbf{B}$, since the basis $\mathbf{B}$ spans $\mathbf{V}$. By Proposition 3.4, no linearly independent subset of $\mathbf{V}$ could have more than n elements. The basis $\mathbf{B}'$ is linearly independent, therefore $n' \leq n$.

The same argument works for $\mathbf{B}'$ instead of $\mathbf{B}$, yielding $n \leq n'$. Therefore $n = n'$. QED

DEFINITION 4.7 Let $\mathbf{V}$ be finitely generated.
If $\mathbf{V} \neq \{0\}$, then the *dimension* of $\mathbf{V}$ is the number of vectors in a basis of $\mathbf{V}$. The dimension is denoted by $dim\,\mathbf{V}$.
If $\mathbf{V} = \{0\}$, then $dim\,\{0\} = 0$ by definition.

Finitely generated vector spaces are also called *finite dimensional* vector spaces. The other vector spaces, which have infinite bases are called *infinite dimensional.*

The dimension of a finite dimensional vector space is a natural number; $dim\,\mathbf{V} = 0$ if and only if $\mathbf{V} = \{0\}$.

When it is necessary to specify the field, we write $dim_{\mathbf{K}}\mathbf{V}$ instead of $dim\,\mathbf{V}$. For example, $dim_{\mathbf{C}}\mathbf{C} = 1$, $dim_{\mathbf{R}}\mathbf{C} = 2$, since $\mathbf{C}$ can be regarded as a complex vector space, as well as a real vector space.

Examples

(*i*) $e_1 = (1, 0, \ldots, 0)$; $e_2 = (0, 1, \ldots, 0); \ldots$ $e_n = (0, 0, \ldots, 1)$ form a basis of $\mathbf{K}^n$, and $dim\,\mathbf{K}^n = n$.

(*ii*) $\mathbf{B} = \{E_{ij}\,|\,1 \leq i \leq m,\ 1 \leq j \leq n\}$ is a basis of $M_{m,n}(\mathbf{K})$, where E_{ij} is the matrix whose (i, j)–entry is 1 and all other entries are zero; $dim\,M_{m,n}(\mathbf{K}) = mn$.

(*iii*) $\mathbf{B} = \{1, X, \ldots, X^n\}$ is a basis of $\mathbf{K}_n[X]$; $dim\,\mathbf{K}_n[X] = n + 1$.

(*iv*) $\mathbf{B} = \{1, X, \ldots, X^n, \ldots\}$ is a basis of $\mathbf{K}[X]$; $\mathbf{K}[X]$ is infinite dimensional.

(*v*) The solution set of a linear homogeneous system of rank r with n unknowns and coefficients in $\mathbf{K}$ is a $\mathbf{K}$-vector space of dimension $n - r$.

THEOREM 4.8 *Let* $\mathbf{V} \neq \{0\}$ *be a finitely generated vector space. Any linearly independent subset of* $\mathbf{V}$ *is contained into a basis.*

Proof. Let $\mathbf{L}$ be a linearly independent subset of $\mathbf{V}$ and $\mathbf{S}$ be a finite set which spans $\mathbf{V}$. Since $\mathbf{V}$ is finitely generated, the set $\mathbf{L}$ is finite, by the previous corollary. If $\mathbf{S} \subset Span\,\mathbf{L}$, then $\mathbf{L}$ is a basis.

If $\mathbf{S} \not\subset Span\,\mathbf{L}$, then choose $v \in \mathbf{S}$, $v \notin Span\,\mathbf{L}$. It follows that $\mathbf{L} \cup \{v\}$ is linearly independent, by Proposition 3.3. Continue until we get a basis. QED

COROLLARY 4.9 *Let* $\mathbf{L}$, $\mathbf{S}$ *be finite subsets of* $\mathbf{V}$ *such that* $\mathbf{L}$ *is linearly independent and* $\mathbf{V} = Span\,(\mathbf{S})$. *Then:*

(i) $card\,\mathbf{L} \leq dim\,\mathbf{V} \leq card\,\mathbf{S}$;

(ii) $card\,\mathbf{L} = \dim\,\mathbf{V}$ *if and only if* $\mathbf{L}$ *is a basis;*

(iii) $card\,\mathbf{S} = \dim\,\mathbf{V}$ *if and only if* $\mathbf{S}$ *is a basis.*

PROPOSITION 4.10 *If* $\mathbf{W}$ *is a vector subspace of* $\mathbf{V}$ *and* $dim\,\mathbf{V} = n,\ n \geq 1$, *then* $\mathbf{W}$ *is finite–dimensional and* $dim\,\mathbf{W} \leq n$. *Equality holds only if* $\mathbf{W} = \mathbf{V}$.

Proof. Assume $\mathbf{W} \neq \{0\}$ and let $v_1 \in \mathbf{W}$, $v_1 \neq 0$. Then $\{v_1\}$ is linearly independent.

If $\mathbf{W} = Span\,\{v_1\}$, then we are done. If $\mathbf{W} \not\supseteq Span\,\{v_1\}$, then there exists $v_2 \in \mathbf{W} \setminus Span\,\{v_1\}$. Proposition 3.3 (ii) applies for $\mathbf{L} = \{v_1\}$, $v = v_2$, so $\{v_1, v_2\}$ is linearly independent.

Now either $\mathbf{W} = Span\{v_1, v_2\}$ or there exists $v_3 \in \mathbf{W} \setminus Span\{v_1, v_2\}$. In the latter case we apply again Proposition 3.3 (ii) and we continue the process. The process must stop after at most n steps, otherwise at step $n+1$ we would find v_{n+1} such that $\{v_1, \ldots, v_n, v_{n+1}\}$ is linearly independent, which contradicts Proposition 3.4.

By the above procedure we found a basis of $\mathbf{W}$ which contains at most n vectors, thus $\mathbf{W}$ is finite dimensional and $dim\,\mathbf{V} \leq n$.

Assume that $dim\,\mathbf{W} = n$. Then any basis of $\mathbf{W}$ is linearly independent in $\mathbf{V}$ and contains n elements; by the previous corollary it is also a basis of $\mathbf{V}$. QED

THEOREM 4.11 *If* $\mathbf{U}$, $\mathbf{W}$ *are finite–dimensional vector subspaces of* $\mathbf{V}$, *then* $\mathbf{U} + \mathbf{W}$ *and* $\mathbf{U} \cap \mathbf{W}$ *are finite dimensional and*

$$\dim\,\mathbf{U} + \dim\,\mathbf{W} = \dim\,(\mathbf{U} + \mathbf{W}) + dim\,(\mathbf{U} \cap \mathbf{W})\ .$$

Sketch of proof. The conclusion is obvious if $\mathbf{U} \cap \mathbf{W} = \mathbf{U}$ or $\mathbf{U} \cap \mathbf{W} = \mathbf{W}$. If not, assume $\mathbf{U} \cap \mathbf{W} \neq \{0\}$ and let $\{v_1, \ldots, v_p\}$ be a basis of $\mathbf{U} \cap \mathbf{W}$. Then, there exist $u_{p+1}, \ldots, u_{p+q} \in \mathbf{U}$ and $w_{p+1}, \ldots, w_{p+r} \in \mathbf{W}$ such that

$\{v_1, \ldots, v_p, u_{p+1}, \ldots, u_{p+q}\}$ is a basis of $\mathbf{U}$,

$\{v_1, \ldots, v_p, w_{p+1}, \ldots, w_{p+r}\}$ is a basis of $\mathbf{W}$,

and show that $\{v_1, \ldots, v_p, u_{p+1}, \ldots, u_{p+q}, w_{p+1}, \ldots, w_{p+r}\}$ is a basis of $\mathbf{U} + \mathbf{W}$.

The idea of proof is similar if $\mathbf{U} \cap \mathbf{W} = \{0\}$. QED

5 Coordinates. Isomorphisms. Change of Coordinates

Let $\mathbf{V}$ be a finite dimensional vector space. In this section we are going to make explicit computations with bases. For, it will be necessary to work with (finite) *ordered* sets of vectors. Consequently, a finite basis $\mathbf{B}$ of $\mathbf{V}$ has three qualities: it is linearly independent, it spans $\mathbf{V}$, and it is ordered.

PROPOSITION 5.1 *The set* $\mathbf{B} = \{v_1, \ldots, v_n\} \subset \mathbf{V}$ *is a basis if and only if every vector* $x \in \mathbf{V}$ *can be written in a unique way in the form*

$$x = x_1 v_1 + \ldots + x_n v_n \ , \quad x_j \in \mathbf{K} \ , \quad j = 1, \ldots, n \ . \tag{5.1}$$

Proof. Suppose that $\mathbf{B}$ is a basis. Then every $x \in \mathbf{V}$ can be written in the form (5.1) since $\mathbf{V} = Span\,\mathbf{B}$. If also $x = x'_1 v_1 + \ldots x'_n v_n$, then

$$0 = x - x = (x_1 - x'_1)v_1 + \ldots + (x_n - x'_n)v_n \ .$$

By the linear independence of $\mathbf{B}$ it follows that $x_1 - x'_1 = \ldots = x_n - x'_n = 0$.

Conversely, the existence of the representation (5.1) for each vector implies $\mathbf{V} = Span\,\mathbf{B}$. The uniqueness applied for $x = 0 = 0v_1 + \ldots + 0v_n$ gives the linear independence of $\mathbf{B}$. QED

DEFINITION 5.2 The scalars x_j are called the *coordinates* of x with respect to the basis $\mathbf{B}$. The column vector

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

is called the *coordinate vector* associated to x.

DEFINITION 5.3 Let $\mathbf{V}$ and $\mathbf{W}$ be two vector spaces over the same field $\mathbf{K}$. A map $\mathcal{T} : \mathbf{V} \longrightarrow \mathbf{W}$ compatible with the vector space operations, i.e. satisfying:

(a) $\mathcal{T}(x + y) = \mathcal{T}(x) + \mathcal{T}(y)$, $\forall\, x, y \in \mathbf{V}$ ($\mathcal{T}$ *is additive*)

(b) $\mathcal{T}(kx) = k\mathcal{T}(x)$, $\forall\, k \in \mathbf{K}$, $\forall\, x \in \mathbf{V}$ ($\mathcal{T}$ *is homogeneous*),

is called a *linear transformation* (or a *vector space morphism*).

A bijective linear transformation is called an *isomorphism*.

If there exists an isomorphism $\mathcal{T} : \mathbf{V} \longrightarrow \mathbf{W}$, then $\mathbf{V}$ and $\mathbf{W}$ are *isomorphic* and we write $\mathbf{V} \simeq \mathbf{W}$.

REMARKS 5.4 It is straightforward that:

(i) The composite of two linear transformations is a linear transformation.

(ii) If a linear transformation $\mathcal{T}$ is bijective, then its inverse $\mathcal{T}^{-1}$ is a linear transformation.

(iii) "$\simeq$" is an equivalence relation on the class of vector spaces over the same field $\mathbf{K}$ (i.e. $\forall \mathbf{V},\ \mathbf{V} \simeq \mathbf{V}$; $\forall \mathbf{V}, \mathbf{W},\ \mathbf{V} \simeq \mathbf{W} \Longrightarrow \mathbf{W} \simeq \mathbf{V}$; $\forall \mathbf{V}, \mathbf{W}, \mathbf{U},\ \mathbf{U} \simeq \mathbf{W}$ and $\mathbf{W} \simeq \mathbf{V} \Longrightarrow \mathbf{U} \simeq \mathbf{V}$).

(iv) $\mathcal{T}(0) = 0$, for any linear transformation $\mathcal{T}$.

(v) (a) and (b) in Definition 5.3 are equivalent to:

$$\mathcal{T}(kx + ly) = k\mathcal{T}(x) + l\mathcal{T}(y) \ , \quad \forall\, k, l \in \mathbf{K},\ \forall\, x, y \in \mathbf{V} \ ,$$

(i.e. $\mathcal{T}$ is compatible with linear combinations)

PROPOSITION 5.5 *Every vector space* $\mathbf{V}$ *of dimension* n *is isomorphic to the space* $\mathbf{K}^n$.

Proof. Let $\mathbf{B} = \{v_1, \ldots, v_n\}$ be a basis of $\mathbf{V}$. Define $\mathcal{T} : \mathbf{V} \longrightarrow \mathbf{K}^n$ by $\mathcal{T}(x) = (x_1, \ldots, x_n)$, where $x_1, \ldots, x_n$ are the coordinates of v with respect to the basis $\mathbf{B}$. The map $\mathcal{T}$ is well defined and bijective, by Proposition 5.1. It is obvious that $\mathcal{T}$ is linear. QED

The linear map used in the above proof is called the *coordinate system* associated to the basis $\mathbf{B}$.

LEMMA 5.6 *Let* $\mathcal{T} : \mathbf{V} \longrightarrow \mathbf{W}$ *be a linear transformation. Then the following are equivalent:*

(i) $\mathcal{T}$ *is injective.*

(ii) $\mathcal{T}(x) = 0 \Longrightarrow x = 0$.

(iii) *For any linearly independent set* $\mathbf{S}$ *in* $\mathbf{V}$, $\mathcal{T}(\mathbf{S})$ *is linearly independent in* $\mathbf{W}$.

Proof. $(i) \Longrightarrow (ii)$ Assume $\mathcal{T}$ injective. If $\mathcal{T}(x) = 0$, then $\mathcal{T}(x) = \mathcal{T}(0)$ by Remark 5.4 (iv). Now injectivity gives $x = 0$.

$(ii) \Longrightarrow (i)$ Assume (ii) and let $x, y \in \mathbf{V}$, $\mathcal{T}(x) = \mathcal{T}(v)$. Then $\mathcal{T}(x - y) = 0$, so $x - y = 0$ by (ii). Thus $x = y$.

$(ii) \Longrightarrow (iii)$ It suffices to prove (iii) for $\mathbf{S}$ finite (for, see Remark 3.2). Let $\mathbf{S} = (v_1, \ldots, v_p)$ and $k_1\mathcal{T}(v_1) + \ldots + k_p\mathcal{T}(v_p) = 0$. By the linearity of $\mathcal{T}$, $\mathcal{T}(k_1v_1 + \ldots + k_pv_p) = 0$; then $k_1v_1 + \ldots + k_pv_p = 0$, by (ii). $\mathbf{S}$ linearly independent implies $k_1 = \ldots = k_p = 0$, thus $\mathcal{T}(\mathbf{S})$ is linearly independent.

$(iii) \Longrightarrow (ii)$ Note that (ii) is equivalent to: $\mathcal{T}(x) \neq 0$, for all $x \neq 0$. Let $x \neq 0$ and $\mathbf{S} = \{x\}$ in (iii). Then $\{\mathcal{T}(x)\}$ is linearly independent, i.e. $\mathcal{T}(x) \neq 0$. QED

THEOREM 5.7 *Two finite dimensional vector spaces* $\mathbf{V}$ *and* $\mathbf{W}$ *are isomorphic if and only if* $dim\,\mathbf{V} = dim\,\mathbf{W}$.

Proof. Suppose that $\mathbf{V}$ and $\mathbf{W}$ are isomorphic i.e. there exists a bijective linear transformation $\mathcal{T} : \mathbf{V} \longrightarrow \mathbf{W}$. Let $n = dim\,\mathbf{V}$ and choose a basis $\mathbf{B} = (v_1, \ldots, v_n)$ of $\mathbf{V}$. Then $\mathcal{T}(\mathbf{B})$ is linearly independent by the previous lemma and the injectivity of $\mathcal{T}$.

It remains to show that $\mathbf{W} = Span\mathcal{T}(\mathbf{B})$. For, let $y \in \mathbf{W}$. Then, there exists $x \in \mathbf{V}$ such that $\mathcal{T}(x) = y$, since $\mathcal{T}$ is surjective. But x can be written as $x = x_1v_1 + \ldots + x_nv_n$ since $x \in Span\,\mathbf{B}$. Then

$$y = \mathcal{T}(x_1v_1 + \ldots + x_nv_n) = x_1\mathcal{T}(v_1) + \ldots + x_n\mathcal{T}(v_n) \in Span\,\mathcal{T}(\mathbf{B})\ .$$

We proved that $\mathcal{T}(\mathbf{B})$ is a basis of $\mathbf{W}$. Since $card\,\mathcal{T}(\mathbf{B}) = n$, it follows that $dim\mathbf{V} = dim\,\mathbf{W}$.

Conversely, suppose $dim\ \mathbf{V} = dim\,\mathbf{W} = n$. Then $\mathbf{V} \simeq \mathbf{K}^n$ and $\mathbf{W} \simeq \mathbf{K}^n$ imply $\mathbf{V} \simeq \mathbf{W}$ (see Remarks 5.4). QED

We will use quite often the following lemma. (See the Linear Algebra highschool manual for a proof.)

LEMMA 5.8 *Let* $A = [a_{ij}]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in M_{m,n}(\mathbf{K})$. *Then the rank of* A *is equal to the maximal number of linearly independent columns.*

THEOREM 5.9 **(the change of basis)** *Let* $\mathbf{B} = \{v_1, \ldots, v_n\}$ *be a basis of* $\mathbf{V}$ *and* $\mathbf{B}' = \{w_1, \ldots, w_n\}$ *an ordered subset of* $\mathbf{V}$, *with*

$$w_j = \sum_{i=1}^{n} c_{ij} v_i \ , \quad \forall\, j = 1, \ldots, n \ .$$

Then $\mathbf{B}'$ *is another basis of* $\mathbf{V}$ *if and only if* $det\,[c_{ij}] \neq 0$.

Proof. Let $\mathcal{T} : \mathbf{V} \longrightarrow \mathbf{K}^n$ be the coordinate system associated to $\mathbf{B}$. Then $\mathcal{T}(v_i) = e_i$ and $\mathcal{T}(w_j) = \begin{bmatrix} c_{1j} \\ \vdots \\ c_{nj} \end{bmatrix}$. Denote $C = [c_{ij}]_{i,j} \in M_{n,n}(\mathbf{K})$. It follows that $\mathcal{T}(w_j)$ is the j'th column of C. Then:

$det\, C \neq 0$ if and only if $rank\, C = n$;

$rank\, C = n$ if and only if $\mathcal{T}(\mathbf{B}')$ is linearly independent, by Lemma 5.8;

Lemma 5.6 applies for $\mathcal{T}$ and $\mathcal{T}^{-1}$, thus $\mathbf{B}'$ is linearly independent if and only if $\mathcal{T}(\mathbf{B}')$ is linearly independent;

$\mathbf{B}'$ is a basis of $\mathbf{W}$ if and only if $\mathbf{B}'$ is linearly independent, by Corollary 4.9,

and the conclusion follows. QED

DEFINITION 5.10 If $\mathbf{B}'$ in the previous theorem is a basis, then the matrix C is called the *matrix of change of basis* from $\mathbf{B}$ to $\mathbf{B}'$.

Note that in this case the matrix of change from $\mathbf{B}'$ to $\mathbf{B}$ is C^{-1}.

COROLLARY 5.11 **(the coordinate transformation formula)** *Let* $\mathbf{B}$, $\mathbf{B}'$ *be two bases of* $\mathbf{V}$, $dim\, \mathbf{V} = n$, C *the matrix of change from* $\mathbf{B}$ *to* $\mathbf{B}'$, *and* $x \in \mathbf{V}$.

If $X = {}^t(x_1, \ldots, x_n)$ *and* $X' = {}^t(x'_1, \ldots, x'_n)$ *are the coordinates of* x *with respect to* $\mathbf{B}$ *and* $\mathbf{B}'$ *respectively, then* $X = CX'$.

Proof. We write

$$x = \sum_{j=1}^{n} x'_j w_j = \sum_{j=1}^{n} x'_j \left(\sum_{i=1}^{n} c_{ij} v_i \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} c_{ij} x_j \right) v_i \ .$$

By the uniqueness of the representation of x with respect to the basis $\mathbf{B}$, we get $x_i = \sum_{j=1}^{n} c_{ij} x'_j$, $\forall\, i = 1, \ldots, n$. This means $X = CX'$. QED

6 Euclidean Vector Spaces

Unless otherwise specified, $\mathbf{K}$ will denote either one of the fields $\mathbf{R}$ or $\mathbf{C}$. Euclidean vector spaces are real or complex vector spaces with an additional operation that will be used to define the length of a vector, the angle between two vectors and orthogonality in a way which generalizes the usual geometric properties of space vectors in $\mathbf{V}_3$. The Euclidean vector space $\mathbf{V}_3$ will make the object of a separate chapter.

DEFINITION 6.1 Let $\mathbf{V}$ be a real or complex vector space. An *inner* (*scalar*, *dot*) *product* on $\mathbf{V}$ is a map $\langle\ ,\ \rangle : \mathbf{V} \times \mathbf{V} \longrightarrow \mathbf{K}$ which associates to each pair (v, w) a scalar denoted by $\langle v, w\rangle$, satisfying the following properties:

(i) *Linearity in the first variable*:

$$\langle v_1 + v_2, w\rangle = \langle v_1, w\rangle + \langle v_2, w\rangle\ , \quad \forall\, v_1, v_2, w \in \mathbf{W} \qquad \text{(additivity)}$$
$$\langle kv, w\rangle = k\langle v, w\rangle\ , \quad \forall\, k \in \mathbf{K},\ \forall\, v, w \in \mathbf{W} \qquad \text{(homogeneity)}$$

(ii) *Hermitian symmetry* (*symmetry* in the real case):

$$\langle w, v\rangle = \overline{\langle v, w\rangle}\ ,$$

where the bar denotes complex conjugation.

(iii) *Positivity*:

$$\langle v, v\rangle > 0, \quad \forall\, v \neq 0,\ v \in \mathbf{V}\ .$$

REMARKS 6.2 (i) Note that linearity in the first variable means that if w is fixed, then the resulting function of one variable is a linear transformation from $\mathbf{V}$ into $\mathbf{K}$.

(ii) If $\mathbf{K} = \mathbf{R}$, then the second equality in (ii) is equivalent to: $\langle v, kw\rangle = k\langle v, w\rangle$, implying linearity in the second variable too, and (iii) is equivalent to $\langle w, v\rangle = \langle v, w\rangle$.

(iii) By (i) and (ii) in the definition we deduce the *conjugate linearity* (*linearity* in the real case) *in the second variable*:

$$\langle v, w_1 + w_2\rangle = \langle v, w_1\rangle + \langle v, w_2\rangle \qquad \text{(additivity)}$$
$$\langle v, kw\rangle = \overline{k}\langle v, w\rangle\ . \qquad \text{(conjugate homogeneity)}$$

(iv) Additivity implies that $\langle v, 0\rangle = 0 = \langle 0, w\rangle$, $\forall\, v, w \in \mathbf{V}$. In particular $\langle 0, 0\rangle = 0$. Combining this with positivity, it follows that:

$$\langle v, v\rangle \geq 0\ , \quad \forall\, v \in \mathbf{V} \quad \text{and} \quad \langle v, v\rangle = 0 \text{ if and only if } v = 0\ .$$

(v) $\langle v, w\rangle = 0\ , \quad \forall\, w \in \mathbf{V} \Longrightarrow v = 0$ and

$\langle v, w\rangle = 0\ , \quad \forall\, v \in \mathbf{V} \Longrightarrow w = 0$, since $\langle v, v\rangle = 0 \Longrightarrow v = 0$.

(vi) Hermitian symmetry implies $\langle v, v\rangle \in \mathbf{R}\ , \quad \forall\, v \in \mathbf{V}$ (without imposing positivity).

DEFINITION 6.3 A real or a complex vector space endowed with a scalar product is called a *Euclidean vector space.*

Examples of canonical Euclidean vector spaces

1) $\mathbf{R}^n$ with $\langle x, y\rangle = x_1y_1 + \ldots + x_ny_n$, where $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$.

2) $\mathbf{C}^n$ with $\langle x, y\rangle = x_1\overline{y_1} + \ldots + x_n\overline{y_n}$, where $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$.

3) $\mathbf{V}_3$ with $\langle \bar{a}, \bar{b}\rangle = \|\bar{a}\| \cdot \|\bar{b}\| \cos\theta$, where $\|\bar{a}\|$, $\|\bar{b}\|$ are the lengths of $\bar{a}$ and $\bar{b}$ respectively, and θ is the angle between $\bar{a}$ and $\bar{b}$.

4) $\mathbf{V} = \{f \mid f : [a, b] \longrightarrow \mathbf{R}, f \text{ continuous}\}$ is a real Euclidean vector space with the scalar product given by

$$\langle f, g\rangle = \int_a^b f(t)g(t)\, dt \ .$$

5) $\mathbf{V} = \{f \mid f : [a, b] \longrightarrow \mathbf{C}, f \text{ continuous}\}$ is a complex Euclidean vector space with the scalar product given by

$$\langle f, g\rangle = \int_a^b f(t)\overline{g(t)}\, dt \ .$$

THEOREM 6.4 (**the Cauchy–Schwarz inequality**) *Let* $\mathbf{V}$ *be a Euclidean vector space. Then*

(6.1) $$| \langle v, w\rangle |^2 \leq \langle v, v\rangle\langle w, w\rangle \ , \quad \forall\, v, w \in \mathbf{V} \ .$$

Equality holds if and only if v, w *are linearly dependent* (*collinear*).

Proof. The case $w = 0$ is obvious. Assume $w \neq 0$. By positivity

(6.2) $$\langle v + \alpha w, v + \alpha w\rangle \geq 0 \ , \quad \forall\, \alpha \in \mathbf{K} \ .$$

Take $\alpha = -\dfrac{\langle v, w\rangle}{\langle w, w\rangle}$ and expand to obtain (6.1).

If equality holds in (6.1), then equality holds in (6.2) for $\alpha = -\dfrac{\langle v, w\rangle}{\langle w, w\rangle}$. Then $v - \dfrac{\langle v, w\rangle}{\langle w, w\rangle} w = 0$ by positivity. Thus v, w are linearly dependent.

Conversely, suppose $v = \lambda w$. Then

$$\begin{aligned} | \langle v, w\rangle |^2 &= | \langle v, \lambda v\rangle |^2 = | \overline{\lambda}\langle v, v\rangle |^2 = | \overline{\lambda} |^2 \ \langle v, v\rangle^2 = \lambda\overline{\lambda}\langle v, v\rangle\langle v, v\rangle \\ &= \langle v, v\rangle\langle \lambda v, \lambda v\rangle = \langle v, v\rangle\langle w, w\rangle \ . \qquad \text{QED} \end{aligned}$$

DEFINITION 6.5 Let $\mathbf{V}$ be a real or complex vector space.

The function $\| \ \| : \mathbf{V} \longrightarrow \mathbf{R}$ is called a *norm* on $\mathbf{V}$ if it satisfies:

(*i*) $\|v\| \geq 0$, $\quad \forall\, v \in \mathbf{V}$ and $\|v\| = 0$ if and only if $v = 0$ (positivity) (*ii*) $\|kv\| = \mid k \mid \|v\|$, $\forall k \in \mathbf{K},\ \forall v \in \mathbf{V}$

(*iii*) $\|u+v\| \leq \|u\| + \|v\|$, $\forall u, v \in \mathbf{V}$ (the triangle inequality).

A vector space endowed with a norm is called a *normed vector space.*

PROPOSITION 6.6 *Let* $\mathbf{V}$ *be a Euclidean vector space. Then the function defined by* $\|v\| = \sqrt{\langle v, v\rangle}$ *is a norm on* $\mathbf{V}$.

Proof. Properties (*i*) and (*ii*) in Definition 6.5 are straightforward from the positivity and (conjugate) linearity of the inner product. For (*iii*)

$$\begin{aligned} \|u+v\|^2 &\leq \langle u+v, u+v\rangle = \langle u,u\rangle + 2\mathrm{Re}\langle u,v\rangle + \langle v,v\rangle \\ &\leq \|u\|^2 + \|v\|^2 + 2 \mid \langle u,v\rangle \mid , \qquad \text{since } \mathrm{Re} z \leq \mid z \mid \text{ for any } z \in \mathbf{C} \\ &\leq \|u\|^2 + \|v\|^2 + 2\|u\| \cdot \|v\| , \qquad \text{by Cauchy's inequality} \\ &\leq (\|u\| + \|v\|)^2 . \qquad \text{QED} \end{aligned}$$

The norm defined by an inner product as in the previous proposition is called a *Euclidean norm.*

REMARK 6.7 Let $(\mathbf{V}, \| \ \|)$ be a (real or complex) vector space. If $\| \ \|$ is a Euclidean norm, then it is easy to see that:

$$\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2) , \quad \forall u, v \in \mathbf{V} \qquad \text{(the parallelogram law)} .$$

Note that not all norms satisfy the parallelogram law. For example, let $\| \ \|_\infty : \mathbf{R}^n \longrightarrow [0, \infty)$, $n \geq 2$ defined by $\|x\|_\infty = \max\{\mid x_1 \mid, \ldots, \mid x_n \mid\}$ and take $u = (1, 0, \ldots, 0)$, $v = (1, 1, 0, \ldots, 0)$.

DEFINITION 6.8 A vector $u \in \mathbf{V}$ with $\|u\| = 1$ is called a *unit vector* or *versor*.

Any vector $v \in \mathbf{V} \setminus \{0\}$ can be written as $v = \|v\|u$, where u is a unit vector. The vector $u = \dfrac{1}{\|v\|} v$ is called the *unit vector in the direction of* v.

If $v, w \in \mathbf{V} \setminus \{0\}$ and $\mathbf{K} = \mathbf{R}$, then the Cauchy inequality is equivalent to

$$-1 \leq \frac{\langle v, w\rangle}{\|v\| \cdot \|w\|} \leq 1 . \tag{6.3}$$

This double inequality allows the following definition.

DEFINITION 6.9 Let $\mathbf{V}$ be a real Euclidean vector space and $v, w \in \mathbf{V} \setminus \{0\}$. Then the number $\theta \in [0, \pi]$ defined by

$$\cos\theta = \frac{\langle v, w\rangle}{\|v\| \cdot \|w\|}$$

is called the *angle* between v and w.

THEOREM 6.10 *Let* $\mathbf{V}$ *be a real or complex normed vector space. The function* $d : \mathbf{V} \times \mathbf{V} \longrightarrow \mathbf{R}$*, defined by* $d(v, w) = \|v - w\|$ *is a distance (metric) on* $\mathbf{V}$*, i.e. it satisfies the following properties:*

(*i*) $d(v, w) \geq 0$, $\forall\, v, w \in \mathbf{V}$ *and* $d(v, w) = 0$ *if and only if* $v = w$.

(*ii*) $d(v, w) = d(w, v)$, $\forall\, v, w \in \mathbf{V}$

(*iii*) $d(u, v) \leq d(u, w) + d(w, v)$, $\forall\, u, v, w \in \mathbf{V}$.

The proof is straightforward from the properties of the norm.

A vector space endowed with a *distance* (i.e. a map satisfying (*i*), (*ii*), (*iii*) in the previous theorem) is called a *metric space*. If the distance is defined by a Euclidean norm, then it is called a *Euclidean distance*.

7 Orthogonality

Let $\mathbf{V}$ be a Euclidean vector space. In the last section, we defined the angle between two nonzero vectors. The definition of orthogonality will be compatible with the definition of the angle.

DEFINITION 7.1 (*i*) Two vectors $v, w \in \mathbf{V}$ are called *orthogonal* (or *perpendicular*) if $\langle v, w\rangle = 0$. We write $v \perp w$ when v and w are orthogonal.

(*ii*) A subset $\mathbf{S} \neq \emptyset$ is called *orthogonal* if its vectors are mutually orthogonal, i.e. $\langle v, w\rangle = 0$, $\forall\, v, w \in \mathbf{S}$, $v \neq w$.

(*iii*) A subset $\mathbf{S} \neq \emptyset$ is called *orthonormal* if it is orthogonal and each vector of $\mathbf{S}$ is a unit vector (i.e. $\forall\, v \in \mathbf{S},\ \|v\| = 1$).

PROPOSITION 7.2 *Any orthogonal set of nonzero vectors is linearly independent.*

Proof. Let $\mathbf{S}$ be orthogonal. In order to show that $\mathbf{S}$ is linearly independent, we may assume $\mathbf{S}$ finite (see Remark 3.2 (*iii*)), $\mathbf{S} = \{v_1, \ldots, v_p\}$, where $v_1, \ldots, v_p$ are distinct. Let $k_1, \ldots, k_p \in \mathbf{K}$ such that $k_1v_1 + \ldots + k_pv_p = 0$. Right multiplication by v_j in the sense of the inner product yields

$$k_1\langle v_1, v_j\rangle + \ldots + k_p\langle v_p, v_j\rangle = \langle 0, v_j\rangle = 0\ , \quad \forall\, j = 1, \ldots, p\ .$$

By the orthogonality of $\mathbf{S}$, $\langle v_i, v_j\rangle = 0$ for $i \neq j$, thus

$$k_j\langle v_j, v_j\rangle = 0\ , \quad \forall\, j = 1, \ldots, p\ .$$

But $v_j \neq 0$, since $0 \notin \mathbf{S}$, so $\langle v_j, v_j\rangle \neq 0$ and it follows that $k_j = 0$, $\forall\, j = 1, \ldots, p$, which shows that $\mathbf{S}$ is linearly independent. QED

Combining the above result and Corollary 4.8 (*iii*) we obtain the following result.

COROLLARY 7.3 *If* $\dim \mathbf{V} = n$, $n \geq 1$, *then any orthogonal set of* n *nonzero vectors is a basis of* $\mathbf{V}$.

THEOREM 7.4 (the Gram–Schmidt procedure) *If* $dim\,\mathbf{V} = n$, $n \geq 2$ *and* $\mathbf{B} = \{v_1, \ldots, v_n\}$ *is a basis of* $\mathbf{V}$*, then there exists an orthonormal basis* $\mathbf{B}' = \{u_1, \ldots, u_n\}$ *of* $\mathbf{V}$*, such that*

$$Span\,\{v_1, \ldots, v_m\} = Span\,\{u_1, \ldots, u_m\}\ , \quad \forall\, m = 1, \ldots, n\ .$$

Proof. Consider the set $\mathbf{B}'' = \{w_1, \ldots, w_n\}$, whose elements are defined by

$$\begin{aligned}
&w_1 = v_1\\
&w_2 = v_2 + k_{12}w_1\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&w_m = v_m + k_{1m}w_1 + \ldots + k_{m-1,m}w_{m-1}\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&w_n = v_n + k_{1n}w_1 + \ldots + k_{n-1,n}w_{n-1}\ , \quad k_{ij} \in \mathbf{K},\ i, j \in \{1, \ldots, n\},\ i < j\ .
\end{aligned}$$

It is easy to prove by induction that $Span\,\{v_1, \ldots, v_m\} = Span\,\{w_1, \ldots, w_m\}$ and $w_m \neq 0$, for all $m = 1, \ldots, n$. This is trivial for $m = 1$. Let $2 \leq m \leq n$ and assume

$$Span\,\{v_1, \ldots, v_{m-1}\} = Span\,\{w_1, \ldots, w_{m-1}\}\ .$$

Then

$$\begin{aligned}
Span\{v_1, \ldots, v_m\} &= Span\{v_1, \ldots, v_{m-1}\} + Span\{v_m\}\\
&= Span\{w_1, \ldots, w_{m-1}\} + Span\{v_m\} \qquad \text{(by the induction hypothesis)}\\
&= Span\{w_1, \ldots, w_{m-1}, v_m\}\\
&= Span\{w_1, \ldots, w_{m-1}, w_m\} \qquad \text{(by the definition of } w_m)\ .
\end{aligned}$$

Now suppose $w_m = 0$ for some $m \geq 2$. Then $v_m \in Span\{w_1, \ldots, w_{m-1}\} = Span\{v_1, \ldots, v_{m-1}\}$ implies $\{v_1, \ldots, v_m\}$ linearly dependent, by Prop. 3.3. This yields a contradiction.

The scalars k_{ij} can be chosen such that $\mathbf{B}''$ will become an orthogonal set as follows.

$$\langle w_2, w_1\rangle = 0 \Longleftrightarrow \langle v_2, w_1\rangle + k_{12}\langle w_1, w_1\rangle = 0 \Longleftrightarrow k_{12} = -\frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle}\ .$$

Assume that k_{ij} are scalars such that $\{w_1, \ldots, w_{m-1}\}$ is orthogonal. Then, using the fact that $\langle w_j, w_i\rangle = 0$, $\forall\, i, j \in \{1, \ldots, m-1\}$, $i \neq j$, we get for all $i = 1, \ldots, m-1$:

$$\langle w_m, w_i\rangle = 0 \Longleftrightarrow \langle v_m, w_i\rangle + k_{im}\langle w_i, w_i\rangle = 0 \Longleftrightarrow k_{im} = -\frac{\langle v_m, w_i\rangle}{\langle w_i, w_i\rangle}\ .$$

Note that $\langle w_i, w_i\rangle \neq 0$ since $w_i \neq 0$.

We showed that $\mathbf{B}''$ is an orthogonal set of nonzero vectors, hence an orthogonal basis, where:

$$\begin{aligned}
&w_1 = v_1\\
&w_2 = v_2 - \frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle}\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&w_m = v_m - \frac{\langle v_m, w_1\rangle}{\langle w_1, w_1\rangle}\,w_1 - \ldots - \frac{\langle v_m, w_{m-1}\rangle}{\langle w_{m-1}, w_{m-1}\rangle}\,w_{m-1}\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&w_n = v_n - \frac{\langle v_n, w_1\rangle}{\langle w_1, w_1\rangle}\,w_1 - \ldots - \frac{\langle v_n, w_{n-1}\rangle}{\langle w_{n-1}, w_{n-1}\rangle}\,w_{n-1}\ ,
\end{aligned}$$

and $Span\{v_1, \ldots, v_m\} = Span\{w_1, \ldots, w_m\}$, $\forall\, m = 1, \ldots, n$.

Now set $u_j = \dfrac{1}{\|w_j\|} w_j$, $j = 1, \ldots, n$. The set $\mathbf{B}' = \{u_1, \ldots, u_n\}$ fulfils all the required properties. QED

COROLLARY 7.5 *Any finite dimensional Euclidean space has an orthonormal basis.*

THEOREM 7.6 *If* $\dim \mathbf{V} = n$ *and* $\mathbf{B} = \{u_1, \ldots, u_n\}$ *is an orthogonal basis, then the representation of* $x \in \mathbf{V}$ *with respect to the basis* $\mathbf{B}$ *is*

$$x = \frac{\langle x, u_1\rangle}{\langle u_1, u_1\rangle} u_1 + \ldots + \frac{\langle x, u_n\rangle}{\langle u_n, u_n\rangle} u_n \ .$$

In particular, if $\mathbf{B}$ *is an orthonormal basis, then*

$$x = \langle x, u_1\rangle u_1 + \ldots + \langle x, u_n\rangle u_n \ .$$

Proof. The vector x has an unique representation with respect to the basis $\mathbf{B}$:

$$x = \sum_{i=1}^{n} x_i u_i \ . \qquad \text{(see (5.1))}$$

We multiply this equality by each u_j, $j = 1, \ldots, n$ in the sense of the inner product. It follows that

$$\langle x, u_j\rangle = \sum_{i=1}^{n} x_i \langle u_i, u_j\rangle = x_j \langle u_j, u_j\rangle \ , \quad \forall\, j = 1, \ldots, n \ ,$$

thus

$$x_j = \frac{\langle x, u_j\rangle}{\langle u_j, u_j\rangle} \ , \quad \forall\, j = 1, \ldots, n \ .$$

If $\mathbf{B}$ is an orthonormal basis, then $\langle u_j, u_j\rangle = 1$, so $x_j = \langle x, u_j\rangle$. QED

DEFINITION 7.7 Let $\emptyset \neq \mathbf{S} \subset \mathbf{V}$.

A vector $v \in \mathbf{V}$ is *orthogonal* to $\mathbf{S}$ if $v \perp w$, $\forall\, w \in \mathbf{S}$.

The set of all vectors orthogonal to $\mathbf{S}$ is denoted by $\mathbf{S}^\perp$ and is called $\mathbf{S}$–*orthogonal*.

It is easy to see that $\mathbf{S}^\perp$ is a vector subspace of $\mathbf{V}$.

If $\mathbf{W}$ is a vector subspace of $\mathbf{V}$, then $\mathbf{W}^\perp$ is called the *orthogonal complement* of $\mathbf{W}$.

Examples

(*i*) $\{0\}^\perp = \mathbf{V}$, since $\langle v, 0\rangle = 0$, $\forall\, v \in \mathbf{V}$.

(*ii*) $\mathbf{V}^\perp = \{0\}$, since

$$\langle v, w\rangle = 0 \ , \quad \forall\, w \in \mathbf{V} \Longrightarrow \langle v, v\rangle = 0 \Longrightarrow v = 0 \ .$$

(*iii*) If $\mathbf{S} = \{(1,1,0), (1,0,1)\} \subset \mathbf{R}^3$, then

$$\mathbf{S}^\perp = \{(x,y,z) \in \mathbf{R}^3 \, ; \, x + y = 0, \ x + z = 0\} \ .$$

REMARKS 7.8 (i) $\mathbf{S}^\perp = (Span\,\mathbf{S})^\perp$.

(ii) If $\mathbf{W}$ is a vector subspace, then $\mathbf{W}\cap\mathbf{W}^\perp = \{0\}$. (For, let $w \in \mathbf{W}\cap\mathbf{W}^\perp$. It follows that $\langle w, w\rangle = 0$, thus $w = 0$.)

THEOREM 7.9 *Let* $\mathbf{W}$ *be a vector subspace of a Euclidean vector space* $\mathbf{V}$*,* $dim\,\mathbf{W} = n$*,* $n \in \mathbf{N}^*$*. Then* $\mathbf{V} = \mathbf{W}\oplus\mathbf{W}^\perp$*. Moreover, if* $v = w + w^\perp$*,* $w \in \mathbf{W}$*,* $w^\perp \in \mathbf{W}^\perp$*, then the Pythagorean theorem holds, i.e.* $\|v\|^2 = \|w\|^2 + \|w^\perp\|^2$*.*

Proof. Let $\mathbf{B} = \{u_1, \ldots, u_n\}$ be an orthonormal basis of $\mathbf{W}$ and $v \in \mathbf{V}$. Take $w = \langle v, u_1\rangle u_1 + \ldots + \langle v, u_n\rangle u_n$, and $w^\perp = v - w$. Obviously $w \in \mathbf{W}$ and for any $j = 1, \ldots, n$ we get

$$\langle w^\perp, u_j\rangle = \langle v - w, u_j\rangle = \langle v, u_j\rangle - \sum_{i=1}^{n}\langle v, u_i\rangle\langle u_i, u_j\rangle = \langle v, u_j\rangle - \langle v, u_j\rangle = 0\ .$$

We showed that $w^\perp \in \mathbf{B}^\perp = \mathbf{W}^\perp$, therefore $\mathbf{V} = \mathbf{W} + \mathbf{W}^\perp$. The sum is direct since $\mathbf{W}\cap\mathbf{W}^\perp = \{0\}$. It follows that the decomposition

$$v = w + w^\perp\ ,\quad w \in \mathbf{W},\ w^\perp \in \mathbf{W}^\perp$$

is unique. Also

$$\|v\|^2 = \langle v, v\rangle = \langle w, w\rangle + \langle w, w^\perp\rangle + \langle w^\perp, w\rangle + \langle w^\perp, w^\perp\rangle = \|w\|^2 + \|w^\perp\|^2\ ,$$

since $\langle w, w^\perp\rangle = \langle w^\perp, w\rangle = 0$. QED

The vector $w \in \mathbf{W}$ in the decomposition of v with respect to $\mathbf{V} = \mathbf{W}\oplus\mathbf{W}^\perp$ is called the *orthogonal projection of* v *onto* $\mathbf{W}$.

PROPOSITION 7.10 *Let* $\mathbf{V}$ *be a Euclidean vector space,* $dim\,\mathbf{V} = n$*,* $n \in \mathbf{N}^*$*,* $\mathbf{B} = \{u_1, \ldots, u_n\}$ *an orthonormal basis and* $x = \sum_{j=1}^{n} x_j u_j$*,* $y = \sum_{j=1}^{n} y_j u_j \in \mathbf{V}$*. If* $\mathbf{V}$ *is a real vector space, then* $\langle x, y\rangle = \sum_{j=1}^{n} x_j y_j$ *and* $\|x\|^2 = \sum_{j=1}^{n} x_j^2$*.*

If $\mathbf{V}$ *is a complex vector space, then* $\langle x, y\rangle = \sum_{j=1}^{n} x_j\overline{y_j}$ *and* $\|x\|^2 = \sum_{j=1}^{n} \mid x_j \mid^2$*.*

Proof. Assume $\mathbf{V}$ is a complex vector space. Then

$$\begin{aligned}\langle x, y\rangle &= \left\langle \sum_{i=1}^{n} x_i u_i, \sum_{j=1}^{n} y_j u_j \right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n}\langle x_i u_i, y_j u_j\rangle \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\overline{y_j}\langle u_i, u_j\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\overline{y_j}\delta_{ij} = \sum_{j=1}^{n} x_j\overline{y_j}\ .\end{aligned}$$

The real case is similar. QED

We conclude this section with a generalization of Theorem 7.4.

THEOREM 7.11 (Gram–Schmidt– infinite dimensional case) *If* $\mathbf{V}$ *is infinite dimensional and* $L = \{v_1, \ldots, v_k, \ldots\} \subset \mathbf{V}$ *is a countable, infinite, linearly independent set of distinct elements, then there exists an orthonormal set* $L' = \{u_1, \ldots, u_k, \ldots\}$ *such that*

$$Span\{v_1, \ldots, v_k\} = Span\{u_1, \ldots, u_k\}\ , \quad \forall\, k \in \mathbf{N}^*.$$

The proof is based on the construction used in Theorem 7.4, i.e.

$$\begin{aligned}
&w_1 = v_1\\
&w_2 = v_2 - \frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle}\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&w_k = v_k - \frac{\langle v_k, w_1\rangle}{\langle w_1, w_1\rangle}\, w_1 - \ldots - \frac{\langle v_k, w_{k-1}\rangle}{\langle w_{k-1}, w_{k-1}\rangle}\, w_{k-1}, \qquad \forall k \geq 2,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned}$$

then $u_j = \dfrac{1}{\|w_j\|}\, w_j$, $\forall k \in \mathbf{N}^*$.

8 Problems

1. Let $\mathbf{V}$ be a vector space over the field $\mathbf{K}$ and S a nonempty set. We define $F = \{f | f : S \to \mathbf{V}\}$,

$$(f + g)(x) = f(x) + g(x), \quad \text{for all} \quad f, g \in F,$$

$$(tf)(x) = tf(x), \quad \text{for all} \quad t \in \mathbf{K},\ f \in F.$$

Show that F is a vector space over $\mathbf{K}$.

2. Let F be the real vector space of the functions $f : \mathbf{R} \to \mathbf{R}$. Show that

$$\cos t, \cos^2 t, \ldots, \cos^n t, \ldots$$

is a linearly independent system in F.

3. Let $\mathbf{V}$ be the real vector space of functions obtained as the composite of a polynomial of degree at most three and the cosine function (polynomials in $\cos x$, of degree at most three). Write down the transformation of coordinates corresponding to the change of basis from $\{1, \cos x, \cos^2 x, \cos^3 x\}$ to the basis $\{1, \cos x, \cos 2x, \cos 3x\}$, and find the inverse of this transformation. Generalization.

4. Let $\mathbf{V}$ be the real vector space of the real sequences $\{x_n\}$ with the property that the series $\sum x_n^2$ is convergent. Let $x = \{x_n\}$, $y = \{y_n\}$ be two elements of $\mathbf{V}$.
1) Show that the series $\sum x_n y_n$ is absolutely convergent.
2) Prove that the function defined by $\langle x, y\rangle = \sum x_n y_n$ is a scalar product on $\mathbf{V}$.

5. Let $\mathbf{V} = \{x |\ x \in \mathbf{R},\ x > 0\}$. For any $x, y \in \mathbf{V}$, and any $s \in \mathbf{R}$, we define

$$x \oplus y = xy, \quad s \odot x = x^s.$$

Show that $(\mathbf{V}, \oplus, \odot)$ is a real vector space.

6. Is the set $\mathbf{K}_n[X]$ of all polynomials of degree at most n, a vector space over $\mathbf{K}$? What about the set of the polynomials of degree at least n?

7. Show that the set of all convergent sequences of real (complex) numbers is a vector space over $\mathbf{R}$ ($\mathbf{C}$) with respect to the usual addition of sequences and multiplication of a sequence by a number.

8. Prove that the following sets are real vector spaces with respect to the usual addition of functions and multiplication of a function by a real number.
1) $\{f|\, f : I \to \mathbf{R}, \quad I = \text{interval} \subseteq \mathbf{R},\ f \text{ differentiable}\}$
2) $\{f|\, f : I \to \mathbf{R}, \quad I = \text{interval} \subseteq \mathbf{R},\ f \text{ admits antiderivatives}\}$
3) $\{f|\, f : [a, b] \to \mathbf{R}, \quad f \text{ integrable}\}$.

9. Which of the following pairs of operations define a real vector space structure on $\mathbf{R}^2$?
1) $(x_1, x_2) + (y_1, y_2) = (x_1 + x_2, x_2y_2)$, $k(x_1, x_2) = (kx_1, kx_2)$, $\quad k \in \mathbf{R}$
2) $(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$, $k(x_1, x_2) = (x_1, kx_2)$, $\quad k \in \mathbf{R}$
3) $(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$, $k(x_1, x_2) = (kx_1, kx_2)$, $\quad k \in \mathbf{R}$.

10. Let P_n be the real vector space of real polynomial functions of degree at most n. Study which of the following subsets are vector subspaces, then determine the sum and the intersection of the vector subspaces you found.

$$A = \{p|\, p(0) = 0\},\ B = \{p|\, p(0) = 1\},\ C = \{p|\, p(-1) + p(1) = 0\}.$$

11. Study the linear dependence of the following sets:
1) $\{1, 1, 1), (0, -3, 1), (1, -2, 2)\} \subset \mathbf{R}^3$,

2) $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} \right\} \subset \mathcal{M}_{\in\times\in}(\mathbf{C})$,

3) $\{1, x, x^2\}, \{e^x, xe^x, x^2e^x\}, \{e^x, e^{-x}, \sinh x\}, \{1, \cos^2 x, \cos 2x\} \subset C^\infty(\mathbf{R}) =$ the real vector space of C^∞ functions on $\mathbf{R}$.

12. Show that the solution set of a linear homogeneous system with n unknowns (and coefficients in $\mathbf{K}$) is a vector subspace of $\mathbf{K}^n$. Determine its dimension.

13. Consider $\mathbf{V} = \mathbf{K}^n$. Prove that every subspace $\mathbf{W}$ of $\mathbf{V}$ is the solution set of some linear homogeneous system with n unknowns.

14. A straight line in $\mathbf{R}^2$ is identified to the solution set of a (nontrivial) linear equation with two unknowns. Similarly, a plane in $\mathbf{R}^3$ is identified to the solution set of a linear equation with three unknowns; a straight line in $\mathbf{R}^3$ can be viewed as the intersection of two planes, so it may be identified to the solution set of a linear system of rank two, with three unknowns.
(a) Prove that the only proper subspaces of $\mathbf{R}^2$ are the straight lines passing through the origin.
(b) Prove that the only proper subspaces of $\mathbf{R}^3$ are the planes passing through the origin, and the straight lines passing through the origin.

15. Which of the following subsets of $\mathbf{R}^3$ are vector subspaces?

$$D_1 : \frac{x}{1} = \frac{y}{2} = \frac{z}{-1},\ D_2 : \frac{x-1}{-1} = \frac{y}{1} = \frac{z-2}{-1}$$

$$P_1 : x + y + z = 0,\ P_2 : x - y - z = 1.$$

Determine the intersection and the sum of the vector spaces found above.

16. In $\mathbf{R}^3$, we consider the vector subspaces $P : x+2y-z = 0$ and $Q : 2x-y+2z = 0$. For each of them, determine a basis and a supplementary vector subspace. Find a basis of the sum $P + Q$ and a basis of the intersection $P \cap Q$.

17. Let S be a subset of $\mathbf{R}^3$ made of the vectors
1) $v_1 = (1, 0, 1), v_2 = (-1, 0, -1), v_3 = (3, 0, 3)$
2) $v_1 = (-1, 1, 1), v_2 = (1, -1, 1), v_3 = (0, 0, 1), v_4 = (1, -1, 2)$.
In each case, determine the dimension of the subspace spanned by S and point out a basis contained in S. What are the Cartesian equations of the subspace ?

18. In each case write down the coordinate transformation formulas corresponding to the change of basis.
1) $e_1 = (-1, 2, 1), e_2 = (1, -2, 1), e_3 = (0, 1, 1)$ in $\mathbf{R}^3$
$e'_1 = (1, -1, 1), e'_2 = (0, 1, -1), e'_3 = (1, 1, 0)$.
2) $e_1 = 1, e_2 = t, e_3 = t^2, e_4 = t^3$ in P_3
$e'_1 = 1 - t, e'_2 = 1 + t^2, e'_3 = t^2 - t, e'_4 = t^3 + t^2$.

19. Let $\mathbf{V}$ be a real vector space. Show that $\mathbf{V} \times \mathbf{V}$ is a complex vector space with respect to the operations

$$(u, v) + (x, y) = (u + x, v + y)$$
$$(a + ib)(u, v) = (au - bv, bu + av).$$

This complex vector space is called the *complexification* of $\mathbf{V}$ and is denoted by ${}^C\mathbf{V}$. Show that:
1) ${}^C\mathbf{R}^n = \mathbf{C}^n$;
2) if S is a linearly independent set in $\mathbf{V}$, then $S \times \{0\}$ is a linearly independent set in ${}^C\mathbf{V}$;
3) if $\{e_1, \ldots, e_n\}$ is a basis of $\mathbf{V}$, then $\{(e_1, 0), \ldots, (e_n, 0)\}$ is a basis of ${}^C\mathbf{V}$, and $dim_C {}^C\mathbf{V} = dim_R \mathbf{V}$.

20. Let $\mathbf{V}$ be a complex vector space. Consider the set $\mathbf{V}$ with the same additive group structure, but scalar multiplication restricted to multiplication by real numbers. Prove that the set $\mathbf{V}$ becomes a real vector space in this way.
Denote this real vector space by ${}^R\mathbf{V}$. Show that
1) ${}^R\mathbf{C}^n = \mathbf{R}^{2n}$,
2) if $dim\mathbf{V} = n$, then $dim{}^R\mathbf{V} = 2n$.

21. Explain why the following maps are not scalar products:
1) $\varphi : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R},\ \varphi(x, y) = \sum\limits_{i=1}^{n} |x_i y_i|$
3) $\varphi : C^0([0, 1]) \times C^0([0, 1]) \to \mathbf{R},\ \varphi(f, g) = f(0)g(0)$.

22. In the canonical Euclidean space $\mathbf{R}^3$, consider the vector $v = (2, 1, -1)$ and the vector subspace $P : x - y + 2z = 0$. Find the orthogonal projection w of v on P and the vector $w^\perp$.

23. Let $\mathbf{R}^4$ be the canonical Euclidean space of dimension 4. Find an orthonormal basis for the subspace generated by the vectors
$u = (0, 1, 1, 0), v = (1, 0, 0, 1), w = (1, 1, 1, 1)$.

Chapter 2

Linear Transformations

1 General Properties

Throughout this section **V** and **W** will be vector spaces over the same field **K**.

We used linear linear transformations, in particular the notion of isomorphism for the the identification of an n–dimensional vector space with $\mathbf{K}^n$ (see Prop. 5.5, Chap.1). In this chapter we will study linear transformations in more detail.

Recall (Def. 5.3) that a linear transformation $\mathcal{T}$ from **V** into **W** is a map $\mathcal{T} : \mathbf{V} \to \mathbf{W}$ satisfying

(1.1) $$\mathcal{T}(x+y) = \mathcal{T}(x) + \mathcal{T}(y)\ , \quad \forall\, x, y \in \mathbf{V} \qquad \text{(additivity)}$$

(1.2) $$\mathcal{T}(kx) = k\mathcal{T}(x)\ , \quad \forall\, k \in \mathbf{K},\ \forall\, x \in \mathbf{V} \qquad \text{(homogeneity).}$$

The definition is equivalent to

(1.3) $$\mathcal{T}(kx+ly) = k\mathcal{T}(x) + l\mathcal{T}(y)\ , \quad \forall\, k, l \in \mathbf{K},\ \forall\, x, y \in \mathbf{V}, \qquad \text{(linearity)}$$

as we pointed out in Chap. 1, Remarks 5.4 (v).

We will write sometimes $\mathcal{T}x$ instead of $\mathcal{T}(x)$.

We are actually quite familiar with some examples of linear transformations. One of them is the following, which is in fact the main example.

MAIN EXAMPLE: left multiplication by a matrix

Let $A \in M_{m,n}(\mathbf{K})$ and consider A as an operator on column vectors. It defines a linear transformation $\mathcal{A} : \mathbf{K}^n \longrightarrow \mathbf{K}^m$ by

$$X \longrightarrow AX\ , \quad \text{where}\ \ X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\ ;$$
$$\mathcal{A}(x) = {}^t(AX)\ , \quad x = (x_1, \ldots, x_n) \in \mathbf{K}^n\ , \quad X = {}^t x\ .$$

Obviously, $\mathcal{A}$ is linear by the known properties of the matrix multiplication.

Note that each component of the vector $\mathcal{A}(x)$ is a linear combination of the components of x.

Consider the particular case $A = \begin{bmatrix} 2 & 1 \\ 3 & 0 \\ -1 & 5 \end{bmatrix} \in M_{3,2}(\mathbf{R})$; then

$$A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 + x_2 \\ 3x_1 \\ -x_1 + 5x_2 \end{bmatrix}$$

and

$$\mathcal{A}(x) = (2x_1 + x_2,\ 3x_1,\ -x_1 + 5x_2)\ .$$

Conversely, any map $\mathcal{A} : \mathbf{K}^n \longrightarrow \mathbf{K}^m$ with the property that each component of $\mathcal{A}(x)$ is a linear combination with constant coefficients of the components of x, is given by a matrix $A \in M_{m,n}(\mathbf{K}$ as above; A is the coefficient matrix of ${}^t\,(\mathcal{A}(x))$, i.e. if for any $x \in \mathbf{K}^n$, the i'th component of $\mathcal{A}(x)$ is $k_1x_1 + \ldots + k_nx_n$, then the i'th row of $\mathcal{A}$ is $(k_1, \ldots, k_n)$.

If $m = n = 1$, then $\mathcal{A} = [a]$ for some $a \in \mathbf{K}$, and $\mathcal{A} : \mathbf{K} \longrightarrow \mathbf{K}$, $\mathcal{A}(x) = ax$.

Linear transformations are also called *vector space homomorphisms* (or *morphisms*), or *linear operators*, or just *linear maps.* Their compatibility with the operations can be regarded as a kind of "transport" of the algebraic structure of $\mathbf{V}$ to $\mathbf{W}$.

Note that (1.1) says that $\mathcal{T}$ is a homomorphism of additive groups.

A linear map $\mathcal{F} : \mathbf{V} \longrightarrow \mathbf{K}$ (i.e. $\mathbf{W} = \mathbf{K}$) is also called a *linear form* on $\mathbf{V}$.

More examples of linear transformations

(*i*) $\mathbf{V} = \mathcal{P}_n$ = the vector space of real polynomial functions of degree $\leq n$, $\mathbf{W} = \mathcal{P}_{n-1}$ and $\mathcal{T}(p) = p'$.

(*ii*) $\mathbf{V} = C^1(a,b)$, $\mathbf{W} = C^0(a,b)$, $\mathcal{T}(f) = f'$.

(*iii*) $\mathbf{V} = C^0[a,b]$, $\mathbf{W} = \mathbf{R}$, $\mathcal{T}(f) = \int_a^b f(t)\ dt$.

(*iv*) $\mathbf{V} = \mathbf{W}$, $c \in \mathbf{K}$, $\mathcal{T} : \mathbf{V} \longrightarrow \mathbf{V}$, $\mathcal{T}(x) = cx$, $\forall\, x \in \mathbf{V}$. (For $c = 1$, $\mathcal{T}$ is the identity map of $\mathbf{V}$)

In Chapter 1, Rem. 5.4 we mentioned that $\mathcal{T}(0) = 0$, for any linear transformation $\mathcal{T}$. We will show next that more properties of $\mathbf{V}$ are transfered to $\mathbf{W}$ via $\mathcal{T}$.

THEOREM 1.1 *Let* $\mathcal{T} : \mathbf{V} \longrightarrow \mathbf{W}$ *be a linear transformation.*

(*i*) *If* $\mathbf{U}$ *is a vector subspace of* $\mathbf{V}$, *then* $\mathcal{T}(\mathbf{U})$ *is a vector subspace of* $\mathbf{W}$.

(*ii*) *If* $v_1, \ldots, v_p \in \mathbf{V}$ *are linearly dependent vectors, then* $\mathcal{T}(v_1), \ldots, \mathcal{T}(v_p)$ *are linearly dependent vectors in* $\mathbf{W}$.

Proof. (*i*) Let $w_1, w_2 \in \mathcal{T}(\mathbf{U})$ and $k, l \in \mathbf{K}$. Then $w_1 = \mathcal{T}(v_1)$, $w_2 = \mathcal{T}(v_2)$, for some $v_1, v_2 \in \mathbf{U}$ and

$$kw_1 + lw_2 = k\mathcal{T}(v_1) + l\mathcal{T}(v_2) = \mathcal{T}(kv_1 + lv_2)\ , \qquad \text{by (1.3)}\ .$$

On the other hand $kv_1+lv_2 \in \mathbf{U}$ since $\mathbf{U}$ is a subspace of $\mathbf{V}$. Therefore $kw_1+lw_2 \in \mathcal{T}(\mathbf{U})$.

(ii) By the linear dependence of $v_1, \ldots, v_p$ there exist $k_1, \ldots, k_p$, not all zero such that $k_1v_1 + \ldots + k_pv_p = 0$. Then $\mathcal{T}(k_1v_1 + \ldots + k_pv_p) = \mathcal{T}(0) = 0$. Now the linearity of $\mathcal{T}$ implies that $k_1\mathcal{T}(v_1) + \ldots + k_p\mathcal{T}(v_p) = 0$. QED

REMARK 1.2 Note that if in (ii) of the previous theorem we replace "dependent" by "independent", the statement we obtain is no longer true. The linear independence of vectors is preserved only by injective linear maps (see Chap. 1, Lemma 5.6).

THEOREM 1.3 *Assume* $dim\, \mathbf{V} = n$ *and let* $\mathbf{B} = \{e_1, \ldots, e_n\}$ *be a basis of* $\mathbf{V}$*, and* $w_1, \ldots, w_n$ *arbitrary vectors in* $\mathbf{W}$*.*

(i) *There is a unique linear transformation* $\mathcal{T} : \mathbf{V} \longrightarrow \mathbf{W}$ *such that*

$$\mathcal{T}(e_j) = w_j \ , \quad \forall\, j = 1, \ldots, n \ . \tag{1.4}$$

(ii) *The linear transformation* $\mathcal{T}$ *defined by* (1.4) *is injective if and only if* $w_1, \ldots, w_n$ *are linearly independent.*

Proof. (i) If $x \in \mathbf{V}$, then x can be represented as $x = \sum_{j=1}^{n} x_je_j$, $x_j \in \mathbf{K}$. Define $\mathcal{T}(x) = \sum_{j=1}^{n} x_jw_j \in \mathbf{W}$. This definition of $\mathcal{T}$ implies $\mathcal{T}(e_j) = w_j$, $\forall\, j = 1, \ldots, n$. Let also $y = \sum_{j=1}^{n} y_je_j \in \mathbf{V}$, and $k, l \in \mathbf{K}$. Then

$$\mathcal{T}(kx+ly) = \sum_{j=1}^{n}(kx_j + ly_j)w_j = k\sum_{j=1}^{n} x_jw_j + l\sum_{j=1}^{n} y_jw_j = k\mathcal{T}(x) + l\mathcal{T}(y) \ ,$$

hence $\mathcal{T}$ is linear. For the uniqueness, suppose $\mathcal{T}_1(e_j) = w_j$, $\forall\, j$.

For any $x = \sum_{j=1}^{n} x_je_j \in \mathbf{V}$ it follows that $\mathcal{T}_1(x) = \sum_{j=1}^{n} x_j\mathcal{T}_1(e_j) = \sum_{j=1}^{n} x_jw_j$, by the linearity of $\mathcal{T}_1$. Hence $\mathcal{T}_1(x) = \mathcal{T}(x)$, $\forall\, x \in \mathbf{V}$.

(ii) In Chap. 1, Lemma 5.6 we gave necessary and sufficient conditions for the injectivity of a linear transformation. From condition (iii) of that result, here it only remains to show that

$$w_1, \ldots, w_n \quad \text{linearly independent} \implies \mathcal{T} \text{ injective} \ .$$

Let $x = \sum_{j=1}^{n} x_je_j$, $y = \sum_{j=1}^{n} y_je_j \in \mathbf{V}$ such that $\mathcal{T}(x) = \mathcal{T}(y)$. Then $\mathcal{T}(x-y) = \sum_{j=1}^{n}(x_j - y_j)w_j = 0$. The linear dependence of $w_1, \ldots, w_n$ implies that $x_j = y_j$, $\forall\, j$, thus $x = y$. QED

Operations with Linear Transformations

Denote by $\mathcal{L}(\mathbf{V},\mathbf{W})$ the set of all linear maps defined on $\mathbf{V}$ with values into $\mathbf{W}$. *Addition* and *scalar multiplication* for the elements of $\mathcal{L}(\mathbf{V},\mathbf{W})$ are the usual operations for maps with values into an $\mathbf{K}$–vector space. (see Examples of vector spaces (v) Chapter 1); if $\mathcal{S},\mathcal{T}\in\mathcal{L}(\mathbf{V},\mathbf{W})$ then $\mathcal{S}+\mathcal{T}$ and $k\mathcal{T}$ are defined by

$$(\mathcal{S}+\mathcal{T})(x)=\mathcal{S}(x)+\mathcal{T}(x)\ ,\quad \forall\, x\in\mathbf{V} \tag{1.5}$$

$$(k\mathcal{T})(x)=k\mathcal{T}(x)\ ,\quad \forall\, x\in\mathbf{V},\ \forall\, k\in\mathbf{K}\ . \tag{1.6}$$

It is easy to see that $\mathcal{S}+\mathcal{T},\ k\mathcal{T}\in\mathcal{L}(\mathbf{V},\mathbf{W})$. Moreover, $\mathcal{L}(\mathbf{V},\mathbf{W})$ becomes a $\mathbf{K}$–vector space with the operations defined by (1.5), (1.6).

The elements of $\mathcal{L}(\mathbf{V},\mathbf{V})$ are called *endomorphisms* of $\mathbf{V}$. Another notation for $\mathcal{L}(\mathbf{V},\mathbf{V})$ is $End\,(\mathbf{V})$.

The vector space $\mathcal{L}(\mathbf{V},\mathbf{K})$ of all linear forms on $\mathbf{V}$ is called *the dual* of $\mathbf{V}$.

Let $\mathbf{U}$, $\mathbf{V}$, $\mathbf{W}$ be vector spaces over the same field $\mathbf{K}$ and $\mathcal{S}\in\mathcal{L}(\mathbf{U},\mathbf{V})$, $\mathcal{T}\in\mathcal{L}(\mathbf{V},\mathbf{W})$. Then the composite $\mathcal{T}\circ\mathcal{S}$ makes sense as a map from $\mathbf{U}$ into $\mathbf{W}$. It is straightforward that $\mathcal{T}\circ\mathcal{S}\in\mathcal{L}(\mathbf{U},\mathbf{W})$. The map $\mathcal{T}\circ\mathcal{S}$ is also written as $\mathcal{T}\mathcal{S}$ and is called the *product* of $\mathcal{T}$ and $\mathcal{S}$. The following properties are immediate:

$$(k\mathcal{A}+l\mathcal{B})\mathcal{C}=k\mathcal{A}\mathcal{C}+l\mathcal{B}\mathcal{C}\ ,\quad \forall\, k,l\in\mathbf{K},\ \mathcal{A},\mathcal{B}\in\mathcal{L}(\mathbf{U},\mathbf{V}),\ \mathcal{C}\in\mathcal{L}(\mathbf{V},\mathbf{W})\ . \tag{1.7}$$

$$\mathcal{C}(k\mathcal{A}+l\mathcal{B})=k\mathcal{C}\mathcal{A}+l\mathcal{C}\mathcal{B}\ ,\quad \forall\, k,l\in\mathbf{K},\ \mathcal{A},\mathcal{B}\in\mathcal{L}(\mathbf{V},\mathbf{W}),\ \mathcal{C}\in\mathcal{L}(\mathbf{U},\mathbf{V})\ . \tag{1.8}$$

If $\mathcal{T}\in End\,(\mathbf{V})$, denote

$$\mathcal{T}^0=\mathcal{I}\ ,\quad \mathcal{T}^2=\mathcal{T}\circ\mathcal{T}\ ,\ldots,\mathcal{T}^n=\mathcal{T}^{n-1}\circ\mathcal{T}=\mathcal{T}\circ\mathcal{T}^{n-1}\ ,$$

where $n\geq 1$ and $\mathcal{I}$ is the identity map of $\mathbf{V}$.

The product of endomorphisms is a binary operation on $End\,(\mathbf{V})$.

We mentioned in Chap. 1, Rem. 5.4 that the inverse of a bijective linear transformation is linear too (it is an easy exercise to prove it).

An injective homomorphism is called a *monomorphism*.

A surjective homomorphism is called an *epimorphism*.

A bijective homomorphism is called an *isomorphism* (see also Chap. 1, Def. 5.3, Rem. 5.4).

REMARK 1.4 Let us return to the main example. Let $\mathcal{A}\in\mathcal{L}(\mathbf{K}^n,\mathbf{K}^m)$, $\mathcal{B}\in\mathcal{L}(\mathbf{K}^m,\mathbf{K}^p)$ be defined as left multiplication by $A\in M_{m,n}(\mathbf{K})$ and $B\in M_{n,p}(\mathbf{K})$ respectively. Then:

(i) $\mathcal{B}\mathcal{A}\in\mathcal{L}(\mathbf{K}^n,\mathbf{K}^p)$ is the left multiplication by the matrix $AB\in M_{m,p}(\mathbf{K})$.

(ii) $\mathcal{A}$ is bijective $\Longleftrightarrow m=n$ and A is invertible.

In this case $\mathcal{A}^{-1}$ is the left multiplication by the matrix A^{-1}.

2 Kernel and Image

Let $\mathbf{V}$ and $\mathbf{W}$ be two vector spaces and $\mathcal{T} \in \mathcal{L}(\mathbf{V}, \mathbf{W})$.

DEFINITION 2.1 The set

$$Ker\,\mathcal{T} = \{x \in \mathbf{V}\,;\, \mathcal{T}(x) = 0\} \subset \mathbf{V}$$

is called the *kernel* of $\mathcal{T}$.

The *image* of $\mathcal{T}$ is the usual

$$Im\,\mathcal{T} = \mathcal{T}(\mathbf{V}) = \{y \in \mathbf{W}\,;\, \exists\, x \in \mathbf{V} \text{ such that } y = \mathcal{T}(x)\} \subset \mathbf{W}\ .$$

REMARKS 2.2 Recall that:

(*i*) $\mathcal{T}(0) = 0$, so $0_{\mathbf{V}} \in Ker\,\mathcal{T}$ and $0_{\mathbf{W}} \in Im\,\mathcal{T}$.

(*ii*) $Ker\,\mathcal{T} = \{0\}$ if and only if $\mathcal{T}$ is injective.

THEOREM 2.3 (*i*) *$Ker\,\mathcal{T}$ is a subspace of $\mathbf{V}$.*
(*ii*) *$Im\,\mathcal{T}$ is a subspace of $\mathbf{W}$.*

Proof. (*i*) Let $u, v \in Ker\,\mathcal{T}$, i.e. $\mathcal{T}(u) = \mathcal{T}(v) = 0$. Then

$$\mathcal{T}(ku + lv) = k\mathcal{T}(u) + l\mathcal{T}(v) = k0 + l0 = 0\ , \quad \forall\, k, l \in \mathbf{K}\ .$$

Thus $ku + lv \in Ker\,\mathcal{T}$, $\forall\, k, l \in \mathbf{K}$.

(*ii*) This is a particular case of Thm. 1.1 (*i*). QED

Example. Let $\mathcal{T} : \mathbf{R}^3 \longrightarrow \mathbf{R}^2$, $\mathcal{T}(x) = (2x_1 - x_2, x_1 + x_2 + 3x_3)$. We find

$$\begin{aligned} Ker\,\mathcal{T} &= \{(x_1, x_2, x_3)\ ;\ 2x_1 - x_2 = 0,\ x_1 + x_2 + 3x_3 = 0\ \} \\ &= \{(\alpha, 2\alpha, -\alpha)\ ;\ \alpha \in \mathbf{R}\ \} = Span\,\{(1, 2, -1)\}\ . \end{aligned}$$

$$Im\,\mathcal{T} = \left\{ (y_1, y_2)\ ;\ \text{the system} \ \left\{ \begin{array}{l} 2x_1 - x_2 = y_1 \\ 2x_1 + x_2 + 3x_3 = y_2 \end{array} \right. \ \text{is compatible} \right\} = \mathbf{R}^2\ .$$

Note that $\mathcal{T}$ is surjective, but not injective. We will see later in this section that a linear transformation between two finite dimensional spaces of the same dimension is surjective if and only if it is injective.

DEFINITION 2.4 The dimensions of $Im\,\mathcal{T}$ and $Ker\,\mathcal{T}$ are called the *rank* and *nullity* of $\mathcal{T}$ respectively.

PROPOSITION 2.5 *Let $y \in Im\,\mathcal{T}$. The general solution of the equation*

$$\mathcal{T}(x) = y \tag{2.1}$$

is the sum of the general solution of $\mathcal{T}(x) = 0$ and a particular solution of (2.1).

Proof. Denote $\mathbf{S}_y = \{x\ ;\ \mathcal{T}(x) = y\} = \mathcal{T}^{-1}(y)$, $\mathbf{S}_0 = \{x\ ;\ \mathcal{T}(x) = 0\} = Ker\,\mathcal{T}$, and fix $x_p \in \mathbf{S}_y$. We must show that $\mathbf{S}_y = \mathbf{S}_0 + \{x_p\}$.

For, let $\tilde{x} \in \mathbf{S}_y$. Then $\tilde{x} = (\tilde{x} - x_p) + x_p \in \mathbf{S}_0 + \{x_p\}$, since $\mathcal{T}(\tilde{x} - x_p) = \mathcal{T}(\tilde{x}) - \mathcal{T}(x_p) = y - y = 0$. Conversely, let $x_0 \in \mathbf{S}_0$; then $\mathcal{T}(x_0 + x_p) = \mathcal{T}(x_0) + \mathcal{T}(x_p) = y$, hence $x_0 + x_p \in \mathbf{S}_y$. QED

THEOREM 2.6 (The dimension formula) *Let $\mathcal{T} : \mathbf{V} \longrightarrow \mathbf{W}$ be a linear transformation and assume that $\mathbf{V}$ is finite dimensional. Then*

(2.2)
$$dim\,\mathbf{V} = dim\,(Ker\,\mathcal{T}) + dim\,(Im\,\mathcal{T})\ .$$

Proof. Say that dim $\mathbf{V} = n$. Then $Ker\,\mathcal{T}$ is finite dimensional too as a subspace of $\mathbf{V}$. Let $p = dim\,(Ker\,\mathcal{T})$, $1 \le p \le n-1$ and $\mathbf{B}_1 = \{u_1, \ldots, u_p\}$ a basis of $Ker\,\mathcal{T}$ (the cases $p = 0$, $p = n$ are left to the reader). Extend $\mathbf{B}_1$ to a basis of $\mathbf{V}$ (see Chap. 1, Thm. 4.8):
$$\mathbf{B} = \{u_1, \ldots, u_p, v_1, \ldots, v_{n-p}\}\ .$$

Denote $w_j = \mathcal{T}(v_j)$, $j = 1, \ldots, n-p$ and $\mathbf{B}_2 = \{w_1, \ldots, w_{n-p}\}$.

If we show that $\mathbf{B}_2$ is a basis of $Im\,\mathcal{T}$, it will follow that $dim\,(Im\,\mathcal{T}) = n - p$, and the proof of the theorem will be done. For, let $y \in Im\,\mathcal{T}$. Then $y = \mathcal{T}(x)$, for some $x \in \mathbf{V}$. We write x in terms of the basis $\mathbf{B}$ of $\mathbf{V}$:

$$x = a_1u_1 + \ldots + a_pu_p + b_1v_1 + \ldots + b_{n-p}v_{n-p}\ ,$$

and apply $\mathcal{T}$, using $\mathcal{T}(u_i) = 0$. We obtain:

$$y = \mathcal{T}(x) = b_1\mathcal{T}(v_1) + \ldots + b_{n-p}\mathcal{T}(v_{n-p}) = b_1w_1 + \ldots + b_{n-p}w_{n-p}\ .$$

This proves that $Im\,\mathcal{T} = Span\,\mathbf{B}_2$.

Now suppose that

(2.3)
$$k_1w_1 + \ldots + k_{n-p}w_{n-p} = 0\ , \quad k_1, \ldots, k_{n-p} \in \mathbf{K}\ .$$

From $w_j = \mathcal{T}(v_j)$ it follows that

$$k_1\mathcal{T}(v_1) + \ldots + k_{n-p}\mathcal{T}(v_{n-p}) = \mathcal{T}(k_1v_1 + ... + k_{n-p}v_{n-p}) = 0\ .$$

Denote $v = k_1v_1 + \ldots + k_{n-p}v_{n-p}$; thus $v \in Ker\,\mathcal{T}$. So we may write v in terms of the basis $\mathbf{B}_1$ of $Ker\,\mathcal{T}$, say $v = c_1u_1 + \ldots + c_pu_p$. Then

$$c_1u_1 + \ldots + c_pu_p - k_1v_1 - \ldots - k_{n-p}v_{n-p} = 0\ .$$

By the linear independence of $\mathbf{B}$, $c_1 = \ldots = c_p = k_1 = \ldots = k_{n-p} = 0$. Therefore all scalars in (2.3) vanish, which shows that $\mathbf{B}_2$ is linearly independent. Consequently, $\mathbf{B}_2$ is a basis of $Im\,\mathcal{T}$. QED

In Chapter 1, Lemma 5.6 we proved that for a linear map $\mathcal{T}$ there are equivalent characterizations of injectivity. Using the dimension formula, we can now state more such characterizations.

COROLLARY 2.7 *Let $\mathcal{T} \in \mathcal{L}(\mathbf{V}, \mathbf{W})$, $dim\, \mathbf{V} = n$, $n \in \mathbf{N}^*$. Then the following statements are equivalent:*

(i) *$\mathcal{T}$ is injective.*

(ii) *$dim\,(Im\, \mathcal{T}) = n$.*

(iii) *If $\{v_1, \dots, v_n\}$ is a basis of $\mathbf{V}$, then $\{\mathcal{T}(v_1), \dots, \mathcal{T}(v_n)\}$ is a basis of $Im\, \mathcal{T}$.*

Proof. "$(i) \Longrightarrow (ii)$" $\mathcal{T}$ injective$\Longrightarrow$ $Ker\, \mathcal{T} = \{0\} \Longrightarrow dim\,(Ker\, \mathcal{T}) = \{0\} \Longrightarrow dim\,(Im\, \mathcal{T}) = dim\, \mathbf{V} - 0 = n$, by the dimension formula (2.3).

"$(ii) \Longrightarrow (iii)$" It is straightforward that $\{\mathcal{T}(v_1), \dots, \mathcal{T}(v_n)\}$ spans $Im\, \mathcal{T}$. Then $dim\,(Im\, \mathcal{T}) = n$ implies that $\{\mathcal{T}(v_1), \dots, \mathcal{T}(v_n)\}$ is a basis of $Im\, \mathcal{T}$ (see Cor. 4.8, Chap. 1).

"$(iii) \Longrightarrow (i)$" (iii) says that $Im\, \mathcal{T}$ has a basis with n elements, so $dim\,(Im\, \mathcal{T}) = n = dim\, \mathbf{V}$. By the dimension formula, $dim\,(Ker\, \mathcal{T}) = 0$, so $Ker\, \mathcal{T} = \{0\}$, which is equivalent to (i). QED

3 The Matrix of a Linear Transformation

Throughout this section $\mathbf{V}$ and $\mathbf{W}$ denote two finite dimensional vector spaces over the same field $\mathbf{K}$, of dimensions n and m respectively; $\mathcal{T} : \mathbf{V} \longrightarrow \mathbf{W}$ is a linear transformation.

Let $\mathbf{B} = \{v_1, \dots, v_n\}$ and $\mathbf{C} = \{w_1, \dots, w_m\}$ be bases of $\mathbf{V}$ and $\mathbf{W}$ respectively.

DEFINITION 3.1 The matrix $T \in M_{m,n}(\mathbf{K})$ whose j'th column is the coordinate column vector of $\mathcal{T}(v_j)$, $\forall j \in \{1, ..., n\}$, is called the *matrix associated to* $\mathcal{T}$ (or the *matrix of* $\mathcal{T}$) *with respect to the bases* $\mathbf{B}$ *and* $\mathbf{C}$.

So, if for each $j \in \{1, ..., n\}$ we write $\mathcal{T}(v_j) = \sum\limits_{i=1}^{m} t_{ij} w_i$, then $T = [t_{ij}]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$.

Examples

(i) $\mathcal{T} : \mathbf{R}^2 \longrightarrow \mathbf{R}^3$, $\mathcal{T}(x) = (2x_1 + x_2,\, 3x_1,\, -x_1 + 5x_2)$.

Let $\mathbf{B} = \{v_1, v_2\}$; $\mathbf{C} = \{w_1, w_2\, w_3\}$ be the canonical bases of $\mathbf{R}^2$ and $\mathbf{R}^3$ respectively. Denote by T the associated matrix.

$T = M^{\mathcal{T}}_{\mathbf{B},\mathbf{C}}$.

Using the definition we find:

$$\mathcal{T}(v_1) = \mathcal{T}(1, 0) = (2, 3, -1) = 2f_1 + 3f_2 - f_3;$$

$$\mathcal{T}(v_2) = \mathcal{T}(0, 1) = (1, 0, 5) = f_1 + 5f_3.$$

Thus the 1'st column of T is $\begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}$ and the 2'nd column of T is $\begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix}$;

$$T = \begin{bmatrix} 2 & 1 \\ 3 & 0 \\ -1 & 5 \end{bmatrix}.$$

(*ii*) $\mathbf{V} = Span\{f_1, f_2\} \subset C^\infty(\mathbf{R})$; $f_1(x) = e^x \cos x$, $f_2(x) = e^x \sin x$, $\forall x \in \mathbf{R}$. $\mathcal{T} : \mathbf{V} \longrightarrow \mathbf{V}$, $\mathcal{T}(f) = f'$. $\mathbf{B} = \mathbf{C} = \{f_1, f_2\}$ is a basis of $\mathbf{V}$.

$\mathcal{T}(f_1)(x) = f_1'(x) = e^x \cos x - e^x \sin x$, $\forall x \in \mathbf{R}$. So, $\mathcal{T}(f_1) = f_1 - f_2$. Similarly, $\mathcal{T}(f_2) = f_1 + f_2$. Then $T = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$.

REMARKS 3.2 (*i*) $\mathcal{T}$ is uniquely determined by T and T is uniquely determined by $\mathcal{T}$. (This is immediate, since a linear transformation is uniquely determined by its values on the basis $\mathbf{B}$, and each $\mathcal{T}(v_j)$ has a unique representation w.r.t the basis $\mathbf{C}$.)

(*ii*) Note that in example (*i*) above, using the definition of the associated matrix w.r.t. the canonical bases we recovered the matrix mentioned in a particular case of the main example; if $\mathcal{T} : \mathbf{K}^n \longrightarrow \mathbf{K}^m$, is defined by

$$\mathcal{T}(x) = (\sum_{j=1}^{n} t_{1j}x_j, ..., \sum_{j=1}^{n} t_{mj}x_j),$$

then, according to Definition 3.1, the matrix associated to $\mathcal{T}$ w.r.t. the canonical bases of $\mathbf{K}^n$ and $\mathbf{K}^m$ is obviously the same as the coefficient matrix of the column ${}^t(\mathcal{T}(x))$, namely $T = [t_{ij}]$.

So, $\mathcal{T}$ acts as left multiplication by T.

Moreover, the next proposition points out that any linear transformation of finite dimensional vector spaces reduces to left multiplication by a matrix. This is why left multiplication by a matrix was called "the main example".

PROPOSITION 3.3 (the matrix of a linear transformation)
Let $\mathbf{V}$, $\mathbf{W}$, $\mathbf{B}$, $\mathbf{C}$, $\mathcal{T}$, T *as in Def.3.1.*

Let also $x \in \mathbf{V}$, $x = \sum\limits_{j=1}^{n} x_j v_j$, *and* $y \in \mathbf{W}$, $y = \sum\limits_{i=1}^{m} y_i w_i$ *such that* $y = \mathcal{T}(x)$. *Denote* $X =^t [x_1, ..., x_n]$, $Y =^t [y_1, ..., y_n]$. *Then* $\mathcal{T}$ *can be written in the matrix form as:*

$$Y = TX, \quad i.e.\ y_i = \sum_{j=1}^{n} t_{ij}x_j, \quad \forall i = 1, ..., m. \tag{3.4}$$

Proof. $\mathcal{T}(v_j) = \sum\limits_{i=1}^{m} t_{ij} w_i$, $\forall j = 1, ..., n$, since T is the matrix of $\mathcal{T}$.

Then $y = \mathcal{T}(x) = \sum\limits_{j=1}^{n} x_j \sum\limits_{i=1}^{m} t_{ij} w_i = \sum\limits_{i=1}^{m} (\sum\limits_{j=1}^{n} t_{ij}) w_i$.

By the uniqueness of the representation of y in the basis $\mathbf{C}$, $y_i = \sum\limits_{j=1}^{n} t_{ij}x_j$, $\forall i = 1, ..., m$. QED

COROLLARY 3.4 *Let* $\mathbf{U}$ *be another finite dimensional vector space, of basis* $\mathbf{D}$, $\mathcal{T} \in \mathcal{L}(\mathbf{V}, \mathbf{W})$, $\mathcal{S} \in \mathcal{L}(\mathbf{W}, \mathbf{U})$,

$T =$ *the matrix of* $\mathcal{T}$ *w.r.t.* $\mathbf{B}$ *and* $\mathbf{C}$;

$S =$ *the matrix of* $\mathcal{S}$ *w.r.t.* $\mathbf{C}$ *and* $\mathbf{D}$.

Then $ST =$ *the matrix of* $\mathcal{ST}$ *w.r.t.* $\mathbf{B}$, $\mathbf{D}$.

PROPOSITION 3.5 *Let $r = rank\, T$.*
Then $dim\,(Im\mathcal{T}) = r$ and $dim\,(Ker\mathcal{T}) = n - r$.

Proof. Denote $C_j =^t [t_{1j}, ..., t_{mj}]$ the j'th column of T. By Lemma 5.8 in Chapter 1, we may assume without loss of generality, that $C_1, ..., C_r$ are linearly independent in $M_{m,1} \simeq \mathbf{K}^m$ and any p columns of T are linearly dependent for $p \geq r+1$ (if $r < n$).

Consider the coordinate system $\psi : \mathbf{W} \to \mathbf{K}^m$ associated to the basis $\mathbf{C}$, $\psi(y) = (y_1, ..., y_m)$, for $y = \sum_{i=1}^{m} y_i w_i$. Then $\psi(\mathcal{T}(v_j)) =^t C_j$. Since ψ is an isomorphism, it preserves the linear dependence and the linear independence of vectors, and so does ψ^{-1}. Thus $\mathcal{T}(v_1), ..., \mathcal{T}(v_r)$ are linearly independent and any $p \geq r+1$ elements of $\{\mathcal{T}(v_1), ..., \mathcal{T}(v_n)\}$ are linear dependent.

On the other hand, $Im\,\mathcal{T} = Span\,(\mathcal{T}(v_1), ..., \mathcal{T}(v_n))$. It follows that $\{\mathcal{T}(v_1), ..., \mathcal{T}(v_r)\}$ is a basis of $Im\,\mathcal{T}$, therefore $dim\,(Im\,\mathcal{T}) = r$, then we apply the dimension formula. QED

We notice that the rank of the associated matrix T does not depend on the chosen bases. Also, the notation $rank\,\mathcal{T} = dim\,(Im\,\mathcal{T})$ makes now more sense, since the previous result shos that $dim\,(Im\,\mathcal{T})$ is indeed the rank of a certain matrix.

Using the above result, it is not hard to deduce the next corollary.

COROLLARY 3.6 *Let $\mathcal{T} \in \mathcal{L}(\mathbf{V}, \mathbf{W})$, $\mathbf{V}, \mathbf{W}$ finite dimensional, $T =$ the matrix of $\mathcal{T}$ w.r.t. $\mathbf{B}$ and $\mathbf{C}$.*

(*i*) *$\mathcal{T}$ is surjective if and only if $dim\,\mathbf{W} = rank\,T$.*

(*ii*) *$\mathcal{T}$ is injective if and only if $dim\,\mathbf{V} = rank\,T$.*

(*iii*) *$\mathcal{T}$ is bijective if and only if $dim\,\mathbf{V} = dim\,\mathbf{W} = rank\,T$.*
In this case T is invertible, and the matrix of $\mathcal{T}^{-1}$ w.r.t. $\mathbf{C}$, $\mathbf{B}$ is T^{-1}.

(*iv*) *If $dim\,\mathbf{V} = dim\,\mathbf{W} = n$, then:*

$$\mathcal{T} \textit{ is injective} \iff \mathcal{T} \textit{ is surjective} \iff \mathcal{T} \textit{ is bijective} \iff n = rank\,T.$$

Note that a particular case of (*iv*) is the case $\mathbf{V} = \mathbf{W}$, i.e. $\mathcal{T} \in End(\mathbf{V})$, $\mathbf{V}$ finite dimensional.

PROPOSITION 3.7 *Let $\mathcal{T} \in End(\mathbf{V})$, $dim\,\mathbf{V} = n$, $n \in \mathbf{N}$ and $\mathbf{B} = \{v_1, \ldots, v_n\}$ $\mathbf{B}' = \{v'_1, \ldots, v'_n\}$ two bases of $\mathbf{V}$. Denote by A and B the matrices associated to $\mathcal{T}$ w.r.t. the basis $\mathbf{B}$, and $\mathbf{B}'$ respectively.*
Then $B = C^{-1}AC$, where C is the matrix of change from $\mathbf{B}$ to $\mathbf{B}'$.

Proof. Let $C = [c_{ij}]$, $A = [a_{ij}]$, $B = [b_{ij}]$ and $C^{-1} = [d_{ij}]$. Then

$$v'_j = \sum_{i=1}^{n} c_{ij} v_i, \qquad v_j = \sum_{i=1}^{n} d_{ij} v'_i, \qquad \forall j = 1, ..., n. \tag{3.5}$$

$$\mathcal{T}(v_i) = \sum_{k=1}^{n} a_{ki} v_k, \quad \forall i = 1, ..., n; \qquad \mathcal{T}(v'_j) = \sum_{l=1}^{n} b_{lj} v'_l, \quad \forall j = 1, ..., n. \tag{3.6}$$

Combining (3.5) and (3.6) we obtain:

$$\mathcal{T}(v_j') = \sum_{i=1}^{n} c_{ij}\mathcal{T}(v_i) = \sum_{i=1}^{n}\sum_{k=1}^{n} c_{ij}a_{ki}v_k = \sum_{i=1}^{n}\sum_{k=1}^{n} c_{ij}a_{ki}\sum_{l=1}^{n} d_{lk}v_l' =$$

$$= \sum_{l=1}^{n}(\sum_{i=1}^{n}\sum_{k=1}^{n} d_{lk}a_{ki}c_{ij})v_l', \qquad \forall j = 1, ...n.$$

By the uniqueness of the representation of $\mathcal{T}(v_j')$ w.r.t. $\mathbf{B}'$ and (3.6), it follows that $b_{lj} = \sum\limits_{k=1}^{n}\sum\limits_{i=1}^{n} d_{lk}a_{ki}c_{ij}$, $\forall\, l, j = 1, ..., n$, thus $B = C^{-1}AC$. QED

DEFINITION 3.8 The matrices $A, B \in M_{n,n}(\mathbf{K})$ are called *similar* if there exists a nonsingular matrix $C \in M_{n,n}(\mathbf{K})$, such that $B = C^{-1}AC$.

By the previous proposition, two matrices are similar if and only if they represent the same endomorphism (with respect to different bases).

REMARKS 3.9 Properties of similar matrices

- (i) Similarity of matrices is an equivalence relation on $M_{n,n}(\mathbf{K})$.
- (ii) Similar matrices have the same rank.
- (iii) Nonsingular similar matrices have the same determinant. As a consequence, it makes sense to define *the determinant of an endomorphism* as the determinant of the associated matrix with respect to an arbitrary basis.

4 Particular Endomorphisms

DEFINITION 4.1 Let $\mathbf{V}$ be a $\mathbf{K}$-vector space and $\mathcal{F} \in \mathcal{L}(\mathbf{V},\mathbf{V})$.

- (i) $\mathcal{F}$ is an *automorphism* if it is bijective.
- (ii) $\mathcal{F}$ is a *projection* if $\mathcal{F}^2 = \mathcal{F}$.
- (iii) $\mathcal{F}$ is an *involution* if $\mathcal{F}^2 = id_V$.
- (iv) $\mathcal{F}$ is a *complex structure* if $\mathcal{F}^2 = -id_V$, for $\mathbf{K}$=$\mathbf{R}$.
- (v) $\mathcal{F}$ is *nilpotent of index* p if $\mathcal{F}^{p-1} \neq 0$, $\mathcal{F}^k = 0, \forall k \geq p$, where $p \in \{2, 3, \ldots\}$.

Denote $\mathcal{G}l(\mathbf{V}) = \{\mathcal{F} \in \mathcal{L}(\mathbf{V},\mathbf{V}) |\, \mathcal{F}$ is an automorphism $\}$. $\mathcal{G}l(\mathbf{V})$ is not a vector space, but it is a group with respect to the product (composition) of automorphisms. $\mathcal{G}l(\mathbf{V})$ is called the *general linear group* of $\mathbf{V}$.

THEOREM 4.2 *If* $\mathcal{F} : \mathbf{V} \to \mathbf{V}$ *is a projection, then* $\mathbf{V} = Ker\,\mathcal{F} \oplus Im\,\mathcal{F}$.

Proof. Let $v \in \mathbf{V}$, $\mathcal{F}(v) \in Im\mathcal{F}$, $w = v - \mathcal{F}(v) \in \mathbf{V}$. Then $\mathcal{F}(w) = \mathcal{F}(v) - \mathcal{F}^2(v) = 0$, thus $w \in Ker\,\mathcal{F}$, which shows that $\mathbf{V} = Ker\,\mathcal{F} + Im\,\mathcal{F}$. By the dimension formula, $dim\,\mathbf{V} = dim\,Ker\,\mathcal{F} + dim\,Im\,\mathcal{F}$, therefore $dim\,(Ker\mathcal{F} \cap Im\,\mathcal{F}) = 0$, so the sum is direct. QED

Obviously if $\mathcal{F}$ is a projection, then $id_{\mathbf{V}} - \mathcal{F}$ is a projection too. The theorem could be reformulated:

If $\mathcal{F}_1$ and $\mathcal{F}_2$ are projections such that $\mathcal{F}_1 + \mathcal{F}_2 = id_{\mathbf{V}}$, then $\mathbf{V} = Im\,\mathcal{F}_1 \oplus Im\,\mathcal{F}_2$.

The latter form has the following generalization.

THEOREM 4.3 *If $\mathcal{F}_i : \mathbf{V} \to \mathbf{V}$, $i = 1, ..., p$ are projections such that $\mathcal{F}_i\mathcal{F}_j = 0$, $\forall i \neq j$ and $\sum_{i=1}^{p} \mathcal{F}_i = id_V$, then $\mathbf{V} = Im\,\mathcal{F}_1 \oplus ... \oplus Im\,\mathcal{F}_p$.*

Proof. The conclusion follows easily from the previous theorem, by induction. Note that for the induction step we need $\mathcal{F}_1 + ... + \mathcal{F}_{p-1}$ to be a projection. To show this, use $\mathcal{F}_i^2 = \mathcal{F}_i$, and $\mathcal{F}_i\mathcal{F}_j = 0$, $\forall i \neq j$.

The details of the proof are left to the reader. QED

THEOREM 4.4 *A finite dimensional real vector space $\mathbf{V}$ admits a complex structure if and only if $dim\,\mathbf{V}$ is even.*

Proof. Suppose $dim\,\mathbf{V} = n = 2m$. Let $\mathbf{B} = \{e_1, ..., e_m, e_{m+1}, ..., e_{2m}\}$ be a basis of $\mathbf{V}$. Define the endomorphism $\mathcal{F} : \mathbf{V} \to \mathbf{V}$ by

$$\mathcal{F}(e_i) = e_{m+i}, \qquad \mathcal{F}(e_{m+i}) = -e_i, \quad i = 1, ..., m.$$

Then $\mathcal{F}(e_j) = -e_j$, $\forall j = 1, ..., n$, so $\mathcal{F}^2 = -id_V$. Conversely, let $\mathcal{F}$ be a complex struture on $\mathbf{V}$. Choose $v_1 \in \mathbf{V}$, $v_1 \neq 0$; then v_1, $\mathcal{F}(v_1)$ are linearly independent. For, suppose by contradiction that $\mathcal{F}(v_1) = kv_1$, for some $k \in \mathbf{R}$. It follows that $-v_1 = \mathcal{F}^2(v_1) = k\mathcal{F}(v_1) = k^2 v_1$; but this is impossible for $k \in \mathbf{R}$.

If $dim\,\mathbf{V} > 2$, then we can choose $v_2 \in \mathbf{V}$ such that v_1, v_2, $\mathcal{F}(v_1)$ are linearly independent. We can show similarly that v_1, v_2, $\mathcal{F}(v_1)$, $\mathcal{F}(v_2)$, are linearly independent. Continuing the process we obtain a basis which has an even number of elements. QED

Note that the matrix associated to the complex structure $\mathcal{F}$ with respect to the basis $\{v_1, ..., v_m, \mathcal{F}(v_1), ...\mathcal{F}(v_m)\}$ is $\begin{bmatrix} 0 & -I_m \\ -I_m & 0 \end{bmatrix}$.

THEOREM 4.5 *If $\mathcal{N} \in \mathcal{L}(\mathbf{V}, \mathbf{V})$ is a nilpotent endomorphism of index p, and $x_0 \in \mathbf{V}$ such that $\mathcal{N}^{p-1} \neq 0$, then the vectors x_0, $, \mathcal{N}(x_0), ..., \mathcal{N}^{p-1}(x_0)$ are linearly independent.*

Proof. Let $k_0, ..., k_{p-1} \in \mathbf{K}$ such that

$$\sum_{i=0}^{p-1} k_i \mathcal{N}^i(x_0) = 0.$$

Applying $\mathcal{N}^{p-1}$ to this equality we obtain $k_0 = 0$. Next, apply successively $\mathcal{N}^{p-2}$, $\ldots$, $\mathcal{N}^2$, $\mathcal{N}$ to obtain $k_1 = ... = k_{p-1} = 0$. QED

Note that $Span\,\{x_0, \mathcal{N}(x_0), ...\mathcal{N}^{p-1}(x_0)\}$ is an invariant subspace of $\mathbf{V}$.

THEOREM 4.6 *If* $dim\,\mathbf{V} = n \geq 1$, $\mathcal{T} \in \mathcal{L}(\mathbf{V}, \mathbf{V})$, *then there exist two subspaces* $\mathbf{U}$ *and* $\mathbf{W}$ *of* $\mathbf{V}$, *invariant with respect to* $\mathcal{T}$ *such that:*

(*i*) $\mathbf{V} = \mathbf{U} \oplus \mathbf{W}$;

(*ii*) $\mathcal{T}|_{\mathbf{U}}$ *is nilpotent;*

(*iii*) $\mathcal{T}|_{\mathbf{W}}$ *is invertible, when* $\mathbf{W} \neq \{0\}$.

Proof. Denote $N_k = Ker\,(\mathcal{T}^k)$ and $R_k = Im\,(\mathcal{T}^k),\ k \in \mathbf{N}^\star$. Obviously, these are invariant subspaces with respect to $\mathcal{T}$, and

$$N_k \subseteq N_{k+1},\ R_k \supseteq R_{k+1},\ \forall k.$$

Moreover,

$$N_k = N_{k+1} \Rightarrow N_{k+1} = N_{k+2} \text{ and } R_k = R_{k+1} \tag{4.7}$$

Let us prove (4.7). Assume $N_k = N_{k+1}$ and $v \in N_{k+2}$.

$$\mathcal{T}^{k+1}(\mathcal{T}v) = 0 \Rightarrow \mathcal{T}v \in N_{k+1} = N_k \Rightarrow \mathcal{T}^k(\mathcal{T}v) = 0 \Rightarrow v \in N_{k+1}.$$

To prove $R_k = R_{k+1}$, use the dimension formula

$$dim\,\mathbf{V} = dim\,N_k + dim\,R_k = dim\,N_{k+1} + dim\,R_{k+1},$$

which implies $dim\,R_k = dim\,R_{k+1}$, thus $R_k = R_{k+1}$, as $R_k \supseteq R_{k+1}$.

Now, since $\mathbf{V}$ is finite dimensional, there exists the smallest $p \in \mathbf{N}^\star$ such that

$$N_1 \subset N_2 \subset ... \subset N_p = N_{p+1} = N_{p+2} = ...$$

$$\text{and } R_1 \supset R_2 \supset ... \supset R_p = R_{p+1} = R_{p+2} = ...$$

Take $\mathbf{U} = N_p$, $\mathbf{W} = R_p$. Using the dimension formula it suffices to prove either $\mathbf{V} = \mathbf{U} + \mathbf{W}$ or $\mathbf{U} \cap \mathbf{W} = \{0\}$ in order to obtain (i). Let $v \in \mathbf{U} \cap \mathbf{W}$. Then $\mathcal{T}^p v = 0$ and $v = \mathcal{T}^p w$ for some w; it follows that $w \in N_{2p}$. But $N_{2p} = N_p$, so $\mathcal{T}^p w = 0$, which proves $\mathbf{U} \cap \mathbf{W} = \{0\}$, therefore (i).

Since $\mathbf{U}$ is $\mathcal{T}$-invariant, we can consider $\mathcal{T}|_{\mathbf{U}} : \mathbf{U} \to \mathbf{U}$; by the definition of p and N_p it follows that $\mathcal{T}^p(N_p) = \{0\}$ and $\mathcal{T}^{p-1}(N_p) \neq \{0\}$, so $\mathcal{T}|_{\mathbf{U}}$ is nilpotent of index p.

For (iii), assume $\mathbf{W} \neq \{0\}$. If $w \in \mathbf{W}$ and $\mathcal{T}w = 0$, then $w = \mathcal{T}^p u$, some u; it follows that $u \in N_{p+1}$. From $N_p = N_{p+1}$ follows $w = 0$. This shows that $\mathcal{T}|_{\mathbf{W}} : \mathbf{W} \to \mathbf{W}$ is injective, hence invertible (as $\mathbf{W}$ is finite dimensional). QED

Examples

(*i*) $\mathcal{T} : \mathbf{R}^3 \to \mathbf{R}^3$ defined by the matrix $T = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ is nilpotent of index 2 since $T^2 = 0$.

(*ii*) Let $\mathcal{D} : \mathbf{P}_3 \to \mathbf{P}_3$, where $\mathbf{P}_3$ is the space of polynomials of degree less or equal to 3, with real coefficients, $\mathcal{D}(p) = p'$. Then $\mathcal{D}$ is nilpotent of index 4, because $p^{(4)} = 0$, $\forall\, p \in \mathbf{P}_3$, while $\mathcal{D}^3 \neq 0$, since the third derivative of polynomials of degree 3 is not 0.

5 Endomorphisms of Euclidean Vector Spaces

Let $(\mathbf{V}, \langle\ ,\ \rangle)$ be a Euclidean $\mathbf{K}$-vector space, $\mathbf{K} = \mathbf{R}$ or $\mathbf{K} = \mathbf{C}$, and $\mathcal{T} \in \mathrm{End}(\mathbf{V})$.

When $\mathbf{V}$ is finite dimensional it is not hard to prove that there exists an endomorphism $\mathcal{T}^* \in \mathrm{End}(\mathbf{V})$, uniquely determined by $\mathcal{T}$ which satisfies

$$\langle x, \mathcal{T}y\rangle = \langle \mathcal{T}^* x, y\rangle\,, \quad \forall\, x, y \in \mathbf{V}\,. \tag{5.1}$$

We will accept without proof that all (possibly infinite dimensional) Euclidean vector spaces we are working with in this course have the property mentioned above for the finite dimensional case. Then the following definition makes sense.

DEFINITION 5.1 The endomorphism $\mathcal{T}^*$ defined by (5.1) is called the *adjoint* of $\mathcal{T}$. If $\mathbf{K} = \mathbf{R}$, then $\mathcal{T}^*$ is also called the *transpose* of $\mathcal{T}$.

REMARKS 5.2

(i) $id^*_{\mathbf{V}} = id_{\mathbf{V}}$.

(ii) Note that (5.1) is eqiuvalent to

$$\langle \mathcal{T}x, y\rangle = \langle x, \mathcal{T}^* y\rangle\,, \quad \forall\, x, y \in \mathbf{V}\,. \tag{5.2}$$

(iii) $(\mathcal{T}^*)^* = \mathcal{T}, \quad \forall\, \mathcal{T} \in End(\mathbf{V})$.

DEFINITION 5.3 If $\mathcal{T} = \mathcal{T}^*$, then $\mathcal{T}$ is called $Hermitian$, when $\mathbf{K} = \mathbf{C}$,
and *symmetric* when $\mathbf{K} = \mathbf{R}$, or
self adjoint in both cases (either $\mathbf{K} = \mathbf{C}$ or $\mathbf{K} = \mathbf{R}$).

If $\mathcal{T} = -\mathcal{T}^*$, then $\mathcal{T}$ is called *skew Hermitian* when $\mathbf{K} = \mathbf{C}$,
and *skew symmetric* (or *antisymmetric*), when $\mathbf{K} = \mathbf{R}$.

If $\mathcal{T}$ satisfies

$$\langle \mathcal{T}x, \mathcal{T}y\rangle = \langle x, y\rangle\,, \quad \forall\, x, y \in \mathbf{V} \quad \text{(i.e. } \mathcal{T} \text{ preserves the scalar product)}, \tag{5.3}$$

then $\mathcal{T}$ is called $unitary$, when $\mathbf{K} = \mathbf{C}$,
and $orthogonal$, when $\mathbf{K} = \mathbf{R}$.

The notions of Hermitian, symmetric, unitary etc. are used for matrices too. A matrix $A \in M_{n,n}(\mathbf{C})$ is called

symmetric if $A = {}^tA$

Hermitian if $A = \overline{{}^tA}$

skew symmetric if $A = -\,{}^tA$

skew Hermitian if $A = -\overline{{}^tA}$

orthogonal if $A\,{}^tA = I$ and $A \in M_{n,n}(\mathbf{R})$

unitary if $A\,\overline{{}^tA} = I$ and $A \in M_{n,n}(\mathbf{C})$

We will see later that the definitions given above for endomorphisms and matrices are related in a natural way.

PROPOSITION 5.4 *Let $\mathcal{T}$, $\mathcal{S} \in End(\mathbf{V})$.*

(*i*) $(\mathcal{T}\mathcal{S})^* = \mathcal{S}^*\mathcal{T}^*$.

(*ii*) *If $\mathcal{T}$ is invertible, then $\mathcal{T}^*$ is invertible too; moreover, $(\mathcal{T}^*)^{-1} = (\mathcal{T}^{-1})^*$.*

(*iii*) $(\mathcal{T} + \mathcal{S})^* = \mathcal{T}^* + \mathcal{S}^*$.

(*iv*) *If $\mathbf{K} = \mathbf{C}$, $k \in \mathbf{C}$, then $(k\mathcal{T})^* = \overline{k}\mathcal{T}^*$.*

(*v*) *If $\mathbf{K} = \mathbf{R}$, $k \in \mathbf{R}$, then $(k\mathcal{T})^* = k\mathcal{T}^*$.*

Proof.

(*i*) $$\langle x, (\mathcal{T}\mathcal{S})y\rangle = \langle x, \mathcal{T}(\mathcal{S}y)\rangle = \langle \mathcal{T}^*x, \mathcal{S}y\rangle = \langle \mathcal{S}^*(\mathcal{T}^*x), y\rangle = \langle (\mathcal{S}^*\mathcal{T}^*)x, y\rangle.$$

The adjoint of $\mathcal{T}\mathcal{S}$ is unique, thus $(\mathcal{T}\mathcal{S})^* = \mathcal{S}^*\mathcal{T}^*$.

(*ii*) We have already noticed that $id_{\mathbf{V}}{}^* = id_{\mathbf{V}}$. If we take the adjoint of both sides in $\mathcal{T}\mathcal{T}^{-1} = id_{\mathbf{V}}$ and $\mathcal{T}^{-1}\mathcal{T} = id_{\mathbf{V}}$, and apply (*i*), we get $(\mathcal{T}^{-1})^*\mathcal{T}^* = id_{\mathbf{V}} = \mathcal{T}^*(\mathcal{T}^{-1})^*$.

(*iii*) is immediate by the additivity of the inner product with respect to each argument.

Now, let us prove (*iv*).

$$\begin{aligned}\langle x, (k\mathcal{T})y\rangle &= \langle x, k(\mathcal{T}y)\rangle = \overline{k}\langle x, \mathcal{T}y\rangle \quad \text{by conjugate linearity}\\ &= \overline{k}\langle \mathcal{T}^*x, u\rangle = \langle (\overline{k}\mathcal{T}^*)x, y\rangle.\end{aligned}$$

Again, apply the uniqueness of the adjoint.

To prove (*v*) we proceed like in (*iv*), but use linearity with respect to the second argument instead of conjugate linearity. QED

We deduce easily the next corollary for self-adjoint endomorphisms.

COROLLARY 5.5 *Let $\mathcal{T}$, $\mathcal{S}$ be two self-adjoint endomorphisms of $\mathbf{V}$ ($\mathbf{K} = \mathbf{R}$ or $\mathbf{K} = \mathbf{C}$) and $k \in \mathbf{R}$. Then*

(*i*) *The endomorphism $\mathcal{T}\mathcal{S}$ is self-adjoint if and only if $\mathcal{T}\mathcal{S} = \mathcal{S}\mathcal{T}$.*

(*ii*) *If $\mathcal{T}$ is invertible, then $\mathcal{T}^{-1}$ is self-adjoint too.*

(*iii*) *$k\mathcal{T} + \mathcal{S}$ is self-adjoint.*

THEOREM 5.6 *Assume $\mathbf{K} = \mathbf{C}$. Then $\mathcal{T}$ is Hermitian if and only if*

$$\langle x, \mathcal{T}x\rangle \in \mathbf{R}\,, \quad \forall\, x \in \mathbf{V}\,.$$

Proof. If $\mathcal{T} = \mathcal{T}^*$, then $\langle \mathcal{T}x, x\rangle = \langle x, \mathcal{T}x\rangle$, by (5.1). But $\langle x, \mathcal{T}x\rangle = \overline{\langle \mathcal{T}x, x\rangle}$, by Hermitian symmetry of the inner product, so the scalar product $\langle x, \mathcal{T}x\rangle = \overline{\langle \mathcal{T}x, x\rangle}$ is real, $\forall\, x \in \mathbf{V}$.

Conversely, if $\langle x, \mathcal{T}x\rangle \in \mathbf{R}$, $\forall\, x \in \mathbf{V}$, it follows that

$$\langle x, \mathcal{T}x\rangle = \overline{\langle x, \mathcal{T}x\rangle} = \overline{\langle \mathcal{T}^*x, x\rangle} = \langle x, \mathcal{T}^*x\rangle\,, \quad \forall\, x \in \mathbf{V}\,.$$

Then $\langle x, (\mathcal{T} - \mathcal{T}^*)x \rangle = 0$, $\forall\, x \in \mathbf{V}$. If we replace x by $x + \alpha y$, $\alpha \in \mathbf{C}$, we obtain

$$\alpha \langle y, (\mathcal{T} - \mathcal{T}^*)x \rangle + \overline{\alpha} \langle x, (\mathcal{T} - \mathcal{T}^*)y \rangle = 0\,, \quad \forall\, x, y \in \mathbf{V}\,,\ \forall\, \alpha \in \mathbf{C}\,.$$

Now take $\alpha = 1$, then $\alpha = i$ to get $\langle x, (\mathcal{T} - \mathcal{T}^*)y \rangle = 0$, $\forall\, x, y \in \mathbf{V}$. If in the last equality $x = (\mathcal{T} - \mathcal{T}^*)y$, it follows that $\mathcal{T} - \mathcal{T}^* = 0$. QED

THEOREM 5.7 *The endomorphism $\mathcal{T}$ preserves the scalar product if and only if it preserves the norm, i.e.*

$$\langle \mathcal{T}x, \mathcal{T}y \rangle = \langle x, y \rangle\,, \quad \forall\, x, y \in \mathbf{V} \iff \|\mathcal{T}x\| = \|x\|\,, \quad \forall\, x \in \mathbf{V}\,.$$

Proof. "$\Rightarrow$" For $y = x$ we obtain $\|\mathcal{T}x\|^2 = \|x\|^2$, thus $\|\mathcal{T}x\| = \|x\|$.
"$\Leftarrow$" If $\mathbf{K} = \mathbf{C}$, then

$$\langle x, y \rangle = \frac{1}{4}\big(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\big)\,;$$

If $\mathbf{K} = \mathbf{R}$, then

$$\langle x, y \rangle = \frac{1}{4}\big(\|x+y\|^2 - \|x-y\|^2\big)\,.$$

In the left hand side of each identity replace $\langle x, y \rangle$ by $\langle \mathcal{T}x, \mathcal{T}y \rangle$, then use $\|\mathcal{T}(x + cy)\| = \|x + cy\|$, $c \in \{\pm 1, \pm i\}$, to end up with $\langle x, y \rangle$. QED

PROPOSITION 5.8 *$\mathcal{T}$ preserves the inner product if and only if $\mathcal{T}^*\mathcal{T} = id_{\mathbf{V}}$.*

Proof.

$$\begin{aligned} \langle \mathcal{T}x, \mathcal{T}y \rangle = \langle x, y \rangle,\ \forall\, x, y \in \mathbf{V} \quad &\Leftrightarrow \quad \langle \mathcal{T}^*\mathcal{T}x, y \rangle = \langle x, y \rangle,\ \forall\, x, y \in \mathbf{V} \\ &\Leftrightarrow \quad \langle (\mathcal{T}^*\mathcal{T} - id_{\mathbf{V}})x, y \rangle = 0,\ \forall\, x, y \in \mathbf{V} \\ &\Leftrightarrow \quad \mathcal{T}^*\mathcal{T} = id_{\mathbf{V}}. \qquad \text{QED} \end{aligned}$$

REMARKS 5.9

(i) If $\mathcal{T}$ preserves the inner product, then $\mathcal{T}$ is injective. (This may be deduced either from Prop.5.8 or Thm.5.7. By Prop.5.8, $\mathcal{T}$ admits a left inverse, thus it is injective. It is an easy exercise for the reader to prove that $Ker\,\mathcal{T} = \{0\}$ using Thm. 5.7.

(ii) If $\mathbf{V}$ is finite dimensional, then $\mathcal{T}$ preserves the inner product if and only if $\mathcal{T}^*\mathcal{T} = \mathcal{T}\mathcal{T}^* = id_{\mathbf{V}}$.

In some books an unitary endomorphism $\mathcal{T}$ is defined as an endomorphism which satisfies $\mathcal{T}^*\mathcal{T} = \mathcal{T}\mathcal{T}^* = id_{\mathbf{V}}$. This condition is stronger than the one we used in the definition here, but they are equivalent in the finite dimensional case. The equivalence comes up easily from Prop.5.8, if we recall that a linear transformation of finite dimensional vector spaces is injective if and only if it is bijective.

THEOREM 5.10 *Let $\mathbf{V}$ be finite dimensional, $dim\,\mathbf{V} = n$, $\mathbf{B} = \{e_1, \ldots, e_n\}$ an orthonormal basis of $\mathbf{V}$, and $T = [t_{ij}]$ the matrix of $\mathcal{T}$ w.r.t. $\mathbf{B}$.*

(I) Assume $\mathbf{K} = \mathbf{C}$. Then:

$(I.i)$	$\mathcal{T}$is Hermitian	$\Leftrightarrow$	$T$is Hermitian.
$(I.ii)$	$\mathcal{T}$is skew Hermitian	$\Leftrightarrow$	$T$is skew Hermitian.
$(I.iii)$	$\mathcal{T}$is unitary	$\Leftrightarrow$	$T$is unitary.

(II) Assume $\mathbf{K} = \mathbf{R}$. *Then:*

$(II.i)$	$\mathcal{T}$ is symmetric	$\Leftrightarrow$	T is symmetric.
$(II.ii)$	$\mathcal{T}$ is skew symmetric	$\Leftrightarrow$	T is skew symmetric.
$(II.iii)$	$\mathcal{T}$ is orthogonal	$\Leftrightarrow$	T is orthogonal.

Proof. The proofs of (II) are almost the same as the ones for (I). We will only prove $(I.i)$, leaving the rest to the reader.

Denote by $[t^*_{ij}]$ the matrix of $\mathcal{T}^*$ w.r.t. $\mathbf{B}$. Multiplying $\mathcal{T}e_j = \sum_{k=1}^{n} t_{kj}e_k$ by e_i in the sense of the inner product, we obtain

$$(5.4) \qquad \langle \mathcal{T}e_j, e_i\rangle = \left\langle \sum_k t_{kj}e_k, e_i \right\rangle = t_{ij}\,.$$

Similarly, $\langle \mathcal{T}^*e_j, e_i\rangle = t^*_{ij}$. But

$$(5.5) \qquad \langle \mathcal{T}^*e_j, e_i\rangle = \langle e_j, \mathcal{T}e_i\rangle = \overline{\langle \mathcal{T}e_i, e_j\rangle} = \overline{t_{ji}}\,.$$

If $\mathcal{T}$ is Hermitian, i.e. $\mathcal{T} = \mathcal{T}^*$, from (5.4), (5.5) follows $t_{ij} = \overline{t_{ji}}\,,\ \forall\, i, j$, hence $T = {}^t\overline{T}$.

Conversely, assume $t_{ij} = \overline{t_{ji}}\,,\ \forall\, i, j$. then

$$\begin{aligned}
\langle x, \mathcal{T}x\rangle &= \left\langle \sum_{j=1}^{n} x_je_j, \sum_{k=1}^{n} x_k\mathcal{T}e_k \right\rangle = \sum_{j,k=1}^{n} x_j\,\overline{x_k}\,\langle e_j, \mathcal{T}e_k\rangle \\
&= \sum_{j,k=1}^{n} x_j\,\overline{x_k}\,\overline{\langle \mathcal{T}e_k, e_j\rangle} = \sum_{j,k=1}^{n} x_j\,\overline{x_k}\,\overline{t_{jk}} = \sum_{j,k=1}^{n} x_j\,\overline{x_k}\,t_{kj} \\
&= \overline{\sum_{k,j=1}^{n} x_k\,\overline{x_j}\,\overline{t_{kj}}} = \overline{\langle x, \mathcal{T}x\rangle}\,,
\end{aligned}$$

thus $\langle x, \mathcal{T}x\rangle \in \mathbf{R}$ and $\mathcal{T}$ is Hermitian by Theorem 5.6. QED

Note that the theorem is no longer true if we remove the condition on the basis to be orthonormal, as we see in the next example.

Example

Let $\mathcal{T} : \mathbf{C}^2 \longrightarrow \mathbf{C}^2$ such that the matrix of $\mathcal{T}$ with respect to the canonical basis $\mathbf{B}$ is $T = \begin{bmatrix} 0 & i \\ -i & 2 \end{bmatrix}$. The canonical basis of $\mathbf{C}^2$ is orthonormal, and $T = {}^t\overline{T}$, thus $\mathcal{T}$ is Hermitian by the theorem. Consider the basis $\mathbf{B}_1 = \{f_1 = (1,0),\, f_2 = (1,1)\}$; $\langle f_1, f_2\rangle = 1 \neq 0$. Denote by T_1 the matrix of $\mathcal{T}$ w.r.t. $\mathbf{B}_1$. Then $T_1 = \begin{bmatrix} 2-i & i \\ 2+2i & -i \end{bmatrix}$, which is not a Hermitian matrix.

6 Isometries

Throughout this section $\mathbf{V}$ denotes a real Euclidean vector space, and $d : \mathbf{V}\times\mathbf{V} \to \mathbf{R}$ is the Euclidean distance. Orthogonal endomorphisms of $\mathbf{V}$ preserve the inner product, therefore they preserve the Euclidian norm and the Euclidean distance. The origin is a fixed point for any endomorphism. Let us introduce other maps which are distance preserving, but in general do not fix the origin.

DEFINITION 6.1 Let $a \in \mathbf{V}$. The map $\mathcal{T}\colon \mathbf{V} \to \mathbf{V}$ defined by $\mathcal{T}(x) = x + a$, $\forall x \in \mathbf{V}$, is called *the translation by the vector a*. In order to avoid confusion we denote sometimes this map by $\mathcal{T}_a$.

Note that in general translations are not linear; the only linear translation is the one corresponding to $a = 0$, namely the identity map.

The next properties of translations are straightforward.

THEOREM 6.2 *If $a, b \in \mathbf{V}$ and $\mathcal{T}_a$, $\mathcal{T}_b$ are the translations by a, and b respectively, then*

1) $\mathcal{T}_a \circ \mathcal{T}_b = \mathcal{T}_b \circ \mathcal{T}_a = \mathcal{T}_{a+b}$;

2) $\mathcal{T}_a$ is bijective, and its inverse is $(\mathcal{T}_a)^{-1} = \mathcal{T}_{-a}$.

It follows that composition of functions defines an abelian group structure on the set of all translations of $\mathbf{V}$, and this group is isomorphic to the additive group $(\mathbf{V}, +)$.

PROPOSITION 6.3 *Translations are distance preserving, i.e.*

$$d(\mathcal{T}(x), \mathcal{T}(y)) = d(x, y),\ \forall x, y \in \mathbf{V}.$$

Proof.

$$d(\mathcal{T}(x), \mathcal{T}(y)) = \|(y + a) - (x + a)\| = \|y - x\| = d(x, y),\ \forall x, y \in \mathbf{V}. \quad QED$$

DEFINITION 6.4 An *isometry* of $\mathbf{V}$ is a surjective map $\mathcal{F} : \mathbf{V} \to \mathbf{V}$ which preserves the Euclidean distance, i.e.

$$d(\mathcal{F}(x), \mathcal{F}(y)) = d(x, y),\ \forall x, y \in \mathbf{V}.$$

REMARKS 6.5 1) By the properties of distance, it is obvious that any distance preserving map is injective, therefore isometries are bijective maps.

2) Surjective orthogonal transformations are isometries.

3) Orthogonal transformations of finite dimensional Euclidean vector spaces are isometries.

4) Translations are isometries.

5) The product (composite map) of two isometries is an isometry. Moreover, the product of maps defines a group structure on the set of all isometries of the Euclidean vector space $\mathbf{V}$. The group of translations is a subgroup in this group.

The next theorem is the main step in proving that any isometry is the product of a translation and an orthogonal transformation.

THEOREM 6.6 *An isometry $\mathcal{F} : \mathbf{V} \to \mathbf{V}$ satisfying $\mathcal{F}(0) = 0$ is an orthogonal transformation.*

Proof. It is easy to see that $\mathcal{F}$ preserves the norm:

$$\|\mathcal{F}(x)\| = \|\mathcal{F}(x) - 0\| = d(0, \mathcal{F}(x)) = d(\mathcal{F}(0), \mathcal{F}(x)) =$$

$$= d(0, x) = \|x - 0\| = \|x\|, \ \forall x \in \mathbf{V}.$$

Using this property, we show next that $\mathcal{F}$ preserves the inner product.

$$d(\mathcal{F}(x), \mathcal{F}(y)) = d(x, y) \Leftrightarrow \|\mathcal{F}(y) - \mathcal{F}(x)\| = \|y - x\| \Leftrightarrow$$

$$\langle \mathcal{F}(y) - \mathcal{F}(x), \mathcal{F}(y) - \mathcal{F}(x) \rangle = \langle y - x, y - x \rangle \Leftrightarrow$$

$$\langle \mathcal{F}(y), \mathcal{F}(x) \rangle = \langle x, y \rangle, \quad \forall x, y \in \mathbf{V}.$$

It remains to show that $\mathcal{F}$ is linear. We show first that $\mathcal{F}$ is homogeneous.

$$\langle \mathcal{F}(kx), \mathcal{F}(y) \rangle = \langle kx, y \rangle = k\langle x, y \rangle =$$

$$= k\langle \mathcal{F}(x), \mathcal{F}(y) \rangle = \langle k\mathcal{F}(x), \mathcal{F}(y) \rangle, \ \forall x, y \in \mathbf{V}, \ \forall k \in \mathbf{R}.$$

Then

$$\langle \mathcal{F}(kx), \mathcal{F}(y) \rangle = \langle k\mathcal{F}(x), \mathcal{F}(y) \rangle \Rightarrow \langle \mathcal{F}(kx) - k\mathcal{F}(x), \mathcal{F}(y) \rangle = 0, \ \forall x, y \in \mathbf{V}, \ \forall k \in \mathbf{R}.$$

Fix now $x \in \mathbf{V}$ and $k \in \mathbf{R}$. By the surjectivity of $\mathcal{F}$, there exits $y \in \mathbf{V}$ such that $\mathcal{F}(kx) - k\mathcal{F}(x) = \mathcal{F}(y)$. For this y, using the positivity of the inner product we get $\mathcal{F}(kx) - k\mathcal{F}(x) = 0$, thus $\mathcal{F}$ is homogeneous.

We proceed similarly to prove the additivity of $\mathcal{F}$. Let $x, y \in \mathbf{V}$.

$$\langle \mathcal{F}(x + y), \mathcal{F}(z) \rangle = \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle =$$

$$= \langle \mathcal{F}(x), \mathcal{F}(z) \rangle + \langle \mathcal{F}(y), \mathcal{F}(z) \rangle = \langle \mathcal{F}(x) + \mathcal{F}(y), \mathcal{F}(z) \rangle.$$

Then

$$\langle \mathcal{F}(x + y) - \mathcal{F}(x) - \mathcal{F}(y), \mathcal{F}(z) \rangle = 0, \ \forall z \in \mathbf{V}.$$

Take z such that $\mathcal{F}(x+y) - \mathcal{F}(x) - \mathcal{F}(y) = \mathcal{F}(z)$, then use the positivity of the inner product to obtain $\mathcal{F}(x + y) - \mathcal{F}(x) - \mathcal{F}(y) = 0$, i.e. $\mathcal{F}$ additive. QED

THEOREM 6.7 (characterization of isometries) *If $\mathcal{F}$ is an isometry, then there exist a translation $\mathcal{T}$ and an orthogonal transformation $\mathcal{R}$ such that $\mathcal{F} = \mathcal{T} \circ \mathcal{R}$.*

Proof. Let $\mathcal{T}$ be the translation by vector $\mathcal{F}(0)$. Then $\mathcal{T}^{-1}$ is the translation by $-\mathcal{F}(0)$. The map $\mathcal{T}^{-1} \circ \mathcal{F}$ is an isometry which fixes the origin. By the previuos theorem, $\mathcal{T}^{-1} \circ \mathcal{F}$ is an orthogonal transformation. Denote $\mathcal{R} = \mathcal{T}^{-1} \circ \mathcal{F}$. Then $\mathcal{F} = \mathcal{T} \circ \mathcal{R}$. QED

Suppose now $dim\, \mathbf{V} = n \in \mathbf{N}^\star$ and consider an orthonormal basis $\mathbf{B} = \{e_1, \ldots, e_n\}$. If $\mathcal{R}$ is an orthogonal transformation, $\{\mathcal{R}(e_1), \ldots, \mathcal{R}(e_n)\}$ is an orthonormal basis too. The matrix R associated to $\mathcal{R}$ w.r.t. $\mathbf{B}$ is orthogonal, thus $(det\, \mathcal{R})^2 = 1$. If $det\, \mathcal{R} = +1$, $\mathcal{R}$ is called a *rotation.*

If $\mathcal{F} = \mathcal{T} \circ \mathcal{R}$ is an isometry such that $\mathcal{R}$ is a rotation, $\mathcal{F}$ is said to be a *positive* (or *orientation preserving) isometry.* When $det\, \mathcal{R} = -1$, $\mathcal{F}$ is said to be a *negative* (or *orientation reversing) isometry.*

7 Problems

1. Let $\bar{a} \neq \bar{0}$ be a fixed vector in the space of the free vectors $\mathbf{V}_3$, and the map $\mathcal{T} : \mathbf{V}_3 \to \mathbf{V}_3$, $\mathcal{T}(\bar{x}) = \bar{a} \times \bar{x}$.

1) Show that $\mathcal{T}$ is a linear transformation.
2) Show that $\mathcal{T}$ is neither injective, nor surjective.
3) Find $Ker(\mathcal{T})$, $Im(\mathcal{T})$, and show that $Ker(\mathcal{T}) \oplus Im(\mathcal{T}) = \mathbf{V}_3$.

2. Let P_n be the complex vector space of the polynomial functions of degree at most n. Show that the map $\mathcal{T} : P_n \to P_n$, defined by $\mathcal{T}p(x) = p(x+3) - p(x)$, $\forall x \in \mathbf{C}$, is a linear transformation. Is $\mathcal{T}$ injective?

3. In each of the following cases determine the matrix associated to $\mathcal{T}$ with respect to the canonical bases, the rank, and the nullity of $\mathcal{T}$.

1) $\mathcal{T} : \mathbf{R}^3 \to \mathbf{C}^3$, $\mathcal{T}(x) = ix$.

2) $\mathcal{T} : \mathcal{M}_{2\times 2}(K) \to \mathcal{M}_{2\times 2}(K)$, $\mathcal{T}(A) = {}^tA$.

3) $\mathcal{T} : \mathbf{C} \to \mathcal{M}_{2\times 2}(\mathbf{C})$, $\mathcal{T}(x) = x \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$.

4. Let $\bar{a} \neq \bar{0}$ be a fixed vector in $\mathbf{V}_3$, and the map $\mathcal{T} : \mathbf{V}_3 \to \mathbf{R}$, $\mathcal{T}(x) = \langle \bar{a}, \bar{x} \rangle$.

1) Show that $\mathcal{T}$ is a linear form.
2) Study the injectivity and the surjectivity $\mathcal{T}$.
3) Determine $Ker(\mathcal{T})$, $Im(\mathcal{T})$ and their dimensions.

5. Which of the following maps is a linear transformation? Determine its kernel and its image.

1) $\mathcal{T} : \mathbf{R}^3 \to \mathbf{R}^3, \mathcal{T}(x) = (x_1 + x_2, x_1 - x_2, 2x_1 + x_2 - x_3)$

2) $\mathcal{T} : \mathbf{R}^3 \to \mathbf{R}^3, \mathcal{T}(x) = (x_1 + x_2, x_1 - x_2 + 5, 2x_1 + x_2 - x_3)$

6. Determine the matrices of the endomorphisms $\mathcal{T}_j : \mathbf{R}^3 \to \mathbf{R}^3$, $j = 1, 2$, with respect to the basis $\{v_1 = (1,1,1),\ v_2 = (1,0,1),\ v_3 = (0,1,1)\}$, knowing that

$$1)\ T_1 = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 1 & 2 \\ 3 & 2 & 1 \end{bmatrix}, \quad 2)\ T_2 = \begin{bmatrix} -1 & 0 & 1 \\ 2 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix}$$

are the matrices of the transformations, with respect to the basis $\{w_1 = (1,2,3),$ $w_2 = (3,1,2),\ w_3 = (2,3,1)\}$.

7. 1) Show that the endomorphism $\mathcal{T} : \mathcal{M}_{3\times 3}(\mathbf{R}) \to \mathcal{M}_{3\times 3}(\mathbf{R})$, $\mathcal{T}(A) = {}^tA$ is a symmetric involution.

2) Let $\mathbf{V}$ be the real Euclidean vector space of C^∞ functions f defined on [a,b], with $f(a) = f(b)$. Show that the differentiation operator $D : \mathbf{V} \to \mathbf{V}$, $D(f) = f'$ is antisymmetric.

3) Show that the endomorphism $\mathcal{T} : \mathbf{R}^2 \to \mathbf{R}^2$,

$$\mathcal{T}(x) = (x_1 \cos\alpha - x_2 \sin\alpha, x_1 \sin\alpha + x_2 \cos\alpha), \quad \alpha \in \mathbf{R},$$

is orthogonal.

8. The endomorphism $\mathcal{T} : \mathbf{R}^3 \to \mathbf{R}^3$ is defined by

$$\mathcal{T}(x, y, z) = (x - y + z, y, y).$$

1) Show that $\mathcal{T}$ is a projection.
2) Determine the dimension and a basis for $Ker(\mathcal{T})$ and for $Im(\mathcal{T})$.

9. Let $\mathcal{T} : \mathbf{C}^2 \to \mathbf{C}^2$ be the endomorphism defined by the matrix

$$T = \begin{bmatrix} 6 - 2i & 2i \\ 4 + 8i & 0 \end{bmatrix}$$

with respect to the canonical basis of the complex vector space $\mathbf{C}^2$.

Find the Hermitian endomorphisms $\mathcal{T}_1$ and $\mathcal{T}_2$ such that $\mathcal{T} = \mathcal{T}_1 + i\mathcal{T}_2$, and their matrices w.r.t. the canonical basis.

Chapter 3

Eigenvectors and Eigenvalues

1 General Properties

Let $\mathbf{V}$ be a $\mathbf{K}$-vector space, and $\mathcal{T} : \mathbf{V} \to \mathbf{V}$ an endomorphism.

The concept of eigenvector is closely related to that of invariant subspace.

DEFINITION 1.1 A vector $x \in \mathbf{V} \setminus \{0\}$ is an *eigenvector* of $\mathcal{T}$ if there exists $\lambda \in \mathbf{K}$ such that

$$\mathcal{T}x = \lambda x. \tag{1.1}$$

The scalar λ is called the *eigenvalue* associated to the eigenvector x. The set of all eigenvalues of the endomorphism $\mathcal{T}$ is called the *spectrum* of $\mathcal{T}$ and is denoted by $Spec\,\mathcal{T}$.

Obviously, an eigenvector can be described as a basis of a one-dimensional $\mathcal{T}$-invariant subspace.

Geometrically, an eigenvector is a nonzero vector v, such that v and $\mathcal{T}v$ are collinear.

If $dim\,\mathbf{V} = n$ and v is an eigenvector of $\mathcal{T}$, $\mathcal{T}v = \lambda v$, we can extend $\{v\}$ to a basis $\{v_1 = v,\, v_2, \ldots, v_n\}$. Then the matrix of $\mathcal{T}$ with respect to this basis will be of the form: $T = \begin{bmatrix} \lambda & A \\ 0 & B \end{bmatrix}$, for some $A \in M_{1,n-1}(\mathbf{K})$, $B \in M_{n-1,n-1}(\mathbf{K})$.

DEFINITION 1.2 Let T be an $n \times n$ matrix with entries in $\mathbf{K}$. An *eigenvector* of T is a column vector $X \in \mathbf{K}^n \setminus \{0\}$, which is an eigenvector for left multiplication by T, i.e. $\exists\, \lambda \in \mathbf{K}$ such that

$$TX = \lambda X, \tag{1.2}$$

As above, λ is called an *eigenvalue* (of the matrix T).

REMARK 1.3 Let $\mathbf{V}$ be finite dimensional, $\mathbf{B}$ a fixed basis, T the matrix of $\mathcal{T}$ w.r.t. $\mathbf{B}$, $x \in \mathbf{V}$, $\lambda \in \mathbf{K}$, and $X \in \mathbf{K}^n$ the coordinate column of x with respect to $\mathbf{B}$. Then

$$\mathcal{T}x = \lambda x \Longleftrightarrow TX = \lambda X. \tag{1.3}$$

We saw in Chap.2, Section 3 that similar matrices represent the same endomorphism (with respect to different bases). By (1.1), (1.2), (1.3) follows that similar matrices have the same eigenvalues; if $B = C^{-1}AC$ and X an eigenvector for A, then $C^{-1}X$ is an eigenvector for B corresponding to the same eigenvalue.

Sometimes eigenvectors and eigenvalues are called *characteristic* (or *proper*) vectors, and *characteristic* (or *proper*) values, respectively.

DEFINITION 1.4 Let $\lambda \in \mathbf{K}$ and denote

$$\mathbf{S}(\lambda) = \{x \in \mathbf{V} \mid \mathcal{T}x = \lambda x\}.$$

If λ is an eigenvalue of $\mathcal{T}$, $\mathbf{S}(\lambda)$ is called the *eigenspace* of λ.

The following proposition is immediate; this type of result, that describes eigenvalues without involving eigenvectors will prove very helpful.

PROPOSITION 1.5 *(i)* $\mathbf{S}(\lambda) = Ker\,(\mathcal{T} - \lambda\, id_V)$.
(ii) λ *is an eigenvalue of* $\mathcal{T}$ *if and only if* $Ker\,(\mathcal{T} - \lambda\, id_V) \neq \{0\}$.

From (i) follows that $\mathbf{S}(\lambda)$ is a subspace of $\mathbf{V}$, since it is the kernel of a linear operator.

Examples

(i) Consider $\mathbf{V} = C^{\infty}(\mathbf{R})$ and the endomorphism $\mathcal{D} : \mathbf{V} \to \mathbf{V}$, $\mathcal{D}(f) = f'$.

Each $\lambda \in \mathbf{R}$ is an eigenvalue of $\mathcal{D}$; it is easy to check that the function f_λ, $f_\lambda(x) = e^{\lambda x}$ is an eigenvector for λ. Moreover, for a fixed λ we can solve the differential equation $f' = \lambda f$ whose solution is

$\mathbf{S}(\lambda) = \{f \mid f(x) = ce^{\lambda x},\ \forall x \in \mathbf{R}, \text{ for some } c \in \mathbf{R}\} = Span\,\{f_\lambda\}$.

In this example it happens that all eigenspaces are one-dimensional.

(ii) $\mathbf{V} = \mathbf{R}^3$, $\mathcal{T} : \mathbf{V} \to \mathbf{V}$, $\mathcal{T}(x, y, z) = (4x + 6y,\ -3x - 5y,\ -3x - 6y + z)$.

We can check that $\lambda_1 = -2$ is an eigenvalue with a corresponding eigenvector $v_1 = (1, -1, 3)$, and $\lambda_2 = 1$ is an eigenvalue with a corresponding eigenvector $v_2 = (-2, 1, 0)$, and another one $v_3 = (0, 0, 1)$.

At the end of the next section it will be striaghtforward that that -2 and 1 are the only eigenvalues, and $\mathbf{S}(-2) = \{(x, y, z) \mid x+y = 0,\ x-2y+z = 0\} = Span\,\{v_1\}$, $\mathbf{S}(1) = \{(x, y, z) \mid x + 2y = 0\} = Span\,\{v_2, v_3\}$; $dim\,\mathbf{S}(-2) = 1$, $\mathbf{S}(-2)$ is a straight line, and $dim\,\mathbf{S}(1) = 2$, $\mathbf{S}(1)$ is a plane.

THEOREM 1.6 *(i) For an eigenvector of* $\mathcal{T}$ *corresponds a single eigenvalue.*
(ii) Eigenvectors corresponding to distinct eigenvalues are linearly independent.
(iii) $\mathbf{S}(\lambda)$ *is an invariant subspace of* $\mathcal{T}$.
(iv) Eigenspaces corresponding to two distinct eigenvalues are independent.

Proof. (i) Let $v \in \mathbf{V} \setminus \{0\}$ such that $\mathcal{T}v = \lambda v$ and $\mathcal{T}v = \lambda_1 v$. Then $(\lambda - \lambda_1)v = 0$, $v \neq 0$ imply $\lambda = \lambda_1$.

(ii) Let $v_1, \ldots, v_p$ be eigenvectors, corresponding to the distinct eigenvalues $\lambda_1, \ldots, \lambda_p$. We proceed by induction on p. For $p = 1$, *(ii)* is obvious since $v_1 \neq 0$, as v_1 is an eigenvector. Assume *(ii)* is true for $p - 1$, $p \geq 2$. Suppose that

$$k_1 v_1 + \ldots + k_p v_p = 0. \tag{1.4}$$

Then $\mathcal{T}(k_1 v_1 + \ldots + k_p v_p) = 0$, and thus

$$k_1 \lambda_1 v_1 + \ldots + k_p \lambda_p v_p = 0 \tag{1.5}$$

If we multiply (1.4) by λ_p and subtract from (1.5), we get

$$k_1(\lambda_1 - \lambda_p)v_1 + \ldots + k_{p-1}(\lambda_{p-1} - \lambda_p)v_{p-1} = 0$$

By the inductive hypothesis this yields $k_1 = \ldots = k_{p-1} = 0$, since λ_j are distinct. Replacing in (1.4), follows $k_p v_p = 0$, therefore $k_p = 0$.

(iii) We have already noticed that $\mathbf{S}(\lambda)$ is a subspace. If $x \in \mathbf{S}(\lambda)$, then $\mathcal{T}(\mathcal{T}x) = \mathcal{T}(\lambda x) = \lambda(\mathcal{T}x)$. Thus $\mathcal{T}x \in \mathbf{S}(\lambda)$.

(iv) Let λ_1, λ_2 be distinct eigenvalues. We will show that $\mathbf{S}(\lambda_1) \cap \mathbf{S}(\lambda_2) = \{0\}$, i.e. $\mathbf{S}(\lambda_1)$ and $\mathbf{S}(\lambda_2)$ are independent. Let $v \in \mathbf{S}(\lambda_1) \cap \mathbf{S}(\lambda_2)$; $\mathcal{T}v = \lambda_1 v = \lambda_2 v$ implies $(\lambda_1 - \lambda_2)v = 0$. Since $\lambda_1 \neq \lambda_2$, it follows that $v = 0$. QED.

2 The Characteristic Polynomial

Throughout this section, $\mathbf{V}$ denotes a finite dimensional $\mathbf{K}$-vector space, $\mathbf{K} = \mathbf{R}$ or $\mathbf{K} = \mathbf{C}$, $dim_K \mathbf{V} = n$, and $\mathcal{T} : \mathbf{V} \to \mathbf{V}$ is an endomorphism.

Let also $A = [a_{ij}]$ be a $n \times n$ matrix with entries in $\mathbf{K}$.

LEMMA 2.1 *Let T be the matrix of $\mathcal{T}$ with respect to an arbitrary basis, and $\lambda \in \mathbf{K}$. Then λ is an eigenvalue of $\mathcal{T}$ if and only if $det\,(T - \lambda I) = 0$.*

Proof. We saw that λ is an eigenvalue of $\mathcal{T}$ if and only if it is an eigenvalue for T. The existence of an eigenvector of T is equivalent to the fact that the linear homogeneous system $(T - \lambda I)X = 0$ has nontrivial solutions. A necessary and sufficient condition for the existence of nontrivial solutions is that the determinant of the system is zero. QED

Now, in order to determine the eigenvalues of $\mathcal{T}$, it suffices to solve a polynomial equation of degree n. Once the eigenvalues are found, for each of them it remains to solve a linear system.

Next we will take a closer look to the case of matrices. For the matrix $A = [a_{ij}] \in M_{n,n}(\mathbf{K})$, the matrix equation $(A - \lambda I)X = 0$ can be written as a system

$$\begin{cases} (a_{11} - \lambda)x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + (a_{22} - \lambda)x_2 + \ldots + a_{2n}x_n = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + (a_{nn} - \lambda)x_n = 0 \end{cases}$$

which has nontrivial solutions if and only if

$$det(A-\lambda I) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ & & \ddots & \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = 0.$$

DEFINITION 2.2 The polynomial $P(\lambda) = det(A-\lambda I)$ is called the *characteristic polynomial of the matrix A,* and the equation $P(\lambda) = det(A-\lambda I) = 0$, $\lambda \in \mathbf{K}$, is called the *characteristic equation of the matrix A*. Sometimes we write $P_A(\lambda)$ in order to avoid confusion.

The eigenvalues of the matrix A are the solutions of its characteristic equation. This equation has always n complex solutions, counted with their multiplicities (the fundamental theorem of algebra!), but it might have no real ones. Sometimes the eigenvalues of A are understood to be complex, even if $\mathbf{K} = \mathbf{R}$; in this case the corresponding eigenvectors are in the complexification of $\mathbf{R}^n$, denoted by ${}^C\mathbf{R}^n \simeq \mathbf{C}^n$.

THEOREM 2.3 *(i) The characteristic polynomial of A has the form*

$$P(\lambda) = (-1)^n[\lambda^n - (tr\,A)\lambda^{n-1} + \ldots + (-1)^n det\,A]$$

(ii) Similar matrices have the same characteristic polynomial.
(iii) The matrices A and ${}^tA$ have the same characteristic polynomial.

Proof. (i) We omit the general proof; it can be shown that the coefficient of λ^k is the sum of principal minors of order k of A, multiplied by $(-1)^k$. A direct computation sorts out the problem for low dimensional cases.

For instance, suppose A is of order 3. Then

$$P(\lambda) = \begin{vmatrix} a_{11}-\lambda & a_{12} & a_{1n} \\ a_{21} & a_{22}-\lambda & a_{2n} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix} = -\lambda^3 + \lambda^2(tr\,A) - \lambda J + det\,A,$$

where $tr\,A = a_{11}+a_{22}+a_{33}$, $J = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$.

(ii) Assume $B = C^{-1}AC$. Then

$$\begin{aligned} det\,(B-\lambda I) &= det\,(C^{-1}AC-\lambda I) = det\,(C^{-1}(A-\lambda I)C) \\ &= det\,C^{-1}det\,(A-\lambda I)det\,C = det\,(A-\lambda I). \end{aligned}$$

(iii) $P_A(\lambda) = det\,(A-\lambda I) = det\,{}^t(A-\lambda I) = det\,({}^tA-\lambda I) = P_{{}^tA}(\lambda)$.

THEOREM 2.4 *The eigenvalues of a complex Hermitian matrix are real. In particular, all eigenvalues of a symmetric real matrix are real.*

Proof. Recall that a Hermitian matrix A is equal to the complex conjugate of its transpose: $A = \overline{{}^tA}$, or equvalently $\overline{A} = {}^tA$.

Let λ be an eigenvalue of A, and X an associated eigenvector. By conjugating the relation

$$(\star) \quad AX = \lambda X, \qquad \text{we find} \qquad (\star\star) \quad \overline{A}\overline{X} = \overline{\lambda}\overline{X}.$$

In $(\star)$ we left multiply by $\overline{{}^tX}$, then take the transpose of both sides to get ${}^tX{}^tA\overline{X} = \lambda{}^tX\overline{X}$. Left multiplicatin by ${}^tX$ in $(\star\star)$ implies ${}^tX\bar{A}X = \bar{\lambda}\,{}^tX\overline{X}$. Then $(\lambda-\bar{\lambda}){}^tX\overline{X} = 0$ since $\bar{A} = {}^tA$. But $X \neq 0$ since it is an eigenvector, so ${}^tX\overline{X} = \|X\|^2 \neq 0$. Therefore $\lambda = \bar{\lambda}$, i.e. $\lambda \in \mathbf{R}$. QED.

DEFINITION 2.5 Let T be the matrix of $\mathcal{T}$ associated to an arbitrary basis. The polynomial $P(\lambda) = det(T - \lambda I)$ is called the *characteristic polynomial* of the endomorphism $\mathcal{T}$. When we want to point out the endomorphism in order to avoid confusion, we will use the notation $P_{\mathcal{T}}(\lambda)$. The equation $P(\lambda) = 0$, $\lambda \in \mathbf{K}$, is called the *characteristic equation* of the endomorphism $\mathcal{T}$

Note that for each fixed $\lambda \in \mathbf{K}$, $P_{\mathcal{T}}(\lambda)$ does not depend on the basis we choose; for any other basis the matrix of $(\mathcal{T} - \lambda\, id_{\mathbf{V}})$ is similar to $T - \lambda I$, thus it has the same determinant. This shows that the characteristic polynomial of an endomorphism of a finite dimensional vector space depends on the endomorphism only, so the definition makes sense.

We may also define the terms *trace* and *determinant* of $\mathcal{T}$ to be those obtained using the matrix of $\mathcal{T}$ with respect to an arbitrary basis: $tr\,\mathcal{T} = tr\,T$, $det\,\mathcal{T} = det\,T$, since these are coefficients of $P_{\mathcal{T}}(\lambda)$, thus independent of the choice of basis (see also Chap2, Remarks 3.9).

The degree of $P(\lambda)$ is n, since $dim\,\mathbf{V} = n =$ the order of the matrix $(T - \lambda I)$. It follows that $\mathcal{T}$ has at most n distinct eigenvalues in $\mathbf{K}$. If this is this is the case, and we pick one eigenvector for each eigenvalue, we obtain a basis (see Thm 1.6 (ii)). The associated matrix of $\mathcal{T}$ is a diagonal matrix whose diagonal entries are the eigenvalues of $\mathcal{T}$. In the next section we will see what happens when $P(\lambda)$ has multiple roots.

Examples

(i) Consider again the endomorphism $\mathcal{T} : \mathbf{R}^3 \to \mathbf{R}^3$, $\mathcal{T}(x, y, z) = (4x + 6y,\ -3x - 5y,\ -3x - 6y + z)$, from the previous section. Our purpose is to determine all eigenvalues and the eigenvectors of $\mathcal{T}$.

The associated matrix with respect to the canonical basis is $T = \begin{bmatrix} 4 & 6 & 0 \\ -3 & -5 & 0 \\ -3 & -6 & 1 \end{bmatrix}$,
and the characteristic polynomial

$$P(\lambda) = \begin{vmatrix} 4-\lambda & 6 & 0 \\ -3 & -5-\lambda & 0 \\ -3 & -6 & 1-\lambda \end{vmatrix} = -(\lambda+2)(\lambda-1)^2.$$

Then the eigenvalues are the roots of $P(\lambda)$, namely -2 and 1. In order to find the eigenvectors, for each eigenvalue λ we solve the system $(T - \lambda I)X = 0$. Here, for $\lambda = -2$, we have to solve:

$$(T + 2I)X = 0 \Leftrightarrow \begin{cases} 6x + 6y = 0 \\ -3x - 3y = 0 \\ -3x - 6y + 3z = 0 \end{cases} \Leftrightarrow \begin{cases} x + y = 0 \\ -x - 2y + z = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} x = \alpha \\ y = -\alpha \\ z = -\alpha \end{cases} \Leftrightarrow X = \begin{bmatrix} \alpha \\ -\alpha \\ -\alpha \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix},\ \alpha \in \mathbf{R}.$$

The eigenspace of $\lambda = -2$ is the solution space of the above system, namely $\mathbf{S}(-2) = Span\{(1,-1,-1)\}$

For $\lambda = 1$, solve

$$(T-I)X = 0 \Leftrightarrow \begin{cases} 3x+6y &= 0 \\ -3x-6y &= 0 \\ -3x-6y &= 0 \end{cases} \Leftrightarrow x+2y = 0$$

$$\Leftrightarrow X = \begin{bmatrix} -2\alpha \\ \alpha \\ \beta \end{bmatrix} = \alpha \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \ \alpha, \beta \in \mathbf{R}.$$

and $\mathbf{S}(1) = Span\{(-2,1,0),(0,0,1)\}$.

The eigenvectors associated to an eigenvalue λ are <u>all</u> the nonzero elements of $\mathbf{S}(\lambda)$. We accept, however the answer of this problem given in the form:

"the eigenvalues are: $\lambda_1 = -2$, with eigenvector $v_1 = (1,-1,3)$; $\lambda_2 = 1$ with eigenvectors $v_2 = (-2,1,0)$ and $v_3 = (0,0,1)$." We will understand that for each eigenspace the answer points out the vectors of a basis.

(ii) Let $\mathbf{P}_3$ be the vector space of polynomials with real coefficients, of degree at most 3, and $\mathcal{T} : \mathbf{P}_3 \to \mathbf{P}_3$, $\mathcal{T}(p) = q$, where $q(X) = p(X+1)$. We want to determine the eigenvalues and the eigenvectors of $\mathcal{T}$.

First we need to fix a basis of $\mathbf{P}_3$. Let us choose the canonical basis $\{1, X, X^2, X^3\}$. The associated matrix is $T = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, and $P(\lambda) = (\lambda-1)^4$ with the only root $\lambda = 1$, of multiplicity 4. The coordinates of the eigenvectors are the nontrivial solutions of the system

$$(T-I)X = 0 \Leftrightarrow \begin{cases} x_2 + x_3 + x_4 = 0 \\ 2x_3 + 3x_4 = 0 \\ 3x_4 = 0 \end{cases} \Leftrightarrow X = \alpha \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \ \alpha \in \mathbf{R}.$$

The column ${}^t[1,0,0,0]$ represents the constant polynomial 1; $\mathbf{S}(1) = Span\{1\} = \{\text{the constant polynomials}\} \simeq \mathbf{R}$.

(iii) Let us determine also the eigenvalues and the eigenvectors of the matrix

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

From

$$P(\lambda) = \begin{vmatrix} -\lambda & 1 & 0 \\ 1 & 1-\lambda & 1 \\ 0 & 1 & -\lambda \end{vmatrix} = -\lambda(\lambda^2 - \lambda - 2)$$

we obtain: $\lambda_1 = -1$, with eigenvector $v_1 = {}^t[1,0,-1]$; $\lambda_2 = 0$, with eigenvector $v_2 = {}^t[1,-1,1]$; $\lambda_3 = 2$, with eigenvector $v_3 = {}^t[1,2,1]$.

3 The Diagonal Form

$\mathbf{V}$, $\mathbf{K}$ and $\mathcal{T}$ are as in the previous section; $n = dim_K\mathbf{V}$.

Working with endomorphisms of finite dimensinal spaces, usually requires a choice of basis. It is desirable to choose a basis such that the associated matrix is as simple as possible. We noticed that when the characteristic polynomial has n distinct roots in $\mathbf{K}$, there exists a basis made of eigenvectors; the matrix of $\mathcal{T}$ with respect to that basis is diagonal.

DEFINITION 3.1 The endomorphism $\mathcal{T}$ is called *diagonalizable* if there exists a basis of $\mathbf{V}$ such that the matrix of $\mathcal{T}$ with respect to this basis is diagonal.

A matrix $A \in M_{n,n}(\mathbf{K})$ is *diagonalizable* if it is similar to a diagonal matrix.

REMARK 3.2 *Let T be the matrix of $\mathcal{T}$ with respect to an arbitrary basis. Then $\mathcal{T}$ is diagonalizable if and only if T is diagonalizable.*

In this section we will give necessary and sufficient conditions for an endomorphism of a finite dimensional vector space (and for a square matrix) to be diagonalizable.

We call *diagonalization* of a (diagonalizable) endomorphism $\mathcal{T}$ the process of finding a basis with the property in Definition 3.1; we say that $\mathcal{T}$ has a *diagonal form* with respect to that basis.

For a matrix A of order n we call *diagonalization* of A the process of finding (if possible) an invertible matrix C and a diagonal matrix D of order n such that $D = C^{-1}AC$; the matrix D is called a *diagonal form of A* and the matrix C is called the *diagonalizing matrix.*

PROPOSITION 3.3 *The endomorphism $\mathcal{T}$ is digonalizable if and only if there exists a basis made up of eigenvectors of $\mathcal{T}$.*

Proof. If $\mathcal{T}$ is diagonalizable, then there exists a basis $\mathbf{B} = \{e_1, \ldots, e_n\}$ such that $M_{\mathbf{B}}^{\mathcal{T}} = D$ is diagonal. Let $D = \begin{bmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 \\ & & \ddots & \\ 0 & 0 & \ldots & d_n \end{bmatrix}$.

This means that $\mathcal{T}e_j = d_je_j$ $j = 1, ..., n$ i.e. the vectors in $\mathbf{B}$ are eigenvectors of $\mathcal{T}$, and $d_1, ..., d_n$ are the associated (not necessarily distinct) eigenvalues.

Conversely, If $\mathbf{B}_1 = \{v_1, ..., v_n\}$ is a basis of $\mathbf{V}$ such that each v_j is an eigenvector, then $\mathcal{T}v_j = \lambda_jv_j$, for some $\lambda_j \in \mathbf{K}$. It follows that the matrix of $\mathcal{T}$ with respect to $\mathbf{B}_1$ is

$$\begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ & & \ddots & \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}$$

(where the eigenvalues λ_j need not be distinct). QED

THEOREM 3.4 *The dimension of an eigenspace of the endomorphism $\mathcal{T}$ is less or equal to the multiplicity order of the corresponding eigenvalue, as a root of the characteristic polynomial.*

Proof. Let λ_0 be an eigenvalue of multiplicity order m. Denote $dim\,\mathbf{S}(\lambda_0) = p$. If $p = n$, then $\mathbf{S}(\lambda_0) = \mathbf{V}$ and $\mathcal{T} = \lambda_0 id_V$. It follows that $P(\lambda) = (\lambda_0 - \lambda)^n$, thus $m = n > p$.

Assume now $p < n$. Let $\{v_1, ..., v_p\}$ be a basis of $\mathbf{S}(\lambda_0)$. We extend this basis to a basis $\mathbf{B}$ of $\mathbf{V}$, $\mathbf{B} = \{v_1, ..., v_p, v_{p+1}, ..., v_n\}$. Since the first p vectors of $\mathbf{B}$ are eigenvectors associated to λ_0, we have:

$$\mathcal{T}(v_j) = \lambda_0 v_j,\ j = 1, ..., p; \quad \mathcal{T}(v_j) = \sum_{k=1}^{n} a_{kj} e_j,\ j = p+1, ..., n.$$

The matrix of $\mathcal{T}$ with respect to $\mathbf{B}$ is

$$T = \begin{bmatrix} \lambda_0 & \dots & 0 & a_{1p+1} & \dots & a_{1n} \\ & \ddots & & & & \\ 0 & \dots & \lambda_0 & a_{pp+1} & \dots & a_{pn} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & a_{np+1} & \dots & a_{nn} \end{bmatrix},$$

thus $P(\lambda) = (\lambda_0 - \lambda)^p D(\lambda)$, where $D(\lambda)$ is a determinant of order $n - p$.

On the other hand, the multiplicity of λ_0 as a root of $P(\lambda)$ is m, so $P(\lambda) = (\lambda - \lambda_0)^m Q(\lambda)$, for some polynomial $Q(\lambda)$ such that λ_0 is not a root of $Q(\lambda)$. From these two factorizations of $P(\lambda)$ follows that $p \leq m$. QED

THEOREM 3.5 *The endomorphism $\mathcal{T}$ is diagonalizable if and only if its characteristic polynomial has all its n roots (counted with their multiplicities) in $\mathbf{K}$, and the dimension of each eigenspace is equal to the multiplicity order of the corresponding eigenvalue.*

Proof. Let $\lambda_1, \lambda_2, \dots, \lambda_p \in \mathbf{K}$ be the distinct eigenvalues of $\mathcal{T}$, and denote by $m(\lambda_j) = m_j$ the multiplicity of λ_j, $j = 1, ..., p$. Then $P(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \dots (\lambda - \lambda_p)^{m_p} Q(\lambda)$, and $\sum_{j=1}^{p} m_j \leq n$. Obviously, the condition that $P(\lambda)$ has all roots in $\mathbf{K}$ is equivalent to $\sum_{j=1}^{p} m_j = n$; if this is the case, $Q(\lambda) \equiv 1$.

Suppose first that $\mathcal{T}$ is diagonalizable, so there is a basis $\mathbf{B} = \{e_1, \dots, e_n\}$ of $\mathbf{V}$, made up of eigenvectors.

Denote $s_j =$ the number of vectors in $\mathbf{B}$ which are eigenvectors for λ_j, and $p_j = dim\,\mathbf{S}(\lambda_j)$. The elements of $\mathbf{B}$ are linearly independent, thus $s_j \leq p_j$. Since $card\,\mathbf{B} = n$, and each eigenvector corresponds to a single eigenvalue, it follows that $\sum_{j=1}^{p} s_j = n$. By the previous theorem, $p_j \leq m_j$. Now

$s_j \leq p_j \leq m_j$, $\forall j = 1, ..., p$ and $\sum_{j=1}^{p} m_j \leq n = \sum_{j=1}^{p} s_j$

force $s_j = p_j = m_j$, $\forall j$, and $\sum_{j=1}^{p} m_j = n$.

Conversely, assume that $\sum_{j=1}^{p} m_j = n$ and $dim\,\mathbf{S}(\lambda_j) = m_j$, $\forall j$. Consider the ordered set $\mathbf{B}_1 = \{v_1, ..., v_n\}$ whose n elements are chosen such that the first m_1 form a basis in $\mathbf{S}(\lambda_1)$, the next m_2 form a basis in $\mathbf{S}(\lambda_2)$, and so on, up to the last m_p elements which form a basis in $\mathbf{S}(\lambda_p)$. Then the elements of $\mathbf{B}$ are distinct eigenvectors (see Thm. 1.6 (*iv*)). Moreover, using induction on p, one shows that $\mathbf{B}$ is a basis of $\mathbf{V}$. By Prop. 3.3, this means that $\mathcal{T}$ is diagonalizable. QED

COROLLARY 3.6 *If $\mathcal{T}$ is diagonalizable and $\lambda_1, \lambda_2, \ldots, \lambda_p$ are its distinct eigenvalues, then*

$$\mathbf{V} = \mathbf{S}(\lambda_1) \oplus \ldots \oplus \mathbf{S}(\lambda_p).$$

It is clear now that not all endomorphisms of finite dimensional spaces (and not all matrices) are diagonalizable.

From the proof of Theorem 3.5 follows that for diagonalizable endomorphisms (matrices) the diagonal form is unique up to the order of the diagonal entries. The diagonalizing matrix is not unique either. Moreover, there are infinitely many diagonalizing matrices corresponding to the same diagonal form.

Diagonalization Algorithm

1) Fix a basis of $\mathbf{V}$ and determine the matrix T of $\mathcal{T}$ with respect to that basis.

2) Determine the eigenvalues of $\mathcal{T}$ by solving the characteristic equation, $P(\lambda) = 0$.

3) If $\mathbf{K} = \mathbf{R}$ and there are non-real roots of $P(\lambda)$, then we stop with the conclusion that $\mathcal{T}$ is not diagonalizable.

Otherwise, move to step 4).

4) For each eigenvalue λ_j check whether the multiplicity m_j is equal to $dim\, \mathbf{S}(\lambda_j)$. For, it suffices to verify if

$$m_j = n - rank\,(T - \lambda_j I),\ \forall j.$$

If there exists at least one j such that $m_j > n - rank\,(T - \lambda_j I)$, then we stop; $\mathcal{T}$ is not diagonalizable, by Thm. 3.5.

If all equalities hold, point out that $\mathcal{T}$ is diagonalizable and go to step 5).

5) Solve the p systems $(T - \lambda_j I)X = 0$, where p is the number of distinct eigenvalues. For each system chose m_j independent solutions, that represent the coordinates of vectors of a basis in $\mathbf{S}(\lambda_j)$. Form a basis of $\mathbf{V}$ such that the first m_1 vectors form a basis of $\mathbf{S}(\lambda_1)$, the next m_2 form a basis of $\mathbf{S}(\lambda_2)$, and so on.

6) The matrix of $\mathcal{T}$ associated to the basis formed in 4) is diagonal; its diagonal entries are:

$$\lambda_1, \ldots \lambda_1; \lambda_2, \ldots, \lambda_2; \lambda_p, \ldots, \lambda_p$$

where each λ_j appears m_j times. Let us denote this diagonal matrix by D.

7) The diagonalizing matrix is C whose columns are the solutions of the systems in 5), i.e. the coordinate-change matrix from the initial basis to the basis formed by eigenvectors.

Examples

(*i*) The endomorphism $\mathcal{T} : \mathbf{R}^3 \to \mathbf{R}^3$, $\mathcal{T}(x, y, z) = (4x+6y,\ -3x-5y,\ -3x-6y+z)$. studied in the previous sections is diagonalizable. Its matrix with respect to the basis

$\{v_1 = (1,-1,-1), v_2 = (-2,1,0), v_3 = (0,0,1)\}$ is $D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. The diagonalizing matrix is $C = \begin{bmatrix} 1 & -2 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$, which satisfies $D = C^{-1}TC$.

(ii) The endomorphism $\mathcal{T} : \mathbf{P}_3 \to \mathbf{P}_3$, $\mathcal{T}(p) = q$, $q(X) = p(X+1)$ whose eigehenvalues and eigenspaces were determined in Section 2 is not diagonalizable, nor is its matrix with respect to the canonical basis, $T = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

We saw that $\lambda = 1$ is the only eigenvalue and $m(1) = 4$, while $dim\, \mathbf{S}(1) = 4 - rank\,(T - I) = 1$.

4 The Canonical Jordan Form

Let $\mathbf{V}$ be a finite dimensional $\mathbf{K}-$vector space, $\mathbf{K} = \mathbf{R}$ or $\mathbf{K} = \mathbf{C}$ as before; $n = dim_{\mathbf{K}}\mathbf{V}$, $\mathcal{T} : \mathbf{V} \to \mathbf{V}$ an endomorphism, $T =$ the matrix of $\mathcal{T}$ with respect to an arbitrary basis, $A \in M_{n,n}(\mathbf{K})$.

In this section we will see that even when an endomorphism (or matrix) is not diagonalizable, but all its eigenvalues are in $\mathbf{K}$, it is still possible to find a "canonical" form that is convenient to work with.

DEFINITION 4.1 Let $\lambda \in \mathbf{K}$, $k \in \mathbf{N}^*$. The matrix

$$\begin{bmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ 0 & 0 & \dots & \dots & \lambda & 1 \end{bmatrix} \in M_{k,k}(\mathbf{K})$$

is said to be *the Jordan cell* (or *Jordan block*) of order k associated to the scalar λ.

The Jordan cells of order 1, 2 and 3 respectively are:

$$[\lambda], \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}.$$

DEFINITION 4.2 We say that the endomorphism $\mathcal{T}$ is *Jordanizable* (or *admits a canonical Jordan form*) if there exists a basis in $\mathbf{V}$ such the associated matrix of $\mathcal{T}$ has the form

$$(4.1) \qquad J = \begin{bmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \dots & J_s \end{bmatrix} \text{ (the canonical Jordan form,)}$$

where each J_i is a Jordan cell associated to some scalar λ_i.; $\lambda_1, ..., \lambda_s$ are not necessarily distinct.

A matrix J of type (4.1) is said to be in *Jordan form* or to be a *Jordan matrix.*

We call the matrix A *Jordanizable* if it is similar to a Jordan matrix; if $J = C^{-1}AC$, then C is called the *Jordanizing matrix* of A.

If T is the matrix of $\mathcal{T}$ with respect to the basis $\mathbf{B}$, and the matrix of $\mathcal{T}$ with respect to some other basis $\mathbf{B}_1$ is the Jordan matrix J, let us denote by C the coordinate-change matrix from $\mathbf{B}$ to $\mathbf{B}_1$.Then $J = C^{-1}TC$, so the endomorphism $\mathcal{T}$ is Jordanizable if and only if its matrix T is Jordanizable.

The basis $\mathbf{B}_1$ is called a *Jordan basis.*

We saw that if an endomorphism admits a diagonal form, the corresponding basis is made up of eigenvectors. Let us take a closer look at the Jordan basis.

Let J be a Jordan matrix, which is the matrix of $\mathcal{T}$, $\{e_1, ..., e_n\}$ the correspoding Jordan basis, and J_1 the first cell of J. Assume J_1 is of dimension $k_1 \geq 2$ and its diagonal entries are equal to $\lambda_1 \in \mathbf{K}$. Then we observe that

$$\mathcal{T}e_1 = \lambda_1 e_1,\ \mathcal{T}e_2 = e_1 + \lambda_1 e_2,\ ...,\ \mathcal{T}e_{k_1} = e_{k_1 - 1} + \lambda_1 e_{k_1}, \tag{4.2}$$

Therefore λ_1 is an eigenvalue of $\mathcal{T}$ with eigenvector e_1. The vectors $e_2, ..., e_{k_1}$ are called *principal vectors* associated to the eigenvector e_1.

Similarly, the scalar on the diagonal of each Jordan block is an eigenvalue, and for each J_i of dimension k_i corresponds a sequence of k_i basis vectors, such that the first one is an eigenvector; if $k_i \geq 2$, the other $k_i - 1$ are principal vectors.

REMARKS 4.3 (*i*) The diagonal form is a particular case of the Jordan form, where the dimension of each cell is 1.

(*ii*) The Jordan form is unique up to the permutation of blocks.

(*iii*) The order of the diagonal blocks depends on the order of basis vectors.

(*iv*) Assume $e_1, \ldots, e_{k_1}$ are nonzero vectors satisfying (4.2). Then

$$(\mathcal{T} - \lambda_1 id_{\mathbf{V}})e_1 = 0, \quad (\mathcal{T} - \lambda_1 id_{\mathbf{V}})^2 e_2 = 0, \quad (\mathcal{T} - \lambda_1 id_{\mathbf{V}})^{k_1} e_{k_1} = 0.$$

Moreover, $e_i \in Ker\,(\mathcal{T} - \lambda_1 id_{\mathbf{V}})^i \setminus Ker\,(\mathcal{T} - \lambda_1 id_{\mathbf{V}})^{i-1}, \quad \forall i = 1, \ldots, k_1$.

Now it takes two more lines to show that $e_1, \ldots, e_{k_1}$ are linearly independent.

(*v*) Let L be an ordered set of distinct t vectors formed by $l \geq 2$ sequences, each of length $s_i \in \mathbf{N}^\star$, $\forall i = 1, \ldots, l$ such that the first vector in each sequence is an eigenvector, and the next ones are associated principal vectors; assume also that the subset of L formed with all eigenvectors in L (which need not correspond to distinct eigenvalues) is linearly independent.

Then it can be shown by induction on t that L is linearly independent.

Note that if we are able to produce such a set L with $t = n$, then we deduce that $\mathcal{T}$ is Jordanizable and L is a Jordan basis.

Based on the fact that if λ is an eigenvalue for the matrix T, then λ^m is an eigenvalue for T^m, $\forall\, m \in \mathbf{N}^*$, we deduce the next lemma; the details of its proof are left as an execise.

LEMMA 4.4 *The only complex eigenvalue of a nilpotent matrix (nilpotent endomorphism of a finite dimensional complex vector space) is zero.*

THEOREM 4.5 *If $\lambda_1, ..., \lambda_p$ are the distinct eigenvalues of $\mathcal{T}$, with multiplicities $m(\lambda_j) = m_j$, $j = 1, ..., p$ and $\sum\limits_{j=1}^{p} m_j = n$, then there exist p $\mathcal{T}$-invariant subspaces $\mathbf{V}_j \subset \mathbf{V}$, $j = 1, \ldots, p$ such that:*

(*i*) $dim\, \mathbf{V}_j = m_j$, $j = 1, \ldots, p$

(*ii*) $\mathbf{V} = \mathbf{V}_1 \oplus \mathbf{V}_2 \oplus \ldots \oplus \mathbf{V}_p$

(*iii*) $\mathcal{T}|_{\mathbf{V}_j} = \mathcal{N}_| + \lambda_|\rangle\lceil_{\mathbf{V}}$, $| = \infty, \ldots, \surd$ *where* $\mathcal{N}_\infty, \ldots, \mathcal{N}_\surd$ *are nilpotent endomorphisms of various orders.*

Proof. for each fixed $j \in \{1, \ldots, p\}$, consider the endomorphisms $\mathcal{T}_j = \mathcal{T} - \lambda_j id_{\mathbf{V}}$ and apply Thm. 4.6, Chap.2 to obtain the subspaces $\mathbf{V}_j$ and $\mathbf{W}_j$ such that $\mathbf{V} = \mathbf{V}_j \oplus \mathbf{W}_j$, and $\mathcal{T}_j|_{\mathbf{V}_j}$ is nilpotent and $\mathcal{T}_j|_{\mathbf{W}_j}$ is invertible. Since $\mathbf{V}_j$ is $\mathcal{T}_j$-invariant, it follows that it is also $\mathcal{T} = \mathcal{T}_j + \lambda_j id_{\mathbf{V}}$-invariant.

Let $\mathcal{T}|_{\mathbf{V}_j} \in End\,(\mathbf{V}_j)$ and $\mathcal{T}|_{\mathbf{W}_j} \in End\,(\mathbf{W}_j)$ be the restrictions of $calT$ to $\mathbf{V} - j$ and to $\mathbf{W}_j$ respectively. From $\mathbf{V} = \mathbf{V}_j \oplus \mathbf{W}_j$ follows

$$det(\mathcal{T} - \lambda id_{\mathbf{V}}) = det(\mathcal{T}|_{\mathbf{V}_j} - \lambda id_{\mathbf{V}_j}) det(\mathcal{T}|_{\mathbf{W}_j} - \lambda id_{\mathbf{W}_j}), ; \text{as polynomials in } \lambda.$$

Then λ_j is an eigenvalue for $\mathcal{T}|_{\mathbf{V}_j}$, of multiplicity m_j, since $\mathcal{T}|_{\mathbf{W}_j} - \lambda_j id_{\mathbf{W}_j}$ is invertible. On the other hand, λ_j is the only eigenvalue of $\mathcal{T}|_{\mathbf{V}_j}$ since 0 is the only eigenvalue of the nilpotent endomorphism $\mathcal{T}_j|_{\mathbf{V}_j}$, by the lemma.

It is clear now that the degree of the polynomial $det(\mathcal{T}|_{\mathbf{V}_j} - \lambda id_{\mathbf{V}_j})$ is m_j, thus $dim\, \mathbf{V}_j = m_j$. Therefore (*i*) and (*iii*) are proved.

(*ii*) is immediate by induction on p, using $\sum\limits_{j=1}^{p} m_j = n$.

We will accept the next theorem without proof, but we note that the missing proof relies on Thm. 4.5.

THEOREM 4.6 (Jordan) *The endomorphism $\mathcal{T}$ admits a Jordan form if and only if its characteristic polynomial has all its n roots (counted with their multiplicities) in* $\mathbf{K}$.

COROLLARY 4.7 *Any endomorphism of a finite dimensional complex vector space (and any complex matrix) admits a Jordan form.*

REMARKS 4.8 We would like to point out in more detail the relationship of the Jordan form J and the decomposition in Theorem 4.5. In the next set of remarks we keep the notation used in Thm. 4.5. Denote also the Jordan basis corresponding to J by $\mathbf{B}_1$, and $d_j = dim\, \mathbf{S}(\lambda_j)$.

(i) The number of Jordan cells having λ_j on the diagonal is d_j, i.e. the maximal number of linearly independent eigenvectors.

(ii) $m_j = dim\, \mathbf{V}_j$= the sum of dimensions of Jordan cells corresponding to λ_j.

(iii) Assuming that the Jordan blocks of J are ordered such that the first d_1 correspond to λ_1, the next d_2 correspond to λ_2, and so on, it follows that the first d_1 vectors of $\mathbf{B}_1$ form a basis of $\mathbf{V}_1$, the next d_2 form a basis of $\mathbf{V}_2$, and so on, up to the last d_p vectors of $\mathbf{B}_1$ which form a basis of $\mathbf{V}_p$.

In practice, if all multiplicities are small enough we may use the following algorithm.

Algorithm for Finding the Jordan Form and the Jordan Basis

1) Find the matrix T of the endomorphism with respect to an arbitrary (fixed) basis.

2) Solve the characteristic equation. If this equation has all n roots in $\mathbf{K}$, then the endomorphism is Jordanizable, otherwise it is not.

3) Compute $dim\, \mathbf{S}(\lambda_j) = n - rank\,(T - \lambda_j I) \leq m_j$. We have already noticed that the number of Jordan blocks corresponding to the eigenvalue λ_j is equal to $dim\, \mathbf{S}(\lambda_j)$. Sometimes, for small values of m_j this fact allows us to figure out the Jordan blocks corresponding to λ_j (see the Remark below).

If $dim\, \mathbf{S}(\lambda_j) = m_j$, then there are m_j Jordan cells of dimension 1, corresponding to λ_j.

4) For each eigenvalue λ_j determine the eigenspace $\mathbf{S}(\lambda_j)$, by solving the linear homogeneous system

$$(Syst.1) \qquad\qquad (T - \lambda_j I)X = 0.$$

Denote the general solution of $(Syst.1)$ by X_1.

5) If $dim\, \mathbf{S}(\lambda_j) < m_j$, some of the eigenvectors admit principal vectors. We impose the compatibility conditions on X_1 and require $X_1 \neq 0$, to solve the system

$$(Syst.2) \qquad\qquad (T - \lambda_j I)X = X_1.$$

Denote general solution of $(Syst.2)$ by X_2. We continue to determine the sequence of principal vectors (actually we determine their coordinates with respect to the basis fixed in 1)) by trying to solve

$$(Syst.3) \qquad\qquad (T - \lambda_j I)X = X_2.$$

This process stops when the compatibility of

$$(Syst.k) \qquad\qquad (T - \lambda_j I)X = X_k$$

leads to a contradiction. The number $k = k(\lambda_j)$ found in this way represents the dimension of the largest Jordan cell having λ_j on the diagonal.

6) Pick particular values for the parameters which appear in $X_1, \ldots, X_k$ to obtain the basis vectors corresponding to the part of the Jordan matrix that has λ_j on the diagonal. These particular values must be chosen such that the compatibility conditions are all satisfied, and the eigenvectors are linearly independent.

REMARK 4.9 Suppose $\lambda_j = \alpha$ is an eigenvalue with $m(\alpha) = m$. Sometimes it is possible to find out the part of the Jordan matrix with diagonal entries α using $dim\,\mathbf{S}(\alpha)$ only, without determining the corresponding part of the basis. For m small, we list below all possible combinations of Jordan blocks with α on the diagonal.

$m = 1 \quad [\alpha] \quad$ (the only possibility) .

$$m = 2 \quad \begin{bmatrix} \alpha & 1 \\ 0 & \alpha \end{bmatrix} \text{ if } dim\,\mathbf{S}(\alpha) = 1; \quad \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} \text{ if } dim\,\mathbf{S}(\alpha) = 2.$$

$$m = 3 \quad \begin{bmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 1 \\ 0 & 0 & \alpha \end{bmatrix} \text{ if } dim\,\mathbf{S}(\alpha) = 1; \quad \begin{bmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{bmatrix} \text{ if } dim\,\mathbf{S}(\alpha) = 2;$$

$$\begin{bmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{bmatrix} \text{ if } dim\,\mathbf{S}(\alpha) = 3.$$

$$m = 4 \quad \begin{bmatrix} \alpha & 1 & 0 & 0 \\ 0 & \alpha & 1 & 0 \\ 0 & 0 & \alpha & 1 \\ 0 & 0 & 0 & \alpha \end{bmatrix} \text{ if } dim\,\mathbf{S}(\alpha) = 1;$$

$$\begin{bmatrix} \alpha & 1 & 0 & 0 \\ 0 & \alpha & 1 & 0 \\ 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & \alpha \end{bmatrix} \text{ or } \begin{bmatrix} \alpha & 1 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ 0 & 0 & \alpha & 1 \\ 0 & 0 & 0 & \alpha \end{bmatrix} \text{ if } dim\,\mathbf{S}(\alpha) = 2;$$

$$\begin{bmatrix} \alpha & 1 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & \alpha \end{bmatrix} \text{ if } dim\,\mathbf{S}(\alpha) = 3; \begin{bmatrix} \alpha & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & \alpha \end{bmatrix} \text{ if } dim\,\mathbf{S}(\alpha) = 4.$$

Note that $dim\,\mathbf{S}(\alpha)$ distinguishes all above cases except for $m = 4$, $dim\,\mathbf{S}(\alpha) = 2$.

Examples (i) Find the Jordan form and the Jordan basis for the endomorphism $\mathcal{T} : \mathbf{R}^4 \to \mathbf{R}^4$,

$$\mathcal{T}(x) = (3x_1 - x_2 - x_3 - 2x_4,\ x_1 + x_2 - x_3 - x_4,\ x_1 - x_4,\ -x_2 + x_3 + x_4).$$

We apply the algorithm.

1) The matrix of $\mathcal{T}$ with respect to the canonical basis is

$$T = \begin{bmatrix} 3 & -1 & -1 & -2 \\ 1 & 1 & -1 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & -1 & 1 & 1 \end{bmatrix}.$$

2) $P(\lambda) = (\lambda - 1)^3(\lambda - 2)$ with roots: $\lambda_1 = 1$, $\lambda_2 = 2$; $m(1) = 3$, $m(2) = 1$. All roots are in $\mathbf{R}$ ($m(1) + m(2) = dim\,\mathbf{R}^4$), thus $\mathcal{T}$ admits a Jordan form.

3) $dim\,\mathbf{S}(1) = 4 - rank\,(T - I) = 2 < m(1)$, so there are 2 Jordan cells with 1 on the diagonal (see the remark above). For the eigenvalue $\lambda_2 = 2$, the multiplicity $m(2) = 1$ forces $dim\,\mathbf{S}(2) = 1$ (Why?). It follows that $\mathcal{T}$ admits a canonical Jordan form

$$J = \begin{bmatrix} 1 & 1 & & \\ 0 & 1 & & \\ & & 1 & \\ & & & 2 \end{bmatrix}.$$

This is the matrix of $\mathcal{T}$ with respect to a Jordan basis $\mathbf{B} = \{e_1, e_2, e_3, e_4\}$, where e_1, e_3 are linearly independent eigenvectors in $\mathbf{S}(1)$, e_2 is a principal vector for e_1, and e_4 is an eigenvector in $\mathbf{S}(2)$.

We will apply the rest of the algorithm for λ_1, then for λ_2.

For $\lambda_1 = 1$:

4) Solve

$$(Syst.1) \qquad (T - I)X = 0; \quad X_1 =^t [\alpha \;\; \alpha - \beta \;\; \alpha - \beta \;\; \beta], \quad \alpha, \beta \in \mathbf{R}.$$

e_1 and e_3 are of the form X_1 with α and β chosen for each of them, such that e_1, e_3 are linearly independent, and e_1 admits a principal vector.

5) The system

$$(Syst.2) \qquad (T - I)X =^t [\alpha \;\; \alpha - \beta \;\; \alpha - \beta \;\; \beta]$$

requires the compatibility condition $\alpha - \beta = 0$, so $e_1 = (\alpha, 0, 0, \alpha)$, $\alpha \neq 0$.

For $\alpha = \beta$, the general solution of $(Syst.2)$ is

$$X_2 =^t [\alpha + \gamma + \delta \;\; \gamma \;\; \alpha + \gamma \;\; \delta], \quad \gamma, \delta \in \mathbf{R}.$$

We know from 3) that there is no need to look for more principal vectors.

6) For e_1, choose $\alpha = 1 = \beta$ in X_1; $e_1 = (1, 0, 0, 1)$. For e_2, $\alpha = 1$ (like in e_1), and pick $\gamma = -1$, $\delta = 0$, which give $e_2 = (0, -1, 0, 0)$.

If in X_1 we take $\alpha = 1$, $\beta = 0$ for e_3, we obtain $e_3 = (1, 1, 1, 0)$.

For $\lambda_2 = 2$:

4) The system $(T - 2I)X = 0$ has the general solution ${}^t[2a \;\; a \;\; a \;\; 0]$ $\; a \in \mathbf{R}$. Any nontrivial solution does the job for e_4; let us take $a = 1$. Then $e_4 = (2, 1, 1, 0)$.

The problem is now completely solved. The Jordanizing matrix is

$$C = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

We may check that $J = C^{-1}TC$, or equivalently, $CJ = TC$.

Note that we determined the Jordan form J immediately after we found out $dim\,\mathbf{S}(1)$; then it was clear that an eighenvector associated to the eigenvalue $\lambda_1 = 1$

admits either one principal vector, or none, so we stoped looking for principal vectors after solving $(Syst.2)$. Let us see what would happen if we tried to determine one more principal vector, namely to solve

$$(T - I) =^t [\alpha + \gamma + \delta \;\; \gamma \;\; \alpha + \gamma \;\; \delta].$$

The compatibility condition for this system is $\alpha = 0$, that contradicts $e_1 \neq 0$.

(ii) Determine the Jordan form and the Jordanizing matrix of

$$A = \begin{bmatrix} 4 & -1 & -1 & 1 \\ 1 & 2 & 0 & 1 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 1 & 3 \end{bmatrix} \in M_{4,4}(\mathbf{R}).$$

We find $P(\lambda) = (\lambda - 3)^4$, $\lambda_1 = 3$, $m(3) = 4 = dim\,\mathbf{R}^4$, thus A is Jordanizable. $dim\,\mathbf{S}(3) = 4 - rank\,(A - 3I) = 2$, thus the Jordan form of A is one of the following:

$$\tilde{J} = \begin{bmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix} \quad \text{or} \quad \tilde{\tilde{J}} = \begin{bmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

We will find out which of $\tilde{J}$ or $\tilde{\tilde{J}}$ is the correct one, while looking for the Jordan basis. The system

$$(Syst.1) \qquad (T - I)X = 0$$

has the general solution $X_1 =^t [\alpha \quad \beta \quad 0 \quad -\alpha + \beta]$, $\quad \alpha, \beta \in \mathbf{R}$.

The next system,

$$(Syst.2) \qquad (T - 3I)X =^t [\alpha \quad \beta \quad 0 \quad -\alpha + \beta]$$

is compatible for $\forall\, \alpha, \beta \in \mathbf{R}$; its general solution is $X_2 =^t [a \quad b \quad -\alpha + \beta \quad \beta - a + b]$.

Following the algorithm, we try to solve the system:

$$(Syst.3) \qquad (T - 3I)X =^t [a \quad b \quad -\alpha + \beta \quad \beta - a + b],$$

which is compatible if and only if $\alpha = \beta = 0$. But this contradicts $e_1 \neq 0$.

Therefore no eigenvector admits sequences of 2 principal vectors.

Consequently, the Jordan form is $J = \tilde{J}$. Each eigenvector e_1, e_3 of the Jordan basis $\{e_1, e_2, e_3, e_4\}$ is followed by one principal vector.

We may pick any linearly independent e_1, $e_3 \in \mathbf{S}(3)$, and for each of them we may take any values of a and b to obtain e_2 and e_4. For $\alpha = 1$, $\beta = 0$, $a = b = 0$, we get $e_1 = (1, 0, 0, -1)$, $e_2 = (0, 0, -1, 0)$, and for $\alpha = 0$, $\beta = 1$, $a = b = 0$, we get $e_3 = (0, 1, 0, 1)$, $e_4 = (0, 0, 1, 1)$.

The Jordanizing matrix $C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ -1 & 0 & 1 & 1 \end{bmatrix}$ and the Jordan form

$J = \begin{bmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix}$ will certainly satisfy $J = C^{-1}AC$.

5 The Spectrum of Endomorphisms on Euclidean Spaces

V is an Euclidean **K**-vector space, $\mathbf{K} = \mathbf{R}$ or $\mathbf{K} = \mathbf{C}$, $\mathcal{T} \in \mathrm{End}(\mathbf{V})$. Recall that the spectrum of $\mathcal{T}$ is

$$\mathrm{Spec}(\mathcal{T}) = \{\lambda \in \mathbf{K} \,|\, \exists v \neq 0,\ \mathcal{T}v = \lambda v\} = \text{the set of all eigenvalues}\,.$$

REMARK 5.1 If λ is an eigenvalue of $\mathcal{T}$ and v a corresponding eigenvector, then

$$\lambda = \frac{\langle \mathcal{T}v, v\rangle}{\langle v, v\rangle}\,.$$

THEOREM 5.2 *If* $\mathbf{K} = \mathbf{C}$ *and* $\mathcal{T}$ *is Hermitian, then*

(*i*) $\mathrm{Spec}(\mathcal{T}) \subset \mathbf{R}$

(*ii*) *Eigenvectors corresponding to distinct eigenvalues are orthogonal.*

(*iii*) *If* $\dim \mathbf{V} = n \in \mathbf{N}^*$*, then there exists an orthonormal basis of* **V***, made up of eigenvectors. Thus a Hermitian endomorphism is always diagonalizable.*

Proof. $\langle x, \mathcal{T}y\rangle = \langle \mathcal{T}x, y\rangle$, $\forall\, x, y \in \mathbf{V}$, by hypothesis.
(*i*) Assume λ is an eigenvalue and v a corresponding eigenvector. Then

$$\lambda = \frac{\langle \mathcal{T}v, v\rangle}{\langle v, v\rangle} = \frac{\langle v, \mathcal{T}v\rangle}{\langle v, v\rangle} = \frac{\overline{\langle \mathcal{T}v, v\rangle}}{\langle v, v\rangle} = \overline{\lambda}.$$

Thus $\lambda \in \mathbf{R}$.

(*ii*) Let $\lambda_1 \neq \lambda_2$ be eigenvalues with corresponding eigenvectors v_1 and v_2 respectively. From

$$\lambda_1\langle v_1, v_2\rangle = \langle \mathcal{T}v_1, v_2\rangle = \langle v_1, \mathcal{T}v_2\rangle = \overline{\lambda_2}\langle v_1, v_2\rangle = \lambda_2\langle v_1, v_2\rangle,$$

follows $(\lambda_1 - \lambda_2)\langle v_1, v_2\rangle = 0$, i.e. $\langle v_1, v_2\rangle = 0$, as $\lambda_1 \neq \lambda 2$.

(*iii*) will be proved by induction on n. For $n = 1$ it is trivially true. Let $n \leq 2$ and assume (*iii*) is true for vector spaces of dimension $n - 1$. From (*i*) follows that $\mathcal{T}$ has at least one real eigenvalue λ_1. Let v_1 be an eigenvector associated to λ_1 and denote $U = \{v_1\}^\perp$. We know from Chapter 1 that U is a subspace of **V**; moreover, $\mathbf{V} = \mathit{Span}\,\{v_1\} \oplus U$, thus $\mathit{dim}\, U = n - 1$.

In order to apply the inductive hypothesis for U we need to show that U is $\mathcal{T}$-invariant. For, let $x \in U$; then $\langle \mathcal{T}x, v_1\rangle = \langle x, \mathcal{T}v_1\rangle = \lambda_1\langle x, v_1\rangle$, thus $\mathcal{T}x \in U$. The restriction $\mathcal{T}|_U : U \to U$ is Hermitian too, so by the inductive hypothesis there exists an orthonormal basis $\{u_2, \ldots, u_n\}$ of U, made up of eigenvectors. Then $\{\frac{1}{\|v_1\|}v_1, u_2, \ldots, u_n\}$ is the required basis of **V**. QED.

Translating the theorem in matrix form we obtain

COROLLARY 5.3 *Let* $A \in M_{n,n}(\mathbf{C})$*, A Hermitian. Then* (*i*) *A is diagonalizable.*

(*ii*) *If D is a diagonal form of A, then $D \in M_{n,n}(\mathbf{R})$.*

(*iii*) *There exists a diagonalizing unitary matrix (in $M_{n,n}(\mathbf{C})$.)*

REMARK 5.4 Let $A \in M_n(\mathbf{R})$, A symmetric. Since real symmetric matrices are particular cases of Hermitian matrices, the previous corollary applies, thus all eigenvalues of A are real and A is diagonalizable; there exists a diagonalizing orthogonal matrix $C \in M_{n,n}(\mathbf{R})$. In particular. Theorem 2.4 comes out as corollary of (*i*).

COROLLARY 5.5 *If $\mathbf{K} = \mathbf{R}$ and $\mathcal{T}$ is symmetric, then*

(*i*) *Assuming $\mathbf{V}$ finite dimensional, all roots of the characteristic polynomial of $\mathcal{T}$ are real.*

(*ii*) *Eigenvectors corresponding to distinct eigenvalues are orthogonal.*

(*iii*) *If $\dim(\mathbf{V}) = n \in \mathbf{N}^*$, then there exists an orthonormal basis of $\mathbf{V}$, made up of eigenvectors. Thus a symmetric endomorphism is always diagonalizable.*

Similarly we can deduce some properties of eigenvalues and eigenvectors of skew Hermitian and skew symmetric endomorphisms.

REMARK 5.6 (*i*) The eigenvalues of a skew Hermitian endomorphism are purely imaginary or zero. Assuming $\mathbf{V}$ finite dimensional over $\mathbf{R}$, it follows that all roots of the characteristic polynomial are purely imaginary or zero.

(*ii*) Eigenvectors corresponding to distinct eigenvalues of a skew Hermitian endomorphism are orthogonal.

(*iii*) Note that skew symmetric endomorphisms are not necessarily diagonalizable.

THEOREM 5.7 *Let $\mathbf{V}$ be a complex (real) Euclidian space and $\mathcal{T} \in \mathrm{End}(\mathbf{V})$ an unitary (orthogonal) endomorphism.*

(*i*) *If there exist any eigenvalues, their absolute value is 1.*

(*ii*) *Eigenvectors corresponding to distinct eigenvalues are orthogonal.*

(*iii*) *If $\mathbf{K} = \mathbf{C}$, $\mathbf{V}$ is finite dimensional, and $\mathcal{T}$ is unitary, then $\mathcal{T}$ is diagonalizable and $\mathbf{V}$ admits an orthonormal basis made up of eigenvalues.*

(*iv*) *If $\mathbf{K} = \mathbf{R}$, $\mathbf{V}$ is finite dimensional, $\mathcal{T}$ is orthogonal and the characteristic polynomial has no other roots in $\mathbf{C}$ besides 1 or -1, then $\mathcal{T}$ is diagonalizable and $\mathbf{V}$ admits an orthonormal basis made up of eigenvectors.*

6 Polynomials and Series of Matrices

$\mathbf{V}$ will denote a $\mathbf{K}$–vector space, $dim\,\mathbf{V} = n \in \mathbf{N}^\star$, and $\mathcal{A} : \mathbf{V} \to \mathbf{V}$ an endomorphism. The matrix of $\mathcal{A}$ w.r.t. a fixed basis of $\mathbf{V}$ is $A = [a_{ij}] \in M_{n,n}(\mathbf{K})$.

DEFINITION 6.1 Consider the polynomial $Q(t) \in \mathbf{K}[t]$, $Q(t) = a_m t^m + a_{m-1}t^{m-1} + \ldots + a_1 t + a_0$. The matrix

$$Q(A) = a_m A^m + a_{m-1}A^{m-1} + \ldots + a_1 A + a_0 I$$

is called *the value of Q for the matrix A*. The endomorphism

$$Q(\mathcal{A}) = a_m \mathcal{A}^m + a_{m-1}\mathcal{A}^{m-1} + \ldots + a_1 \mathcal{A} + a_0 id_V$$

is called *the value of Q for $\mathcal{A}$*.

$Q(A)$ is also called *polynomial of matrix* or *matrix polynomial*; $Q(\mathcal{A})$ is called *polynomial of endomorphism.*

Since endomorphisms (of finitely dimensional vector spaces) are completely determined by their associated matrices, it suffices to study matrix polynomials. All results about matrices may be reformulated in terms of endomorphisms.

REMARK 6.2 If the matrix A admits the canonical Jordan form J, then the following observations make the computation of $Q(A)$ easier.

$$A = CJC^{-1}, \quad A^2 = CJ^2C^{-1}, \quad \ldots, A^m = CJ^mC^{-1}.$$

If A admits the diagonal form, i.e. $J = D$ is a diagonal matrix, then the diagonal form D of A is even more convenient to use for computing powers of A.

$$A = CDC{-}1, \quad A^2 = CD^2C^{-1}, \quad \ldots, A^m = CD^mC^{-1},$$

since $D = diag(d_1, \ldots, d_n) \Rightarrow D^k = diag(d_1^k, \ldots, d_n^k)$, $\forall k \in \mathbf{N}$.

THEOREM 6.3 (Cayley–Hamilton) *If $P(\lambda)$ is the characteristic polynomial of the matrix A, then $P(A) = 0$*

Proof. Recall that for any matrix $M = m_{ij} \in M_{n,n}(\mathbf{K})$, there exists a matrix M^+, called *the reciprocal of M* such that

$$M \cdot M^+ = (det\,M)I.$$

(Some books use the notation M^* instead of M^+ and call it the adjoint of M instead of the reciprocal; here *reciprocal* and *adjoint* have different meanings.)

The (i,j) entry of M^+ is the algebraic complement of m_{ji} in M. For $M = A - \lambda I$, we obtain

$$(A - \lambda I)(A - \lambda I)^+ = P(\lambda)I, \tag{6.1}$$

since $P(\lambda) = det(A - \lambda I)$. Moreover, by the construction of the reciprocal matrix follows that each entry of $(A - \lambda I)^+$ is a polynomial in λ of degree at most $n-1$; consequently, we can express

$$(A - \lambda I)^+ = B_{n-1}\lambda^{n-1} + B_{n-2}\lambda^{n-2} + \ldots + B_0, \tag{6.2}$$

where $B_i \in M_{n,n}(\mathbf{K}),\ i = 0, \ldots, n-1$. Let

$$P(\lambda) = \sum_{k=0}^{n} a_k \lambda^k,\ a_k \in \mathbf{K}. \tag{6.3}$$

Replacing (6.2), (6.3) in (6.1) we obtain

$$(A-\lambda I)(B_{n-1}\lambda^{n-1}+B^{n-2}\lambda^{n-2}+\ldots+B_1\lambda+B_0) = (a_{n-1}\lambda^{n-1}+a^{n-2}\lambda^{n-2}+\ldots+a_1\lambda+a_0)I$$

or

$$(B_{n-1})\lambda^n+(AB_{n-1}-B_{n-2})\lambda^{n-1}+\ldots+(AB_1-B_0)\lambda+AB_0 = (a_nI)\lambda^n+\ldots+(a_1I)\lambda+a_0I.$$

Now we identify the coefficients

$$-B_{n-1} = a_nI,\ AB_{n-1} - B_{n-2} = a_{n-1}I,\ \ldots,\ AB_1 - B_0 = a_1I,\ AB_0 = a_0I.$$

Multiplying on the left by A_n, A_{n-1}, ..., A, I and adding up each side we get

$$\begin{array}{rl} P(A) = & a_nA^n + a_{n-1}A^{n-1} + \ldots + a_1A + a_0I \\ = & \qquad\qquad -A^nB_{n-1}+ \\ & +A^nB_{n-1} \ \ -A^{n-1}B_{n-2}+ \\ & +A^{n-1}B_{n-2} - A^{n-2}B_{n-3}+ \\ & +\ldots\ldots\ldots\ldots\ldots\ldots+ \\ & +A^2B_1 - AB_0+ \\ & +AB_0 \qquad\qquad = \\ = & 0 \end{array} \quad . \qquad QED.$$

COROLLARY 6.4 *If* $dim\,\mathbf{V} = n \in \mathbf{N}^\star$, $\mathcal{A} : \mathbf{V} \to \mathbf{V}$ *is an endomorphism, and* $P(\lambda)$ *the characteristic polynomial of* $\mathcal{A}$*, then* $P(\mathcal{A}) = 0$.

THEOREM 6.5 *If* $A \in M_{n,n}(\mathbf{K})$*, then the value at* A *of any polynomial of degree at least* n *can be expressed as a matrix polynomial in* A*, of degree at most* $n-1$.

Proof. Let $P(\lambda)$ be the characteristic polynomial of A. By the Cayley-Hamilton Theorem, $P(A) = 0$. Since the degree of P is n, the coefficient of A^n in $P(A)$ is not zero, thus we can express

$$A^n = b_{n-1}A^{n-1} + \ldots + b_0I$$

This proves the theorem in the case of polynomials of degree n. If we multiply the equality by A and replace A^n, we get A^{n+1} as a linear combination of $A^{n-1}, \ldots, A, I$. Inductively, the same holds for A^{n+p}, $\forall p \in \mathbf{N}$. QED.

Assume now $\mathbf{K} = \mathbf{C}$ or $\mathbf{R}$ and consider the power series $f(t) = \sum\limits_m a_m t^m$ with coefficients in $\mathbf{K}$. Such a series makes sense for $t \in \mathbf{V}$, when $\mathbf{V}$ is endowed with multiplication, such that the powers t^m are well defined. If $\mathbf{V}$ is a Euclidean vector space, we consider the Euclidean metric on $\mathbf{V}$, and we can discuss about the convergence of the series. In particular, we are interested in $\mathbf{V} = M_{n,n}(\mathbf{K}) \simeq \mathbf{K}^{n^2}$ with the Euclidean norm

$$\|A\| = \sqrt{\sum_{1\le i,j\le n} |a_{ij}|}.$$

DEFINITION 6.6 Let $\mathcal{A} : \mathbf{V} \to \mathbf{V}$ be an endomorphism, $dim\,\mathbf{V} = n \in \mathbf{N}^{\star}$and A the associated basis w.r.t. a fixed basis. We define a *series of matrices* and a *series of endomorphisms* by

$$\sum_m a_m A^m, \text{ and} \sum_m a_m \mathcal{A}^m \text{ respectively.}$$

If the series are convergent, their sums are denoted by $f(A)$, and $f(\mathcal{A})$ respectively, where f is the function defined by the convergent series $f(t) = \sum\limits_m a_m t^m$, $t \in \mathbf{K}$. $f(A)$ is called *function of matrix* or *matrix function*, and $f(\mathcal{A})$ is called *function of endomorphism.*

On finite dimensional vector spaces, the study of series of endomorphisms reduces to the study of matrix series. On the other hand, as a consequence of theorem Cayley–Hamilton, $\sum\limits_m a_m A^m$ can be expressed as a matrix polynomial $Q(A)$, of degree $n-1$ whose coefficients are numerical series. If $\sum\limits_m a_m A^m$ is convergent, then the coefficients of $Q(A)$ are convergent numerical series.

If A admits distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, then the polynomial of degree $n-1$ associated to $\sum\limits_m a_m A^m$ can be written in the Lagrange form

$$f(A) = \sum_{j=1}^{n} \frac{(A-\lambda_1 I)\ldots(A-\lambda_{j-1}I)(A-\lambda_{j+1}I)\ldots(A-\lambda_n I)}{(\lambda_j-\lambda_1)\ldots(\lambda_j-\lambda_{j-1})(\lambda_j-\lambda_{j+1})\ldots(\lambda_j-\lambda_n)} f(\lambda_j),$$

or

$$f(A) = \sum_{j=1}^{n} Z_j f(\lambda_j),$$

where $Z_j \in M_{n,n}(\mathbf{K})$ are matrices independent of the function f.

If A has multiple eigenvalues, a more general formula works

$$f(A) = \sum_{k=1}^{p} \sum_{j=0}^{m_k-1} Z_{kj} f^{(j)}(\lambda_k),$$

where $f^{(j)}$ denotes the derivative of order j of the funtion f, and Z_{kj} are matrices independent of f.

The following series are convergent for any matrix A.

$$e^A = \sum_{m=0}^{\infty} \frac{1}{m!} A^m, \quad sin\,A = \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m+1)!} A^{2m+1}, \quad cos\,A = \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m)!} A^{2m}.$$

We will also consider the series

$$e^{tA} = \sum_{m=0}^{\infty} \frac{t^m}{m!} A^m, \quad \text{where } t \in \mathbf{K}.$$

This series is very useful in the theory of linear differential with constant coefficients.

7 Problems

1. Determine the eigenvalues and the eigenvectors of the endomorphisms in problems 4 - 7, Chapter 2.

2. Let $\mathbf{V}$ be the real Euclidean vector space of the continuous real functions on $[0, 2\pi]$. Let $\mathcal{T} : \mathbf{V} \to \mathbf{V}$ be the endomorphism defined by

$$g = \mathcal{T}(f),\ g(x) = \int_0^{2\pi} [1 + \cos(x - t)] \cdot f(t)dt,\ x \in [0, 2\pi].$$

1) Show that the subspace $Im(\mathcal{T})$ is finite dimensional and find an orthogonal basis of this subspace of $\mathbf{V}$.
2) Determine $Ker(\mathcal{T})$.
3) Show that $\mathcal{T}$ is symmetric; find its eigenvalues and eigenvectors.

3. Determine the diagonal form and the corresponding basis, then compute $e^{\mathcal{A}}$.

1) $\mathcal{A} : \mathbf{R}^3 \to \mathbf{R}^3,\ A = \begin{bmatrix} 7 & 4 & -1 \\ 4 & 7 & -1 \\ -4 & 4 & 4 \end{bmatrix}$ 2) $\mathcal{A} : \mathbf{R}^3 \to \mathbf{R}^3,\ A = \begin{bmatrix} 4 & 6 & 0 \\ -3 & -5 & 0 \\ -3 & -6 & 1 \end{bmatrix}$

3) $\mathcal{A} : \mathbf{R}^4 \to \mathbf{R}^4,\ A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 \\ 1 & 0 & -2 & 5 \end{bmatrix}$

4. Determine the canonical forms (diagonal or Jordan) of the following endomorphisms, and the corresponding bases.

1) $\mathcal{A} : \mathbf{R}^3 \to \mathbf{R}^3,\ A = \begin{bmatrix} -1 & 0 & -3 \\ 3 & 2 & 3 \\ -3 & 0 & -1 \end{bmatrix}$ 2) $\mathcal{A} : \mathbf{R}^3 \to \mathbf{R}^3,\ A = \begin{bmatrix} 6 & 6 & -15 \\ 1 & 5 & -5 \\ 1 & 2 & -2 \end{bmatrix}$

3) $\mathcal{A} : \mathbf{R}^2 \to \mathbf{R}^2,\ A = \begin{bmatrix} 2 & -1 \\ -1 & 4 \end{bmatrix}$ 4) $\mathcal{A} : \mathbf{R}^3 \to \mathbf{R}^3,\ A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{bmatrix}$

5) $\mathcal{A} : \mathbf{R}^4 \to \mathbf{R}^4,\ A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 3 & 1 & 1 \\ 0 & 0 & 0 & -1 \\ -1 & -1 & 0 & 2 \end{bmatrix}$

5. For each matrix A in the previous problem compute A^n, e^A, $\sin A$, $\cos A$ using the Cayley-Hamilton theorem.

6. Use the Cayley - Hamilton theorem to determine A^{-1} and the value of the matrix polynomial $Q(A) = A^4 + 3A^3 - 9A^2 - 28A$, where

$$A = \begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}.$$

Chapter 4

Bilinear Forms. Quadratic Forms

1 Bilinear Forms

Let $\mathbf{V}$ be $\mathbf{K}$-vector space.
The basic example for bilinear forms is the canonical inner product on $\mathbf{R}^n$. On the other hand, the notion of bilinear form extends the one of linear form. Recall that a linear form on $\mathbf{V}$ is a linear map $\omega : \mathbf{V} \to \mathbf{K}$.

DEFINITION 1.1 A map $\mathcal{A} : \mathbf{V} \times \mathbf{V} \to \mathbf{K}$ is called *a bilinear form on* $\mathbf{V}$ if it is linear in each argument i.e.

$$\begin{aligned} \mathcal{A}(kx + ly, z) &= k\mathcal{A}(x, z) + l\mathcal{A}(y, z), \\ \mathcal{A}(x, ky + lz) &= k\mathcal{A}(x, y) + l\mathcal{A}(x, z), \quad \forall x, y, z \in \mathbf{V},\ \forall\, k, l \in \mathbf{K}. \end{aligned}$$

Example Any inner product on a real vector space is a bilinear form.

Counterexample An inner product on a complex vector space is not a bilinear form, since it is not $\mathbf{C}$-linear in the second argument.

Denote by $\mathcal{B}(\mathbf{V}, \mathbf{K})$ the set of all bilinear forms on $\mathbf{V}$. Addition and multiplication of bilinear forms are defined in the usual way for $\mathbf{K}$-valued functions. These operations define a $\mathbf{K}$-vector space structure on $\mathcal{B}(\mathbf{V}, \mathbf{K})$.

REMARK 1.2 Assume that $\mathbf{V}$ is finite dimensional, $\dim \mathbf{V} = n \in \mathbf{N}^\star$, and let $\mathbf{B} = \{e_1, \ldots, e_n\}$ be a basis of $\mathbf{V}$. Let $x, y \in \mathbf{V}$, $x = \sum_{i=1}^{n} x_i e_i$, $y = \sum_{j=1}^{n} y_j e_j$. If $\mathcal{A} \in \mathcal{B}(\mathbf{V}, \mathbf{K})$, then

$$\mathcal{A}(x, y) = \mathcal{A}\Big(\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j\Big) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j \mathcal{A}(e_i, e_j),$$

which shows that $\mathcal{A}$ is uniquely determined by its values on $\mathbf{B} \times \mathbf{B}$.

Using the notation $a_{ij} = \mathcal{A}(e_i, e_j)$, $A = [a_{ij}] \in M_{n,n}(\mathbf{K})$, we may describe the bilinear form $\mathcal{A}$ by the one of the following equalities

$$\mathcal{A}(x,y) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_iy_j \qquad \textit{(the analytic expression of } \mathcal{A}\textit{)}$$

or

$$\mathcal{A}(x,y) = {}^tXAY \qquad \textit{(the matrix expression of } \mathcal{A}\textit{)},$$

where X, Y represent as usually the coordinate columns of x and y respectively.

A is called *the matrix associated to $\mathcal{A}$ w.r.t. the basis* $\mathbf{B}$. It is straightforward that the correspondence between a bilinear form and its associated matrix w.r.t. a fixed basis defines an isomophism from $\mathcal{B}(\mathbf{V}, \mathbf{K})$ onto $M_{n,n}(\mathbf{K})$. It follows that $dim\,\mathcal{B}(\mathbf{V}, \mathbf{K}) = n^2$.

Example Consider the cannonical inner product on $\mathbf{R}^n$,

$$\langle x, y\rangle = x_1y_1 + \ldots + x_ny_n.$$

If $\mathbf{B}$ is the canonical basis, then the associated matrix is the $n \times n$ identity matrix I_n, since $\langle e_i, e_j\rangle = \delta_{ij}$.

DEFINITION 1.3 The bilinear form $\mathcal{A}$ is said to be *symmetric* if

$$\mathcal{A}(x,y) = \mathcal{A}(y,x),\ \forall x, y \in \mathbf{V}.$$

The bilinear form $\mathcal{A}$ is said to be *skew symmetric* if

$$\mathcal{A}(x,y) = -\mathcal{A}(y,x),\ \forall x, y \in \mathbf{V}.$$

From the definition of the inner product on a real vector space follows that inner products on real vector spaces are symmetric bilinear forms.

THEOREM 1.4 *Let $\mathcal{A} : \mathbf{V} \times \mathbf{V} \to \mathbf{K}$ be a bilinear form and $dim\,\mathbf{V} = n \in \mathbf{N}^\star$. Then $\mathcal{A}$ is symmetric (skew symmetric) if and only if the associated matrix w.r.t. a fixed basis is symmetric (skew symmetric).*

Proof. Fix a basis $\mathbf{B} = \{e_1, \ldots, e_n\}$ in $\mathbf{V}$. Denote $A = [a_{ij}]$ the associated matrix.

Assume $\mathcal{A}$ symmetric. Then $a_{ij} = \mathcal{A}(e_i, e_j) = \mathcal{A}(e_j, e_i) = a_{ji}$, $\forall i, j = 1, \ldots, n$, thus $A = {}^tA$.

Conversely, assume $A = {}^tA$. Then

$$\mathcal{A}(y,x) = {}^tYAX = {}^t({}^tYAX) = {}^tX\,{}^tAY = {}^tXAY = \mathcal{A}(x,y),\ \forall x, y \in \mathbf{V}.$$

The skew symmetric case is similar. QED.

THEOREM 1.5 *Let $\mathcal{A}$ be a bilinear form on $\mathbf{V}$, $dim\,\mathbf{V} = n \in \mathbf{N}^\star$, $\mathbf{B}$ and $\mathbf{B}'$ two bases in $\mathbf{V}$ with $C \in M_{n,n}(\mathbf{K})$-the matrix of change from $\mathbf{B}$ to $\mathbf{B}'$.*

If A and $B \in M_{n,n}(\mathbf{K})$ are the matrices associated to $\mathcal{A}$ w.r.t. $\mathbf{B}$ and $\mathbf{B}'$ respectively, then

$$B = {}^tCAC.$$

Proof. Let $x, y \in \mathbf{V}$ arbitrary, and X, Y the coordinate columns of x, respectively y w.r.t. $\mathbf{B}$. Denote by X', Y' the coordinate columns of the same x, y w.r.t. $\mathbf{B}'$. The coordinate columns are related by $X = CX'$, $Y = CY'$. Starting with the matrix form of $\mathcal{A}$ w.r.t. $\mathbf{B}$ we obtain

$$\mathcal{A}(x,y) = {}^tXAY = {}^t(CX')A(CY') = {}^t(X')({}^tCAC)(Y').$$

On the other hand, the matrix form of $\mathcal{A}$ w.r.t. $\mathbf{B}'$ is $\mathcal{A}(x,y) = {}^t(X')B(Y')$, and B is uniquely determined by this equality; therefore $B = {}^tCAC$. QED.

DEFINITION 1.6 If A is nonsingular (singular), then the bilinear form $\mathcal{A}$ is called *nondegenerate (degenerate).* The rank of A is also called *the rank of the bilinear form* $\mathcal{A}$.

The rank of the associated matrix is the same regardless the choice of basis (see Corollary 1.8), so the above definition makes sense.

DEFINITION 1.7 Let $\mathcal{A} : \mathbf{V} \times \mathbf{V} \to \mathbf{K}$ be a symmetric bilinear form. The set

$$Ker\,\mathcal{A} = \{x \in \mathbf{V} |\, \mathcal{A}(x,y) = 0,\ \forall y \in \mathbf{V}\}$$

is called *the kernel of the form* $\mathcal{A}$.

PROPOSITION 1.8 *$Ker\,\mathcal{A}$ is a vector subspace of* $\mathbf{V}$.

Proof. Let $u, v \in Ker\,\mathcal{A}$ i.e. $\mathcal{A}(u,y) = 0$, $\mathcal{A}(v,y) = 0\, \forall y \in \mathbf{V}$. If $k, l \in \mathbf{K}$, then

$$\mathcal{A}(ku + lv, y) = k\mathcal{A}(u,y) + l\mathcal{A}(v,y) = 0,$$

thus $ku + lv \in Ker\,\mathcal{A}$. QED.

THEOREM 1.9 *If $\mathcal{A} : \mathbf{V} \times \mathbf{V} \to \mathbf{K}$ is a symmetric bilinear form, and $dim\,\mathbf{V} = n \in \mathbf{N}^\star$, then*

$$rank\,\mathcal{A} = n - dim\,Ker\mathcal{A}.$$

Proof. Let $\{e_1, \ldots, e_n\}$ be a basis of $\mathbf{V}$ and $A = [a_{ij}]$, $a_{ij} \in \mathbf{K}$ the matrix associated to $\mathcal{A}$ w.r.t. this basis. If $x, y \in \mathbf{V}$ are arbitrary, $x = \sum\limits_{i=1}^{n} x_i e_i$, $y = \sum\limits_{j=1}^{n} y_j e_j$,

$$\mathcal{A}(x,y) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i y_j = \sum_{j=1}^{n}\Big(\sum_{i=1}^{n} a_{ij}x_i\Big)y_j$$

which shows that $x \in Ker\,\mathcal{A}$ if and only if x is a solution of the linear homogeneous system

$$\sum_{i=1}^{n} a_{ij}x_i = 0\ j = 1, \ldots, n.$$

So $Ker\,\mathcal{A}$ is isomorphic to the solution space of this system, whose dimension is $n - rank\,A = n - rank\,\mathcal{A}$. QED.

COROLLARY 1.10 (i) *$rank\,\mathcal{A}$ is independent of the choice of basis; it depends only on the symmetric bilinear form $\mathcal{A}$.*

(ii) *$\mathcal{A}$ is nondegenerate $\Leftrightarrow$ $Ker\,\mathcal{A} = \{0\}$.*

2 Quadratic Forms

Throughout the section $\mathbf{V}$ denotes a $\mathbf{K}$-vector space and $\mathcal{A}$ a symmetric bilinear form on $\mathbf{V}$.

DEFINITION 2.1 A map $Q : \mathbf{V} \to \mathbf{K}$ is called *a quadratic form on* $\mathbf{V}$ if there exists a symmetric bilinear form $\mathcal{A}$ such that $Q(x) = \mathcal{A}(x,x)$, $\forall x \in \mathbf{V}$.

REMARK 2.2 In the definition of the quadratic form, Q and $\mathcal{A}$ are uniquely determined by each other. Obviously Q is determined by $\mathcal{A}$. For the converse, the following computation gives a formula for $\mathcal{A}$ in terms of Q.

$$Q(x+y) = \mathcal{A}(x+y, x+y) = \mathcal{A}(x,x) + \mathcal{A}(y,y) + \mathcal{A}(x,y) + \mathcal{A}(y,x).$$

From the symmetry of $\mathcal{A}$,

$$Q(x+y) = \mathcal{A}(x,x) + \mathcal{A}(y,y) + 2\mathcal{A}(x,y), \text{ therefore}$$

$$\mathcal{A}(x,y) = \frac{1}{2}(Q(x+y) - Q(x) - Q(y)).$$

The symmetric bilinear form $\mathcal{A}$ associated to Q is called *the polar of* Q.

Example The quadratic form corresponding to the inner product on a real vector space is the square of the Euclidean norm

$$Q(x) = \langle x, x \rangle = \|x\|^2,\ x \in \mathbf{V}.$$

Assume $dim\, \mathbf{V} = n \in \mathbf{N}^\star$; let $\{e_1, \ldots, e_n\}$ be a basis of $\mathbf{V}$ and $A = [a_{ij}]$ the matrix of $\mathcal{A}$ w.r.t. this basis. For any $x \in \mathbf{V}$, $x = \sum\limits_{i=1}^{n} x_i e_i$,

$$Q(x) = \mathcal{A}(x,x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j = {}^tXAX,$$

where $X = {}^t[x_1, \ldots, x_n]$. It follows that the matrix A characterizes Q as well as $\mathcal{A}$. The symmetric matrix A is also called *the matrix of the quadratic form* Q; by definition, *the rank of* Q is the rank of the matrix A, so $rank\, Q = rank\, A = rank\, \mathcal{A}$.

DEFINITION 2.3 Let $\mathcal{A} : \mathbf{V} \times \mathbf{V} \to \mathbf{K}$ be a symmetric bilinear form and Q the associated quadratic form. The vectors $x, y \in \mathbf{V}$ are said to be *orthogonal* w.r.t. $\mathcal{A}$ (or Q) if $\mathcal{A}(x,y) = 0$.

DEFINITION 2.4 Let $\mathbf{U} \subset \mathbf{V}$, a vector subspace of $\mathbf{V}$. The vector space

$$\mathbf{U}^\perp = \{y \in \mathbf{V} \,|\, \mathcal{A}(x,y) = 0,\ \forall x \in \mathbf{V}\}$$

is called the *orthogonal complement* of $\mathbf{U}$ in $\mathbf{V}$ w.r.t. $\mathcal{A}$.

THEOREM 2.5 *Let $\mathcal{A} : \mathbf{V} \times \mathbf{V} \to \mathbf{K}$ be a symmetric bilinear form. Denote by $\mathcal{A}|_{\mathbf{U}}$ the restriction of $\mathcal{A}$ to $\mathbf{U} \times \mathbf{U}$.*

(*i*) *$\mathbf{U}^{\perp}$ is a vector subspace of $\mathbf{V}$.*

(*ii*) *If $\{u_1, u_2, \ldots, u_p\}$ is a basis in $\mathbf{U}$, then*

$$y \in \mathbf{U}^{\perp} \Leftrightarrow \mathcal{A}(u_1, y) = \mathcal{A}(u_2, y) = \ldots \mathcal{A}(u_p, y) = 0. \tag{2.1}$$

(*iii*) *If $dim\,\mathbf{V} = n \in \mathbf{N}^{\star}$, then $dim\,\mathbf{U} + dim\,\mathbf{U}^{\perp} \geq dim\,\mathbf{V}$; equality holds if and only if $\mathcal{A}|_{\mathbf{U}}$ is nondegenerate.*

(*iv*) *If $dim\,\mathbf{V} = n \in \mathbf{N}^{\star}$, then*

$$\mathbf{U} \oplus \mathbf{U}^{\perp} = \mathbf{V} \Leftrightarrow \mathcal{A}|_{\mathbf{U}} \text{ is nondegenerate.}$$

Proof. (*i*) Let $y_1, y_2 \in \mathbf{U}^{\perp}$, i.e. $\mathcal{A}(x, y_1) = 0$, $\mathcal{A}(x, y_2) = 0$, $\forall x \in \mathbf{U}$. For any $k, l \in \mathbf{K}$

$$\mathcal{A}(x, ky_1 + ly_2) = k\mathcal{A}(x, y_1) + l\mathcal{A}(x, y_2) = 0,$$

thus $ky_1 + ly_2 \in \mathbf{U}^{\perp}$.

(*ii*) $y \in \mathbf{U}^{\perp}$ implies $\mathcal{A}(x, y) = 0$, $\forall x \in \mathbf{U}$; in particular $\mathcal{A}(u_i) = 0$, $\forall i = 1, \ldots, p$, since $u_i \in \mathbf{U}$. Conversely, if $x \in \mathbf{U}$, then $x = \sum_{i=1}^{n} x_i u_i$, using the given basis of $\mathbf{U}$. Then

$$\mathcal{A}(x, y) = \sum_{i=1}^{n} x_i \mathcal{A}(u_i, y) = 0, \text{ i.e. } y \in \mathbf{U}^{\perp}.$$

(*iii*) If $\mathbf{V}$ is finite dimensional, so is $\mathbf{U}$. Fix a basis of $\mathbf{U}$ using the notation of (*ii*), and complete it to a basis of $\mathbf{V}$; in (2.1) write y as a linear combination of this basis. Then the rank of the system in (2.1) is at most p. Consequently, $dim\,\mathbf{U}^{\perp} \geq n - p$, so $dim\,\mathbf{U} + dim\,\mathbf{U}^{\perp} = n$. Equality holds when the rank of the system is exactly p, i.e. $dim\,\mathbf{U} = rank\,\mathcal{A}|_{\mathbf{U}}$; but this means that $\mathcal{A}|_{\mathbf{U}}$ is nondegenerate (see the last part of Section 1).

(*iv*) As a consequence of (*iii*), it suffices to show that

$$\mathbf{U} \cap \mathbf{U}^{\perp} = \{0\} \Leftrightarrow \mathcal{A}|_{\mathbf{U}} \text{ is nondegenerate.}$$

But this is obvious by Corollary 1.10. QED.

Example Note that if a symmetric bilinear form $\mathcal{A}$ (or the quadratic form Q) is nondegenerate on the space $\mathbf{V}$, it might happen that its restriction to some subspace of $\mathbf{V}$ is degenerate. For example, $Q(x) = x_1^2 - x_2^2 + x_3^2$ is nondegenerate on $\mathbf{R}^3$; however, the restriction of Q to $\mathbf{U} = \{x \in \mathbf{R}^3 \,|\, x_1 + x_2 = 0\}$ is degenerate, as $0 \neq x = (1, -1, 0) \in Ker(\mathcal{A}|_{\mathbf{U}})$. Indeed, for $\mathcal{A}(x, y) = x_1y_1 - x_2y_2 + x_3y_3$, $\mathcal{A}((1, -1, 0), (a, -a, b)) = 0$, $\forall a, b \in \mathbf{R}$.

DEFINITION 2.6 Let $\mathcal{A} : \mathbf{V} \times \mathbf{V} \to \mathbf{K}$ be a symmetric bilinear form. A basis $\{e_1, \ldots, e_n\}$ is called *an orthogonal basis w.r.t. the form* $\mathcal{A}$ (or *w.r.t. the quadratic form* Q) if $\mathcal{A}(e_i, e_j) = 0$ for $i \neq j$, $i, j = 1, \ldots, n$, i.e. the vectors of this basis are pairwise orthogonal. The associated matrix is a diagonal matrix. If we denote its diagonal entries by $a_i = a_{ii} = \mathcal{A}(e_i, e_i)$, $i = 1, \ldots, n$, then the corresponding analytic expressions of $\mathcal{A}$ and Q

$$\mathcal{A}(x, y) = \sum_{i=1}^{n} a_i x_i y_i, \quad Q(x) = \sum_{i=1}^{n} a_i x_i^2$$

are called *canonical expressions* of $\mathcal{A}$ and Q respectively.

3 Reduction of Quadratic Forms to Canonical Expression

Let $\mathbf{V}$ be a $\mathbf{K}$-vector space, $\mathbf{K} = \mathbf{R}$ or $\mathbf{K} = \mathbf{C}$, $dim\, \mathbf{V} = n \in \mathbf{N}^\star$ and $Q(x) = {}^tXAX$, a quadratic form on $\mathbf{V}$, expressed by the associated matrix A w.r.t. a fixed basis of $\mathbf{V}$. A change of basis corresponds to $X = CX'$, where C is the matrix of change. The matrix associated w.r.t. the new basis is $B = {}^tCAC$. Note that both A and B are symmetric matrices.

THEOREM 3.1 (Gauss method) *If* $Q : \mathbf{V} \to \mathbf{R}$ *is a quadratic form, then there exists a basis of* $\mathbf{V}$ *which is orthogonal w.r.t.* Q *(the associated analytic expression is canonical).*

Proof. We will describe an inductive algorithm which reduces the problem to the case of a quadratic form defined on a vector space of smaller dimension.

Denote by $\{e_1, \ldots, e_n\}$ the initial basis and by $A = [a_{ij}]$ the associated matrix. Suppose that the analytic expression $Q(x) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} a_{ij} x_i x_j$ is not canonical and Q is not identically zero (otherwise the problem is solved and nothing needs to be done).

Case 1. There exists i such that $a_{ii} \neq 0$.

We may assume without loss of generality that $i = 1$, so we can write

$$Q(x) = a_{11} x_1^2 + 2 \sum_{j=2}^{n} a_{1j} x_1 x_j + \sum_{i=2}^{n} \sum_{j=2}^{n} a_{ij} x_i x_j.$$

By factoring out $\dfrac{1}{a_{11}}$, from the terms which contain x_1, then completing the square we obtain

$$Q(x) = \frac{1}{a_{11}} (a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n)^2 + \sum_{i=2}^{n} \sum_{j=2}^{n} a'_{ij} x_i x_j.$$

The change of coordinates

$$x'_1 = a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n;\; x'_j = x_j,\; j = 2, \ldots n,$$

which correspond to a new basis $\{e'_1, e'_2 = e_2, \ldots, e'_n = e_n\}$ with matrix of change

$$C' = \begin{bmatrix} \frac{1}{a_{11}} & -\frac{a_{12}}{a_{11}} & \ldots & -\frac{a_{1n}}{a_{11}} \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{bmatrix},$$

gives

$$Q(x) = \frac{1}{a_{11}} x_1'^2 + \sum_{i=2}^{n} \sum_{j=2}^{n} a'_{ij} x'_i x'_j .$$

Now $\sum\limits_{i=2}^{n} \sum\limits_{j=2}^{n} a'_{ij} x_i x_j$ is the analytic expression of a quadratic form on *Span* $\{e'_2, \ldots, e'_n\}$; if this new quadratic form is identically zero, or in canonical form, the reduction is done, otherwise we continue by applying the algorithm for this new quadratic form, which is defined on a vector space of dimension $n-1$.

Case 2. $a_{ii} = 0$, $\forall i = 1, \ldots, n$.

As Q is not identically zero, there is at least one element $a_{ij} \neq 0$, $i \neq j$. After the change of coordinates

$$x_i = x'_i + x'_j ,\ x_j = x'_i - x'_j ,\ x_k = x'_k ,\ k \neq i, j$$

the expression of the quadratic form becomes $Q(x) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} a''_{ij} x'_i x'_j$ where $a''_{ii} \neq 0$, since $x_i x_j = x_i'^2 - x_j'^2$. The matrix of change of basis which corresponds to this change of coordinates is the nonsingular matrix

$$C'' = \begin{bmatrix} 1 & 0 & \ldots & 0 & \ldots & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 & \ldots & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 & \ldots & -1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & \ldots & 0 & \ldots & 1 \end{bmatrix}.$$

Now Case 1. applies for the new analytic expression of Q.

After at most $n-1$ steps we obtain an orthogonal basis w.r.t. Q.

The corresponding canonical expression represents a linear combination of squares of linearly independent forms. The number of these squares in the canonical expression is minimal, and it is equal to *rank* Q. QED.

Example 1. Determine the canonical form of $Q : \mathbf{R}^3 \to \mathbf{R}$, $Q(x) = 2x_1x_2 + 2x_1x_3$, using Gauss method. Point out the change of coordinates and the matrix of change.

Solution. Note that $a_{ii} = 0$, $i = 1, 2, 3$, so we need to start with a change of coordinates described in Case 2. Let

$$x_1 = x'_1 + x'_2 , \quad x'_2 = x'_1 - x'_2 , \quad x_3 = x'_3 ,$$

as $a_{12} = 1 \neq 0$. Then $Q(x) = 2x_1'^2 - 2x_2'^2 + 2x_1'x_3' + 2x_2'x_3'$.

Next, we follow Case 1 and obtain $Q(x) = \frac{1}{2}(2x_1' + x_3')^2 - \frac{1}{2}x_3'^2 - 2x_2'^2 + 2x_2'x_3'$. The change of coordinates

$$x_1'' = 2x_1' + x_3', \quad x_2'' = x_2', \quad x_3'' = x_3',$$

leads to $Q(x) = \frac{1}{2}x_1''^2 - \frac{1}{2}x_2''^2$.

Combining the two changes of coordinates we get

$$x_1'' = x_1 + x_2 + x_3, \quad x_2' = \frac{1}{2}x_1' - \frac{1}{2}x_2', \quad x_3 = x_3'.$$

The corresponding matrix of change is

$$C = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1/2 & 0 & -1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1/2 & 1 & -1/2 \\ 1/2 & -1 & -1/2 \\ 0 & 0 & 1 \end{bmatrix};$$

$$C^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 1/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The columns of C represent the coordinates of the vectors of the new basis w.r.t. the initial basis. Thus the new basis is $\{f_1, f_2, f_3\}$, $f_1 = \frac{1}{2}e_1 + \frac{1}{2}e_2$, $f_2 = e_1 - e_2$, $f_3 = -\frac{1}{2}e_1 - \frac{1}{2}e_2 + e_3$.

THEOREM 3.2 (Jacobi's method) *Let $Q : \mathbf{V} \to \mathbf{R}$ be a quadratic form, and $A = [a_{ij}]$ the associated matrix w.r.t. an arbitrary fixed basis $\mathbf{B} = \{e_1, \ldots, e_n\}$.*

Denote $\Delta_0 = 1$. If all the determinants

$$\Delta_1 = a_{11},\ \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},\ \Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix},\ \ldots,\ \Delta_n = det\, A$$

are not zero, then there exists a basis $\mathbf{B}_1 = \{f_1, \ldots, f_n\}$ such that the expression of Q w.r.t. this new basis is

$$Q(x) = \sum_{i=1}^{n} \frac{\Delta_{i-1}}{\Delta_i} x_i'^2, \tag{3.1}$$

where x_i', $i = 1, \ldots, n$ are the coordinates of x w.r.t. the basis $\mathbf{B}_1$.

Proof. We are looking for the vectors $f_1, \ldots, f_n$ of the form

$$f_1 = c_{11}e_1,\ f_2 = c_{12}e_1 + c_{22}e_2,\ \ldots\ f_n = c_{1n}e_1 + c_{2n}e_2 + \ldots c_{nn}e_n \tag{3.2}$$

satisfying

$$\mathcal{A}(e_i, f_j) = 0,\ if\ 1 \leq i < j \leq n, \qquad \mathcal{A}(e_i, f_i) = 1,\ i = 1, \ldots n \tag{3.3}$$

where $\mathcal{A}$ is the polar of the quadratic form Q. We will prove that this is the desired basis. Using the bilinearity of $\mathcal{A}$ in (3.3), we obtain the system

$$\begin{array}{ccccccccccc}
\mathcal{A}(e_1, f_i) & = & c_{1i}a_{11} & + & c_{2i}a_{12} & + & \dots & + & c_{ii}a_{1i} & = & 0 \\
\mathcal{A}(e_2, f_i) & = & c_{1i}a_{21} & + & c_{2i}a_{22} & + & \dots & + & c_{ii}a_{2i} & = & 0 \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & & \\
\mathcal{A}(e_{i-1}, f_i) & = & c_{1i}a_{i-1,1} & + & c_{2i}a_{i-1,2} & + & \dots & + & c_{ii}a_{i-1,i} & = & 0 \\
\mathcal{A}(e_i, f_i) & = & c_{1i}a_{i1} & + & c_{2i}a_{i2} & + & \dots & + & c_{ii}a_{ii} & = & 1
\end{array}$$

of i equations, with i unknowns: $c_{1i}, c_{2i}, \dots, c_{ii}$.

The determinant of this system is $\Delta_i \neq 0$, thus the solution is unique. By Cramer's rule we obtain

$$c_{ii} = \begin{bmatrix} a_{11} & \dots & a_{1,i-1} & 0 \\ \dots & \dots & \dots & \dots \\ a_{i-1,1} & \dots & a_{i-1,i-1} & 0 \\ a_{i1} & \dots & a_{i,i-1} & 1 \end{bmatrix} = \frac{\Delta_{i-1}}{\Delta_i}.$$

So the basis $\mathbf{B}_1 = \{f_1, \dots, f_n\}$ is uniquely determined by (3.2), (3.3). It remains to show that the matrix $B = [b_{ij}]$ associated to Q w.r.t. this basis is diagonal, with the required diagonal entries. For, compute

$$b_{ij} = \mathcal{A}(f_i, f_j) = \dots = c_{1i}\mathcal{A}(e_1, f_j) + \dots + c_{ii}\mathcal{A}(e_i, f_j),\ i \leq j.$$

Using (3.2), it follows that $b_{ij} = 0$, if $i < j$, and $b_{ii} = c_{ii} = \dfrac{\Delta_{i-1}}{\Delta_i}$. The matrix B is symmetric, since $\mathcal{A}$ is symmetric (as the polar of a quadratic form), thus $b_{ij} = 0$, for $i > j$ too. QED.

Note that Jacobi's method does not apply for **Example 1**.

THEOREM 3.3 (eigenvalues method) *Let* $\mathbf{V}$ *be a real* n*-dimensional vector space as before,* $Q : \mathbf{V} \to \mathbf{R}$ *a quadratic form, and* $A = [a_{ij}]$ *the matrix of* Q *w.r.t. an arbitrary fixed basis* $\mathbf{B}$*. Then there exists a basis* $\mathbf{B}_1 = \{f_1, \dots, f_n\}$ *such that the associated expression of* Q *is*

$$Q(x) = \sum_{i=1}^{n} \lambda_i x_i'^2 ,$$

where λ_i, $i = 1, \dots, n$ *are the eigenvalues of* A *(each eigenvalue appears as many times as its algebraic multiplicity), and* x_i', $i = 1, \dots, n$ *are the coordinates of* x *w.r.t. the basis* $\mathbf{B}_1$.

Proof. The matrix associated to a quadratic form is always symmetric, thus A has n real eigenvalues (counted with their multiplicities) and admits a diagonal form over $\mathbf{R}$. Moreover, there is a basis $\mathbf{B}_1 = \{f_1, \dots, f_n\}$ made of eigenvectors whose coordinate columns w.r.t. the initial basis $\mathbf{B}$ are orthonormal in $\mathbf{R}^n$, or equivalently, the matrix of change C is orthogonal (${}^tC = C^{-1}$). Denote $D = diag(\lambda_1, \dots, \lambda_n)$ the diagonal matrix whose diagonal entries are the eigenvalues of A. Then

$$D = C^{-1}AC = {}^tCAC$$

implies that D is the matrix associated to Q w.r.t. the basis $\mathbf{B}_1$. It follows that the expression of Q w.r.t. $\mathbf{B}_1$ is the one in the conclusion of the theorem. QED.

REMARKS 3.4 Comparison of the three methods

1) Gauss method is the most elementary and it provides a sequence of coordinate changes, not the new basis (which can be deduced by multiplying the corresponding matrices of change obtained at each step). The basis corresponding to the canonical expression has no special properties.

2) Jacobi's method is the most recommended (if it applies) when we need a canonical expression, and we are not interested in the corresponding basis, since it takes longer to deduce it.

3) The eigenvalues method is the most convenient for further purposes. If on $\mathbf{V}$ we consider the scalar product which makes the initial basis into an orthonormal basis, then the basis corresponding to the canonical expression is orthonormal too. For example, if $\mathbf{V} = \mathbf{R}^n$ and the initial basis is the canonical basis in $\mathbf{R}^n$, then the new basis is orthonormal (w.r.t. the canonical inner product).

The method provides the canonical expression and the corresponding basis. The basis has additional useful properties.

4 The Signature of a Real Quadratic Form

Real quadratic forms of constant sign are useful in many applications, for example in extemum problems for functions of several real variables. This why it is helpful to study these forms, and state some methods which indicate whether the sign is constant or not.

$\mathbf{V}$ will denote a real vector space.

DEFINITION 4.1 A quadratic form $Q : \mathbf{V} \to \mathbf{R}$ is said to be:

(i) *positive definite* if $Q(x) > 0$, $\forall x \in \mathbf{V} \setminus \{0\}$;

(ii) *negative definite* if $Q(x) < 0$, $\forall x \in \mathbf{V} \setminus \{0\}$;

(iii) *positive semidefinite* if $Q(x) \geq 0$, $\forall x \in \mathbf{V}$;

(iv) *negative semidefinite* if $Q(x) > 0$, $\forall x \in \mathbf{V}$;

(v) *indefinite* if $\exists\, x_1, x_2 \in \mathbf{V}$ such that $Q(x_1) > 0$ and $Q(x_2) < 0$.

The notions defined above are used for symmetric bilinear forms too, with the obvious meaning. A real symmetric bilinear form $\mathcal{A} : \mathbf{V} \times \mathbf{V} \to \mathbf{R}$ is *positive definite* if the associated quadratic form $Q(x) = \mathcal{A}(x, x)$ is positive definite, etc.

Example The inner product on a real vector space is a positive definite symmetric bilinear form.

From now on, throughout the section we assume that $dim\, \mathbf{V} = n \in \mathbf{N}^{\star}$.

The following remarks are easy observations relating some of the resuls in this chapter.

REMARKS 4.2 1) The number of nonzero terms in a canonical expression of Q is the same, regardless the method used, and it is equal to $rank\, Q$.

2) If Q is positive or negative definite, then Q is nondegenrate, i.e., in the canonical expression appear n nonzero terms.

3) Q is positive definite $\Leftrightarrow$ in a fixed canonical expression all coefficients are strictly positive.

In particular, from the eigenvalues method follows that: Q is positive definite $\Leftrightarrow$ all eigenvalues of Q are strictly positive.

4) Q is negative definite $\Leftrightarrow$ in a fixed canonical expression all coefficients are strictly negative.

From the eigenvalues method follows that: Q is negative definite $\Leftrightarrow$ all eigenvalues of Q are strictly negative.

If we combine Remarks 3), 4) above and Jacobi's method we obtain the following result.

THEOREM 4.3 (Sylvester's criterion) *If the hypotheses of Jacobi's method are fulfilled, then*
(i) Q is positive definite $\Leftrightarrow$ all determinants Δ_i, $i = 1, \ldots, n$ are strictly positive.
(ii) Q is negative definite $\Leftrightarrow$ the determinants Δ_i, $i = 1, \ldots, n$ have alternating signs: $\Delta_1 < 0$, $\Delta_2 > 0$, $\Delta_3 < 0$, ..., $(-1)^k \Delta_k > 0$, ...

DEFINITION 4.4 Let $Q(x) = \sum\limits_{i=1}^{n} a_i x_i^2$ be a canonical expression of $Q : \mathbf{V} \to \mathbf{R}$, $dim\, \mathbf{V} = n \in \mathbf{N}^\star$, where p of the coefficients $a_1, \ldots, a_n$ are strictly positive, q are strictly negative, and $d = n - (p + q)$ coefficients are equal to zero.

The triplet $(p, q, d) \in \mathbf{N}^3$ is called the *signature* of the form Q and is denoted by $sign\, Q$.

As we have seen, the canonical expression of a quadratic form is not unique. The next result states that the signature depends only on the quadratic form.

THEOREM 4.5 (Inertia law) *The signature of a quadratic form Q is the same, for any canonical expression of Q.*

Proof. Consider two bases $\mathbf{B} = \{e_1, \ldots, e_n\}$ and $\mathbf{B}' = \{e'_1, \ldots, e'_n\}$ in $\mathbf{V}$ such that the expression of Q w.r.t. each of them is canonical

$$Q(x) = \sum_{i=1}^{n} a_i x_i^2 \text{ and } Q(x) = \sum_{i=1}^{n} a'_i x_i'^2 \text{ respectively.}$$

We may assume that in the expression of Q w.r.t. $\mathbf{B}$ ($\mathbf{B}'$) the first p (p') coefficients are strictly positive, the next q (q') are strictly negative, and the last d (d') are zero. Morever, we may assume that $a_i, a'_i \in \{-1, 0, 1\}$. Then

$$Q(x) = \sum_{i=1}^{p} x_i^2 - \sum_{i=p+1}^{p+q} x_i^2 = \sum_{i=1}^{p'} x_i'^2 - \sum_{i=p'+1}^{p'+q'} x_i'^2. \tag{4.1}$$

We need to show that $p = p'$, $q = q'$. Suppose $p \neq p'$ and assume $p > p'$. Let $\mathbf{U} = Span(e_1, \ldots, e_p)$ and $\mathbf{U}' = Span(e'_{p'+1}, \ldots, e'_n)$, thus $dim\,\mathbf{U} = p$, $dim\,\mathbf{U}' = n - p'$. This yields

$$dim\,\mathbf{U} + dim\,\mathbf{U}' = p + n - p' > n,$$

which shows that the subspaces $\mathbf{U}$, $\mathbf{U}'$ are not independent. Therefore there exists $x \in \mathbf{U} \cap \mathbf{U}'$, $x \neq 0$, and x can be written as

$$x = x_1 e_1 + \ldots + x_p = x'_{p'+1} e'_{p'+1} + \ldots + x'_n e'_n.$$

Replacing this particular x in (4.1) we obtain

$$0 \leq x_1^2 + \ldots + x_p^2 = Q(x) = -x'^2_{p'+1} - \ldots - x'^2_n \leq 0.$$

But this implies

$$x_1 = \ldots = x_p = 0; \quad x'_{p'+1} = \ldots = x'_n = 0,$$

thus $x = 0$ which contradicts the choice of x. It follows that $p = p'$. Similarly, $q = q'$, hence $d = d'$ and $(p, q, d) = (p', q', d')$. QED.

COROLLARY 4.6 *1)The quadratic form $Q : \mathbf{V} \leftrightarrow \mathbf{R}$, $dim\,\mathbf{V} = n \in \mathbf{N}^\star$is positive definite if and only if one of the following conditions is satisfied:*

(i) *$sign\,Q = (n, 0, 0)$;*

(ii) *the determinants Δ_i, $i = 1, \ldots, n$ are strictly positive;*

(iii) *the eigenvalues of the matrix of Q are all strictly positive.*

2) Q is negative definite if and only if one of the following conditions is satisfied:

(i) *$sign\,Q = (0, 0, n)$;*

(ii) *$(-1)^i \Delta_i > 0$, $i = 1, \ldots, n$;*

(iii) *the eigenvalues of the matrix of Q are all strictly negative.*

Example 2. Determine the canonical form of $Q : \mathbf{R}^3 \to \mathbf{R}$, $Q(x) = x_1^2 + 7x_2^2 + x_3^2 - 8x_1x_2 - 16x_1x_3 - 8x_2x_3$, and the corresponding basis. Use Jacobi's method, then the eigenvalues method. Check the inertia law.

Solution.

$$A = \begin{bmatrix} 1 & -4 & -8 \\ -4 & 7 & -4 \\ -8 & -4 & 1 \end{bmatrix};$$

$$\Delta_0 = 1,\ \Delta_1 = a_{11} = 1,\ \Delta_2 = \begin{vmatrix} 1 & -4 \\ -4 & 7 \end{vmatrix} = -9, \Delta_3 = det\,A = -729.$$

Using formula (3.1) we obtain $Q(x) = x'^2_1 - \dfrac{1}{9} x'^2_2 + \dfrac{1}{81} x'^2_3$.

We look for the corresponding basis $\mathbf{B}_1 = \{f_1, f_2, f_3\}$ of the form (3.2). The coefficients c_{ij}, $1 \leq i < j \leq n = 3$ are given by the solutions of the linear equations (3.3) as follows.

$$\mathcal{A}(e_1, f_1) = 1 \Rightarrow c_{11} = 1.$$

$$\begin{cases} \mathcal{A}(e_1, f_2) = 0 \\ \mathcal{A}(e_2, f_2) = 1 \end{cases} \Rightarrow \begin{cases} c_{12} - 4c_{22} = 0 \\ -4c_{12} + 7c_{22} = 1 \end{cases} \Rightarrow c_{12} = \frac{-4}{9};\ c_{22} = \frac{-1}{9}.$$

$$\begin{cases} \mathcal{A}(e_1, f_3) = 0 \\ \mathcal{A}(e_2, f_3) = 0 \\ \mathcal{A}(e_3, f_3) = 1 \end{cases} \Rightarrow \begin{cases} c_{13} - 4c_{23} - 8c_{33} = 0 \\ -4c_{13} + 7c_{23} - 4c_{33} = 0 \\ -8c_{13} - 4c_{23} + c_{33} = 0 \end{cases} \Rightarrow c_{13} = \frac{-8}{81};\ c_{23} = \frac{-4}{81};\ c_{23} = \frac{1}{81}.$$

It follows that $f_1 = e_1$, $f_2 = -\frac{4}{9}e_1 - \frac{1}{9}e_2$, $f_3 = -\frac{8}{81}e_1 - \frac{4}{81}e_2 + \frac{1}{81}e_3$, and the matrix of change is

$$C = \begin{bmatrix} 1 & -4/9 & -8/81 \\ 0 & -1/9 & -4/81 \\ 0 & 0 & 1/81 \end{bmatrix}.$$

Next we will use the eigenvalues method.

The eigenvalues of A are $\lambda_1 = 9$, $\lambda_2 = 9$, $\lambda_3 = -9$. Therefore $Q(x) = 9y_1^2 + 9y_2^2 - 9y_3^2$, where y_1, y_2, y_3 are the coordinates of x w.r.t. a basis made up of eigenvectors, such that the matrix of change is orthogonal. To determine such a basis $\{u_1, u_2, u_3\}$, pick first some linearly independent eigenvectors. Let us take $v_1 = (-1, 0, 1)$, $v_2 = (-1, 2, 0)$, $v_3 = (2, 1, 2)$. Using the Gram-Schmidt procedure for the first two (which happen not to be orthogonal, as they correspond to the same eigenvalue), and normalizing the third we obtain the $u_1 = (-1/\sqrt{2}, 0, 1/\sqrt{2})$, $u_2 = (-\sqrt{2}/6, 2\sqrt{2}/3, -\sqrt{2}/6)$, $u_3 = (2/3, 1/3, 2/3)$.

As the inertia law states, from either one of the canonical expressions obtained using different methods, the signature is $(2, 1, 0)$.

5 Problems

1. Let V be a real vector space and $\omega : V \to \mathbf{R}$ a linear form. Show that $Q(x, y) = \omega(x)\omega(y)$, $x, y \in V$ is a positive semidefinite symmetric bilinear form on V.

2. Let P_3 be the real vector space of the polynomial functions of degree at most two and $Q : P_3 \times P_3 \to \mathbf{R}$, $Q(x, y) = \int_0^1 \int_0^1 x(t)y(s)dtds$.

1) Show that Q is a positive semidefinite symmetric bilinear form.

2) Determine the matrices of Q with respect to the bases $\{1, t, t^2\}$ and $\{1-t, 2t, 1+t+t^2\}$. What is the relationship between these matrices ?

3) What is the canonical expression of g ?

3. Consider the bilinear form

$$Q(x, y) = x_1y_2 - x_2y_1 + x_1y_3 - x_3y_1 + x_1y_4 - x_4y_1 + x_4y_4.$$

1) Write $Q(x, y)$ in matrix form.

2) Find the matrix associated to Q with respect to the basis $f_1 = (1, 1, 0, 1)$, $f_2 = (0, 1, 1, 0)$, $f_3 = (0, 1, 0, 1)$, $f_4 = (1, 0, 0, 1)$.

4. Reduce the quadratic forms to the canonical expression using orthogonal transformations, determine the signatures and the sets of constant level.

1) $5x_1^2 + 4x_2^2 + 6x_3^2 + 4x_1x_2 + 4x_1x_3$,
2) $3x_1^2 - 5x_2^2 - 7x_3^2 - 8x_1x_2 + 8x_2x_3$,
3) $2x_1x_2 + 2x_3x_4$

5. Show that the quadratic forms
1) $Q(x) = 3x_1^2 + 4x_2^2 + 5x_3^2 + 4x_1x_2 - 4x_2x_3$,
2) $Q(x) = 2x_1^2 + 5x_2^2 + 5x_3^2 + 4x_1x_2 - 4x_1x_3 - 8x_2x_3$,
are positive definite.

Chapter 5

Free Vectors

1 Free Vectors

$\mathbf{E}_3$ will denote the 3-dimensional space of elementary geometry.

For any two fixed points $A, B \in \mathbf{E}_3$ consider the oriented segment $\overrightarrow{AB}$.

A is called the *origin* of $\overrightarrow{AB}$ and B the vertex of $\overrightarrow{AB}$ (Fig. 1).

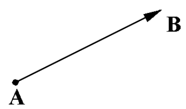

Fig. 1

If $A \neq B$, the straight line AB is called the *supporting line* of the oriented segment $\overrightarrow{AB}$. The *length* (*module* or *norm*) of $\overrightarrow{AB}$ is the length of the segment AB and it is denoted by $\| \overrightarrow{AB} \|$, AB or $d(A, B)$.

If $A = B$ we say that $\overrightarrow{AB}=\overrightarrow{AA}$ is the *zero oriented segment* with the origin at A. The zero segment has length zero.

Note that $\overrightarrow{AB}$ and $\overrightarrow{BA}$ are distinct oriented segments if $A \neq B$.

We say that 2 oriented segments have the *same direction* if their supporting lines are parallel or identical.

Since each straight line admits two different orientations, a nonzero oriented segment is characterized by its *direction, sense, length* and *origin.*

DEFINITION 1.1 Two nonzero oriented segments are said to be *equipolent* if they have the same direction, sense and length (but possibly different origins).

By definition, all zero oriented segments are equipolent.

If $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are equipolent, we write $\overrightarrow{AB} \sim \overrightarrow{CD}$ (Fig. 2).

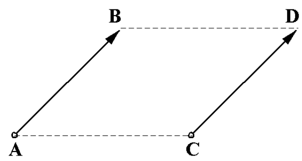

Fig. 2

PROPOSITION 1.2 *The equipolence relation is an equivalence relation on the set of all oriented segments in* $\mathbf{E}_3$*, i.e., it satisfies the following properties*

(i) $\overrightarrow{AB} \sim \overrightarrow{AB}$*, for any* $A, B \in \mathbf{E}_3$ *(reflexivity)*

(ii) $\overrightarrow{AB} \sim \overrightarrow{CD} \Rightarrow \overrightarrow{CD} \sim \overrightarrow{AB}$ *(symmetry)*

(iii) $\overrightarrow{AB} \sim \overrightarrow{CD}$ *and* $\overrightarrow{CD} \sim \overrightarrow{EF} \Rightarrow \overrightarrow{AB} \sim \overrightarrow{EF}$ *(transitivity).*

DEFINITION 1.3 A *free vector* is an equivalence class of the equipolence relation. The set of all free vectors is denoted by $\mathbf{V}_3$.

The equivalence class of the oriented segment $\overrightarrow{AB}$ is denoted by $\overline{AB}$ and it is called *the free vector AB*. Thus $\overline{AB} = \{\overrightarrow{CD} \mid \overrightarrow{CD} \sim \overrightarrow{AB}\}$, and $\overrightarrow{AB} \in \overline{AB}$. Any element of a free vector is called a *representative* of that free vector. $\overrightarrow{AB}$ is a reprezentative of $\overline{AB}$; if $\overrightarrow{AB} \sim \overrightarrow{MN}$ (or equivalently, $\overrightarrow{MN} \in \overline{AB}$), then $\overrightarrow{MN}$ is another representative of $\overline{AB}$, and $\overline{AB} = \overline{MN}$.

The equivalence class of all zero oriented segments is called *zero free vector* and is denoted by $\bar{0}$. Arbitrary free vectors will be denoted by overlined letters $\bar{a}, \bar{b}, \bar{c} \ldots$. According to the previous notation that involves a representative of the free vector, we may write either $\overrightarrow{AB} \in \bar{a}$ or $\overline{AB} = \bar{a}$, if $\bar{a}$ is the set of all oriented segments equipolent to $\overrightarrow{AB}$.

DEFINITION 1.4 The *length (norm)* of a free vector $\bar{a}$ is the length of a representative $\overrightarrow{AB}$, $\overrightarrow{AB} \in \bar{a}$.

We denote the length of $\bar{a}$ by $\|\bar{a}\|$. Thus $\|\bar{a}\| = \|\overline{AB}\| = \| \overrightarrow{AB} \| = d(A, B)$.

It follows that $\|\bar{a}\|$ is a positive real number and $\|\bar{a}\| = 0 \Leftrightarrow \bar{a} = \bar{0}$. Similarly, the *direction* and *sense* of a nonzero free vector are the direction and respectively the sense of a representative.

The above definitions are correct since all representatives of $\overline{AB}$ have the same length, same direction and same sense (for $A \neq B$).

A free vector of length one is called a *unit vector* or a *versor* and it is usually denoted by $\bar{e}$ or $\bar{u}$.

Nonzero free vectors are characterized by direction, sense and length. The length zero determines the zero free vector.

PROPOSITION 1.5 *For any* P *in* $\mathbf{E}_3$ *and any free vector* $\bar{a} \in \mathbf{V}_3$*, there exists a unique oriented segment* $\overrightarrow{PA}$ *(i.e. a unique vertex* A*) such that* $\overrightarrow{PA} \in \bar{a}$*.*

This proposition allows us to use representatives which have a convenient origin.

In $\mathbf{E}_3$ we fix a point O, called the *origin of the space* $\mathbf{E}_3$. By the previous proposition there is a bijective correspondence between $\mathbf{V}_3$ and the set of all oriented segments which have the origin at O.

$\mathbf{E}_3$ can be viewed as the set of all vertices of the oriented segments which have the origin at O. Therefore there is also a bijective correspondence between $\mathbf{E}_3$ and $\mathbf{V}_3$.

DEFINITION 1.6 Two nonzero free vectors of same direction are called *collinear vectors* (Fig. 3).

By definition, $\bar{0}$ is collinear with any other vector.

Three nonzero vectors are called *coplanar* if there is a plane containing their directions (Fig. 4).

By definition, $\bar{0}$ and any other two vectors are coplanar.

REMARKS 1.7 *(i) Let* $\bar{a} = \overline{PA}$, $\bar{b} = \overline{PB} \in \mathbf{V}_3 \setminus \{0\}$. *Then* $\bar{a}, \bar{b}$, *are collinear* $\Leftrightarrow$ *the points* P, A, B *are collinear.*

(ii) Let $\bar{a} = \overline{PA}$, $\bar{b} = \overline{PB}, \bar{c} = \overline{PC} \in \mathbf{V}_3 \setminus \{0\}$. *Then* $\bar{a}, \bar{b}, \bar{c}$ *are coplanar* $\Leftrightarrow$ *the points* P, A, B, C *are coplanar.*

The collinear vectors $\overline{AB}$ and $\overline{BA}$ are called *opposite vectors*. It follows that the opposite of $\bar{0}$ is $\bar{0}$. If $\bar{a} \neq \bar{0}$, then the opposite of $\bar{a}$ is the unique free vector which has the same direction and the same length as $\bar{a}$, but it has opposite sense. In the picture below $-\bar{a}$ is the opposite of $\bar{a}$ (Fig. 3).

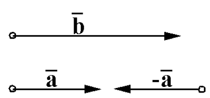

Fig. 3

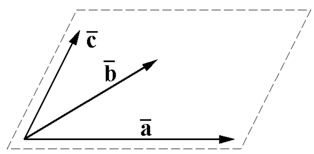

Fig. 4

2 Addition of Free Vectors

DEFINITION 2.1 Let $\bar{a}, \bar{b} \in \mathbf{V}_3$ such that $\overline{PA} = \bar{a}$ and $\overline{AB} = \bar{b}$, where P is an arbitrary point in $\mathbf{E}_3$.

The *sum* of $\bar{a}, \bar{b}$ is the free vector denoted by $\bar{c} = \bar{a} + \bar{b}$ and defined by (Fig. 5)

$$\bar{a} + \bar{b} = \overline{PB}$$

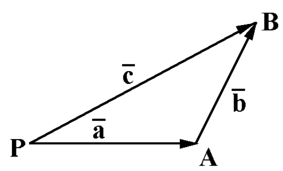

Fig. 5

This addition rule is called the *triangle law.*

If $\bar{a}, \bar{b}$ are not collinear, we may use the equivalent *parallelogram law* and define $\bar{a}+\bar{b}$ as the class of the oriented diagonal $\overrightarrow{PB}$ of the parallelogram $PABC$, where $\overline{PA}=\bar{a}$, $\overline{PC}=\bar{b}$.

Addition of free vectors $+ : \mathbf{V}_3 \times \mathbf{V}_3 \to \mathbf{V}_3$ is a well defined binary operation since it does not depend on the choice of the point P.

THEOREM 2.2 *Addition of free vectors has the following properties:*

(i) $\forall \bar{a}, \bar{b}, \bar{c} \in \mathbf{V}_3, \quad \bar{a}+(\bar{b}+\bar{c})=(\bar{a}+\bar{b})+\bar{c}$ *(associativity)*

(ii) $\forall \bar{a}, \bar{b} \in \mathbf{V}_3, \quad \bar{a}+\bar{b}=\bar{b}+\bar{a}$ *(commutativity)*

(iii) $\forall \bar{a} \in \mathbf{V}_3, \quad \bar{a}+\bar{0}=\bar{0}+\bar{a}=\bar{a}$ *($\bar{0}$ is the identity element)*

(iv) $\forall \bar{a} \in \mathbf{V}_3 \quad \exists -\bar{a} \in \mathbf{V}_3$ *such that* $\bar{a}+(-\bar{a})=(-\bar{a})+\bar{a}=\bar{0}$ *(each vector has an inverse with respect to "+")*

Proof. (ii) and (iii) are immediate from the definition.

For (iv) we check easily that $-\bar{a}$ is the opposite of $\bar{a}$, i.e. if $\bar{a}=\overline{PA}$, then $-\bar{a}=\overline{AP}$.

If $\bar{a}, \bar{b}, \bar{c}$ are pairwise not collinear, then associativity follows from Fig. 6.

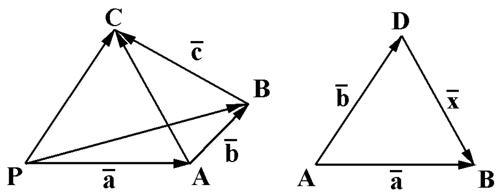

Fig. 6 **Fig. 7**

COROLLARY 2.3 $(\mathbf{V}_3, +)$ *is an abelian group.*

Like in any abelian group with additive notation, we may define the *subtraction* (*difference*) of vectors. The difference $\bar{a}-\bar{b}$ is the unique solution of the equation $\bar{b}+\bar{x}=\bar{a}$; this is $\bar{x}=\bar{a}+(-\bar{b})$.

If $\bar{a}=\overline{AB}$ and $\bar{b}=\overline{AD}$, then $\bar{x}=\bar{a}-\bar{b}=\overline{DB}$ (Fig. 7).

By the observation which proves (iv), it makes sense to write $\overline{AB}=-\overline{BA}$, for any $A, B \in \mathbf{E}_3$.

3 Multiplication by Scalars

In this section we will define on $\mathbf{V}_3$ a natural structure of a real vector space. Addition of free vectors was already defined. We still need an external operation $\mathbf{R} \times \mathbf{V}_3 \to \mathbf{V}_3$ which satisfies the definition of the vector space.

DEFINITION 3.1 (**multiplication of free vectors by scalars**, Fig. 8).

Let $k \in \mathbf{R}$ and $\bar{a} \in \mathbf{V}_3$. Then $k\bar{a} \in \mathbf{V}_3$ is defined as follows

(i) if $\bar{a}=\bar{0}$ or $k=0$, then $k\bar{a}=\bar{0}$;

(ii) if $\bar{a} \neq \bar{0}$ and $k \neq 0$, then $k\bar{a}$ is the vector which has the direction of $\bar{a}$, length $|k|\,\|\bar{a}\|$, the sense of $\bar{a}$ for $k>0$, and the sense of $-\bar{a}$ for $k<0$.

Note that $\bar{a}$ and $k\bar{a}$ are collinear, and $\|k\bar{a}\|=|k| \cdot \|\bar{a}\| \;\forall\, k \in \mathbf{R}$, $\bar{a} \in \mathbf{V}_3$.

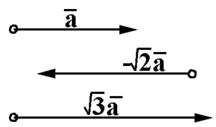

Fig. 8

THEOREM 3.2 *Multiplication of free vectors by real scalars has the following properties:*

(i) $\forall k, l \in \mathbf{R},\ \forall \bar{a} \in \mathbf{V}_3,\ k(l\bar{a}) = (kl)\bar{a}$
(ii) $\forall \bar{a} \in \mathbf{V}_3,\ 1\bar{a} = \bar{a}$
(iii) $\forall k, l \in \mathbf{R},\ \forall \bar{a} \in \mathbf{V}_3,\ (k+l)\bar{a} = k\bar{a} + l\bar{a}$
(iv) $\forall k \in \mathbf{R},\ \forall \bar{a}, \bar{b} \in \mathbf{V}_3,\ k(\bar{a}+\bar{b}) = k\bar{a} + k\bar{b}$.

Proof. $(i) - (iii)$ are left to the reader.

We will prove (iv) for $\bar{a}, \bar{b}$ not collinear. The collinear case is left to the reader as well.

Let $\overline{OA} = \bar{a}$ and $\overline{AB} = \bar{b}$. Then $\bar{a} + \bar{b} = \overline{OA} + \overline{AB} = \overline{OB}$.

Suppose $k > 0$. Then there is A' such that $\overline{OA'} = k\bar{a}$ and B' such that $\overline{OB'} = k(\bar{a} + \bar{b})$.

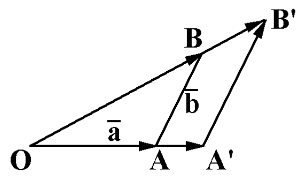

Fig. 9

The triangles $\triangle OAB$ and $\triangle OA'B'$ are similar since they have a common angle whose sides are proportional. This similarity implies $AB \| A'B'$ and $\|\overline{A'B'}\| = k\|\overline{AB}\|$. The same orientation of $\overline{OA}, \overline{OA'}$ and $\overline{OB}, \overline{OB'}$ gives the same orientation of $\overline{AB}$ and $\overline{A'B'}$, thus $\overline{A'B'} = k\overline{AB}$ (Fig. 9).

The case $k < 0$ is similar to $k > 0$, and the case $k = 0$ is trivial. QED.

COROLLARY 3.3 $\mathbf{V}_3$ *is a real vector space.*

4 Collinearity and Coplanarity

The collinearity and coplanarity of free vectors are notions we defined geometrically. In this section we will see they are related to the algebraic structure of a real vector space that was defined on $\mathbf{V}_3$.

PROPOSITION 4.1 *If the free vectors $\bar{a}$ and $\bar{b}$ are collinear and $\bar{a} \neq \bar{0}$, then there exists a unique real number k such that $\bar{b} = k\bar{a}$.*

Proof. If $\bar{b} = \bar{0}$, pick $k = 0$.

If $\bar{b} \neq \bar{0}$ and $\bar{b}$ has the sense of $\bar{a}$, then

$$\bar{b} = \frac{\|\bar{b}\|}{\|\bar{a}\|}\bar{a},$$

since $\dfrac{\|\bar{b}\|}{\|\bar{a}\|}\bar{a}$ has the direction, sense and length of $\bar{b}$.

Similarly, if $\bar{b} \neq 0$ and the sense of $\bar{b}$ is opposite to sense of $\bar{a}$ (i.e., $\bar{b}$ has the sense of $-\bar{a}$), $\bar{b} = -\dfrac{\|\bar{b}\|}{\|\bar{a}\|}\bar{a}$, and the existence of k is proved.

For the uniqueness, let $k, l \in \mathbf{R}$ such that $\bar{b} = k\bar{a} = l\bar{a}$. Then $k = l$ since $\bar{a} \neq 0$, by the properties of vector space operations. QED

COROLLARY 4.2 *(i) The vectors $\bar{a}, \bar{b}$ are collinear if and only if they are linearly dependent.*

(ii) If $\bar{a} \neq 0$, then the set $\mathbf{V}_1 = \{\bar{b} |\ \bar{a}, \bar{b}$ collinear$\}$ is a 1-dimensional vector space. More precisely, $\mathbf{V}_1 = \{\bar{b} |\ \exists k \in \mathbf{R},\ \bar{b} = k\bar{a}\} = Span\,\{\bar{a}\}$.

(iii) Two noncollinear vectors are linearly independent.

PROPOSITION 4.3 *If the vectors $\bar{a}, \bar{b}, \bar{c}$ are coplanar and $\bar{a}, \bar{b}$ are not collinear, then there exist the unique numbers $s,\ t \in \mathbf{R}$ such that $\bar{c} = s\bar{a} + t\bar{b}$.*

Proof. If $\bar{c} = \bar{0}$, take $s = t = 0$. If $\bar{c}$ and $\bar{a}$ are collinear, take s such that $\bar{c} = s\bar{a}$ and $t = 0$.

If $\bar{c}$ and $\bar{a}$ are not collinear and $\bar{c}$, $\bar{b}$ are not collinear either, let A, B, C such that $\overline{OA} = \bar{a}$, $\overline{OB} = \bar{b}$, $\overline{OC} = \bar{c}$. Then O, A, B, C are coplanar points.

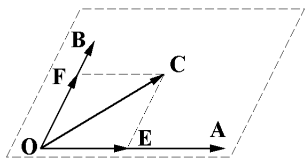

Fig. 10

The parallel lines through C to the supporting lines of $\overline{OA}$ and $\overline{OB}$ respectively, determine the vectors $\overline{OE}$ and $\overline{OF}$ such that (Fig. 10):

$\overline{OC} = \overline{OE} + \overline{OF}$, $\overline{OA}$ and $\overline{OE}$ collinear, and $\overline{OB}$, $\overline{OF}$ collinear. Then we can write $\overline{OE} = s\overline{OA}$, $\overline{OF} = s\overline{OB}$ for some $s, t \in \mathbf{R}$, and the existence of the decomposition is proved.

For the uniqueness assume

$$\bar{c} = s\bar{a} + t\bar{b} = s_1\bar{a} + t_1\bar{b}$$

Then $(s-s_1)\bar{a} + (t-t_1)\bar{b} = \bar{0}$. Since $\bar{a}$, $\bar{b}$ are not collinear, they are linearly independent, so $s - s_1 = 0$ and $t - t_1 = 0$. QED

COROLLARY 4.4 *(i) The vectors $\bar{a}$, $\bar{b}$, $\bar{c}$ are coplanar if and only if they are linearly dependent.*

(ii) If $\bar{a}$, $\bar{b}$ are not collinear, then the set $\mathbf{V}_2 = \{\bar{c} |\ \bar{a}, \bar{b}, \bar{c}$ coplanar$\}$ is a 2-dimensional vector space. More precisely, $\mathbf{V}_2 = \{\bar{c} = s\bar{a} + t\bar{b} |\ s, t \in \mathbf{R}\} = Span\,\{\bar{a}, \bar{b}\}$.

(iii) Three noncoplanar vectors are linearly independent.

THEOREM 4.5 $\mathbf{V}_3$ *is a 3-dimensional real vector space.*

Proof. We saw that any three vectors which are not coplanar, are linearly independent. Let $\bar{a} = \overline{OA}$, $\bar{b} = \overline{OB}$, $\bar{c} = \overline{OC}$ such that O, A, B, C are not coplanar. We will show that $\bar{a}$, $\bar{b}$, $\bar{c}$ span $\mathbf{V}_3$.

Let $\bar{d} = \overline{OD} \in \mathbf{V}_3$. If $\bar{d}$ is coplanar with 2 of the vectors $\bar{a}$, $\bar{b}$, $\bar{c}$, we are done. Assume it is not. Then the planes through D which are parallel to the planes (OAB), (OAC), (OBC) determine a parallelipiped.

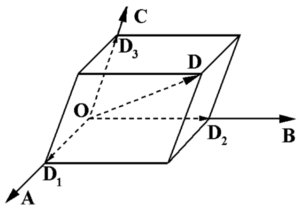

Fig. 11

$\overline{OD}$ is an oriented diagonal of this parallelipiped and it decomposes as (Fig. 11)

$$\overline{OD} = \overline{OD_1} + \overline{OD_2} + \overline{OD_3} = r\overline{OA} + s\overline{OB} + t\overline{OC}.$$

Therefore $\bar{d} = \overline{OD} \in Span\{\bar{a}, \bar{b}, \bar{c}\}$. QED

5 Inner Product in $\mathbf{V}_3$

DEFINITION 5.1 Let $\bar{a}, \bar{b} \in \mathbf{V}_3 \setminus \{\bar{0}\}$ and $\bar{a} = \overline{OA}$, $\bar{b} = \overline{OB}$.

The *angle* $\angle(\bar{a}, \bar{b})$ between $\bar{a}$ and $\bar{b}$ is the angle $\angle AOB \in [0, \pi]$ of their representatives $\overrightarrow{OA}$ and $\overrightarrow{OB}$.

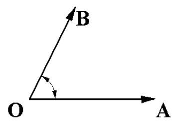

Fig. 12

The definition of the angle does not depend on the representatives chosen.

DEFINITION 5.2 Let $\bar{u} \in \mathbf{V}_3$ be a versor, $\bar{b} \neq \bar{0}$ and $\theta = \angle(\bar{u}, \bar{b})$.

The vector $\pi_{\bar{u}}\bar{b} = \|\bar{b}\| \cos\theta \, \bar{u}$ is called the *orthogonal projection* of $\bar{b}$ onto $\bar{u}$.

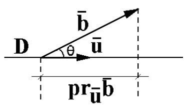

Fig. 13

The number $pr_{\bar{u}}\bar{b} = \|\bar{b}\| \cos\theta$ is called the *algebraic measure* of the *orthogonal projection* of $\bar{b}$ onto $\bar{u}$ (Fig. 13).

By definition $\pi_{\bar{u}}\bar{0} = \bar{0}, \quad pr_{\bar{u}}\bar{0} = 0$.

Note that

$$pr_{\bar{u}}\bar{b} < 0 \Leftrightarrow \theta \in \left(\frac{\pi}{2}, \pi\right]; \quad pr_{\bar{u}}\bar{b} = 0 \Leftrightarrow \theta = \frac{\pi}{2} \text{ or } \bar{b} = \bar{0}.$$

LEMMA 5.3 *For any versor $\bar{u}$, the functions $\pi_{\bar{u}} : \mathbf{V}_3 \to \mathbf{V}_3$ and $pr_{\bar{u}} : \mathbf{V}_3 \to \mathbf{R}$ are linear transformations.*

We leave the proof as an exercise. The picture below (Fig. 14) describes one of the cases that arise in order to prove additivity, where $\overline{OA}$ is collinear to $\bar{u}$, $\overline{OB} = \bar{b}$, $\overline{OC} = \bar{c}$, $OABC$ is a parallelogram, $BB' \perp OA$, $CC' \perp OA$, $DD' \perp OA$.

Then $\overline{AD} = \bar{b} + \bar{c}$, $\pi_{\bar{u}}\bar{b} = \overline{OB'} = \overline{C'D'}$, $\pi_{\bar{u}}\bar{c} = \overline{OC'}$, $\pi_{\bar{u}}(\bar{b} + \bar{c}) = \pi_{\bar{u}}(\overline{OD}) = \overline{OD'}$. The additivity of $\pi_{\bar{u}}$ follows from $\overline{OD'} = \overline{OC'} + \overline{C'D'}$.

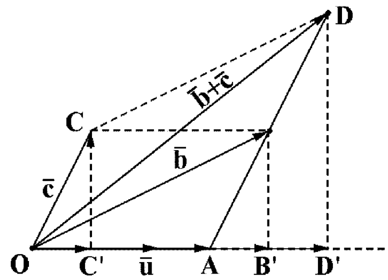

Fig. 14

THEOREM 5.4 *The function $\langle , \rangle : \mathbf{V}_3 \times \mathbf{V}_3 \to \mathbf{R}$ defined by*

$$\langle \bar{a}, \bar{b} \rangle = \begin{cases} \|\bar{a}\| \, \|\bar{b}\| \cos\theta, & \text{if } \bar{a} \neq \bar{0}, \, \bar{b} \neq 0 \text{ and } \theta = \angle(\bar{a}, \bar{b}) \\ 0, & \text{if } \bar{a} = \bar{0} \text{ or } \bar{b} = \bar{0} \end{cases}$$

is a scalar (dot, inner) product in $\mathbf{V}_3$.

Proof. We need to check the symmetry (commutativity), homogeneity, additivity and positivity of the function defined above (see the definition of the scalar product on a real vector space). All are straightforward except for additivity. Let us prove the additivity of $\langle , \rangle$ with respect to the second argument. For, assume $\bar{a} \neq \bar{0}$ (the case $\bar{a} = \bar{0}$ is obvious), and let $\bar{u} = \dfrac{\bar{a}}{\|\bar{a}\|}$ be the versor of $\bar{a}$. Then

$$\langle \bar{a}, \bar{b} + \bar{c} \rangle = \langle \|\bar{a}\|\bar{u}, \bar{b} + \bar{c} \rangle = \|\bar{a}\| \langle \bar{u}, \bar{b} + \bar{c} \rangle, \text{ by homogeneity.}$$

From the definitions

$$pr_{\bar{u}}\bar{b} = \langle \bar{u}, \bar{b} \rangle, \quad \forall \bar{b} \in \mathbf{V}_3.$$

Then

$$\langle \bar{u}, \bar{b} + \bar{c} \rangle = \langle \bar{u}, \bar{b} \rangle + \langle \bar{u}, \bar{c} \rangle \text{ by the previous lemma.}$$

It follows that

$$\begin{aligned}\langle \bar a, \bar b + \bar c\rangle &= \|\bar a\|(\langle \bar u, \bar b\rangle + \langle \bar u, \bar c\rangle)\\ &= \langle \|\bar a\|\bar u, \bar b\rangle + \langle \|\bar a\|\bar u, \bar c\rangle\\ &= \langle \bar a, \bar b\rangle + \langle \bar a, \bar c\rangle.\end{aligned}$$

The scalar product defined above is called the *canonical scalar product* on $\mathbf{V}_3$.

COROLLARY 5.5 $\mathbf{V}_3$ *is a Euclidean vector space.*

REMARKS 5.6 1) $\|\bar a\| = \sqrt{\langle \bar a, \bar a\rangle}, \quad \forall \bar a \in \mathbf{V}_3$. Therefore the length of a free vector is the Euclidean norm of that vector with respect to the canonical scalar product.

2) The Cauchy-Schwarz inequality follows directly from the definition using:

$$|\cos\theta| \le 1 \Rightarrow |\langle \bar a, \bar b\rangle| \le \|\bar a\| \cdot \|\bar b\|.$$

3) $\langle \bar a, \bar b\rangle = 0 \Leftrightarrow \bar a = 0$ or $\bar b = \bar 0$ or $\angle(\bar a, \bar b) = \dfrac{\pi}{2}$.

As in a general Euclidean vector space, the free vectors $\bar a, \bar b$ are called *orthogonal* if $\langle \bar a, \bar b\rangle = 0$.

It is convenient to work with orthonormal bases.

The coordinates of a vector with respect to an orthonormal basis are called *Euclidean coordinates.*

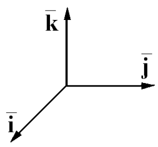

Fig. 15

Let $\{\bar i, \bar j, \bar k\}$ be an orthonormal basis (Fig. 15), i.e. the values of the scalar products are given by the table

$\langle , \rangle$	$\bar i$	$\bar j$	$\bar k$
$\bar i$	1	0	0
$\bar j$	0	1	0
$\bar k$	0	0	1

which leads to the canonical expression of the scalar product of $\bar a = a_1\bar i + a_2\bar j + a_3\bar k$ and $\bar b = b_1\bar i + b_2\bar j + b_3\bar k$, namely

$$\langle \bar a, \bar b\rangle = a_1b_1 + a_2b_2 + a_3b_3.$$

In particular the norm of $\bar a$ is $\|\bar a\| = \sqrt{a_1^2 + a_2^2 + a_3^2}$ and

$$\cos\theta = \frac{\langle \bar a, \bar b\rangle}{\|\bar a\| \cdot \|\bar b\|} = \frac{a_1b_1 + a_2b_2 + a_3b_3}{\sqrt{a_1^2 + a_2^2 + a_3^2} \cdot \sqrt{b_1^2 + b_2^2 + b_3^2}}, \quad \text{for} \quad \bar a \ne \bar 0, \bar b \ne \bar 0,\ \theta \in [0, \pi].$$

6 Vector (cross) Product in $\mathbf{V}_3$

Let $\bar{a}, \bar{b} \in \mathbf{V}_3$. If $\bar{a} \neq \bar{0}$ and $\bar{b} \neq \bar{0}$, then θ denotes the angle of $\bar{a}$ and $\bar{b}$.

DEFINITION 6.1 The vector denoted $\bar{a} \times \bar{b}$ and defined by (Fig. 16)

$$\bar{a} \times \bar{b} = \begin{cases} \|\bar{a}\| \, \|\bar{b}\| \sin\theta \bar{e}, & \text{if } \bar{a}, \bar{b} \text{ noncollinear} \\ \bar{0}, & \text{if } \bar{a}, \bar{b} \text{ collinear (in particular } \bar{a} = \bar{0} \text{ or } \bar{b} = \bar{0}), \end{cases}$$

where $\bar{e}$ is a versor whose direction is given by $\bar{e} \perp \bar{a}$, $\bar{e} \perp \bar{b}$ and the sense of $\bar{e}$ is given by the right-hand rule for $(\bar{a}, \bar{b}, \bar{e})$ is called *the vector (or cross)* product of $\bar{a}$ and $\bar{b}$.

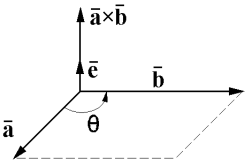

Fig. 16

The function: $\mathbf{V}_3 \times \mathbf{V}_3 \to \mathbf{V}_3$, $\quad (\bar{a}, \bar{b}) \to \bar{a} \times \bar{b}$ is bilinear. This follows from the proposition below.

PROPOSITION 6.2 *The vector product has the following properties:*
(1) $\bar{a} \times \bar{b} = \bar{0} \Leftrightarrow \bar{a}, \bar{b}$ *collinear; in particular* $\bar{a} \times \bar{0}$, $\bar{a} \times \bar{a} = \bar{0}$
(2) $\bar{b} \times \bar{a} = -\bar{a} \times \bar{b}$ *(anticommutativity)*
(3) $t(\bar{a} \times \bar{b}) = (t\bar{a}) \times \bar{b} = \bar{a} \times (t\bar{b}), \quad \forall t \in \mathbf{R}$ *(homogeneity)*
(4) $\bar{a} \times (\bar{b} + \bar{c}) = \bar{a} \times \bar{b} + \bar{a} \times \bar{c}$ *(additivity or distributivity)*
(5) $\|\bar{a} \times \bar{b}\|^2 = \|\bar{a}\|^2 \cdot \|\bar{b}\|^2 - \langle \bar{a}, \bar{b} \rangle^2$ *(the Lagrange identity)*
(6) If $\bar{a}, \bar{b}$ *are not collinear, then* $\|\bar{a} \times \bar{b}\|$ *represents the area of the parallelogram determined by* $\bar{a}$ *and* $\bar{b}$.

Proof. (1), (2), (3) are immediate from the definition. For (5) multiply

$$\sin^2\theta = 1 - \cos^2\theta$$

by $\|\bar{a}\|^2 \cdot \|\bar{b}\|^2$. The area of the parallelogram in Fig.16 is

$$\|\bar{a}\| \cdot \|\bar{b}\| \sin\theta = \|\bar{a} \times \bar{b}\|,$$

so 6) is an obvious geometric interpretation of the vector product. For (4) assume that $\bar{a}$ is a versor and let α be a plane perpendicular on $\bar{a}$ (Fig. 17). Let also $\bar{b}'$ and $\bar{c}'$ be the projections of $\bar{b}$ respectively $\bar{c}$ onto the plane P (i.e. $\bar{b}'$ is determined by the intersections of the lines passing through O and B and perpendicular to P, and P, where $\bar{b} = \overline{OB}$)

Then

$$\bar{a} \times \bar{b} = \bar{a} \times \bar{b}', \quad \bar{a} \times \bar{c} = \bar{a} \times \bar{c}'$$

and

$$\bar{a} \times \bar{b}', \; \bar{a} \times \bar{c}', \; \bar{a} \times (\bar{b}' + \bar{c}')$$

are obtained from $\bar{b}', \bar{c}', \bar{b}' + \bar{c}'$ respectively, by a rotation of angle $\dfrac{\pi}{2}$ about the axis $\bar{a}$. Since the rotation of the sum is the sum of rotations, i.e.

$$\bar{a} \times (\bar{b}' + \bar{c}') = \bar{a} \times \bar{b}' + \bar{a} \times \bar{c}',$$

it follows that

$$\bar{a} \times (\bar{b} + \bar{c}) = \bar{a} \times \bar{b} + \bar{a} \times \bar{c}.$$

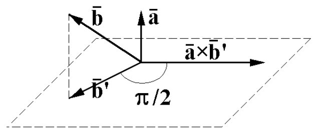

Fig. 17

From now on the canonical basis $\{\bar{i}, \bar{j}, \bar{k}\}$ of $\mathbf{V}_3$ is an orthonormal basis satisfying also $\bar{k} = \bar{i} \times \bar{j}$ (Fig. 15).

i.e. $\bar{i}, \bar{j}, \bar{k}$ are the versors of a Cartesian system of coordinate axes in E_3.

Then the cross products of these basis vectors are in the following table

$\times$	$\bar{i}$	$\bar{j}$	$\bar{k}$
$\bar{i}$	$\bar{0}$	$\bar{k}$	$-\bar{j}$
$\bar{j}$	$-\bar{k}$	$\bar{0}$	$\bar{i}$
$\bar{k}$	$\bar{j}$	$-\bar{i}$	$\bar{0}$

and the expression of $\bar{a} \times \bar{b}$ can be written as a symbolic determinant,

$$\bar{a} \times \bar{b} = \begin{vmatrix} \bar{i} & \bar{j} & \bar{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = (a_2 b_3 - a_3 b_2)\bar{i} + (a_3 b_1 - a_1 b_3)\bar{j} + (a_1 b_2 - a_2 b_1)\bar{k}.$$

DEFINITION 6.3 The vector $\bar{a} \times (\bar{b} \times \bar{c})$ is called the *double vector product* of $\bar{a}, \bar{b}, \bar{c}$.

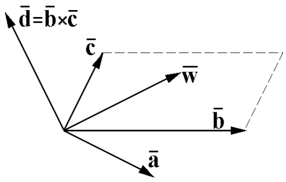

Fig. 18

It can be shown that

$$\bar{a} \times (\bar{b} \times \bar{c}) = \begin{vmatrix} \bar{b} & \bar{c} \\ \langle \bar{a}, \bar{b} \rangle & \langle \bar{a}, \bar{c} \rangle \end{vmatrix} = \langle \bar{a}, \bar{c} \rangle \bar{b} - \langle \bar{a}, \bar{b} \rangle \bar{c}.$$

Note that $\bar{w} = \bar{a} \times (\bar{b} \times \bar{c}), \bar{b}, \bar{c}$ are coplanar (Fig. 18). In general $\bar{a} \times (\bar{b} \times \bar{c}) \neq (\bar{a} \times \bar{b}) \times \bar{c}$ thus the vector product is not associative.

7 Mixed Product

DEFINITION 7.1 Let $\bar{a}, \bar{b}, \bar{c} \in \mathbf{V}_3$. The number $\langle \bar{a}, \bar{b} \times \bar{c} \rangle$ is called the *mixed product* of the vectors $\bar{a}, \bar{b}, \bar{c}$.

PROPOSITION 7.2 *(1)* $\langle \bar{a}, \bar{b} \times \bar{c} \rangle = 0 \Leftrightarrow \bar{a}, \bar{b}, \bar{c}$ *are coplanar.*

(2) If $\bar{a}, \bar{b}, \bar{c}$ *are not coplanar, then the absolute value* $|\langle \bar{a}, \bar{b} \times \bar{c} \rangle|$ *represents the volume of the parallelepiped determined by the representatives with common origin of* $\bar{a}, \bar{b}, \bar{c}$ *(Fig. 19).*

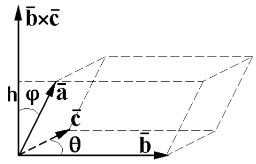

Fig. 19

Proof. (1) Assume $\bar{b}, \bar{c}$ noncollinear (otherwise $\bar{b} \times \bar{c} = \bar{0}$ and the equivalence is obvious). Then

$$\begin{aligned} \langle \bar{a}, \bar{b} \times \bar{c} \rangle = 0 \quad & \Leftrightarrow \bar{a} \perp \bar{b} \times \bar{c} \\ & \Leftrightarrow \bar{a} = \bar{0} \text{ or the direction of } \bar{a} \\ & \qquad \text{is parallel to a plane containing the directions of } \bar{b} \text{ and } \bar{c} \\ & \Leftrightarrow \bar{a}, \bar{b}, \bar{c} \text{ coplanar.} \end{aligned}$$

Let $\mathcal{V}$ be the volume

$$\mathcal{V} = \|\bar{b} \times \bar{c}\| \cdot h = \|\bar{b} \times \bar{c}\| \cdot \|\bar{a}\| \cdot |\cos \varphi|,$$

where $\varphi = \angle(\bar{a}, \bar{b} \times \bar{c})$. Since $\|\bar{b} \times \bar{c}\|$ is the area of the parallelipiped basis. Then $\mathcal{V} = |\langle \bar{a}, \bar{b}, \bar{c} \rangle|$.

PROPOSITION 7.3 (Properties of the mixed product)

(1) $$\textit{If} \quad \begin{matrix} \bar{a} = a_1\bar{i} + a_2\bar{j} + a_3\bar{k} \\ \bar{b} = b_1\bar{i} + b_2\bar{j} + b_3\bar{k} \\ \bar{c} = c_1\bar{i} + c_2\bar{j} + c_3\bar{k} \end{matrix}, \quad \textit{then } \langle \bar{a}, \bar{b} \times \bar{c} \rangle = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

(2) $$\begin{aligned} \langle \bar{a}, \bar{b} \times \bar{c} \rangle &= \langle \bar{c}, \bar{a} \times \bar{b} \rangle = \langle \bar{b}, \bar{c} \times \bar{a} \rangle \\ &= -\langle \bar{a}, \bar{c} \times \bar{b} \rangle \\ &= \langle \bar{a} \times \bar{b}, \bar{c} \rangle. \end{aligned}$$

(3) $$\langle t\bar{a}, \bar{b} \times \bar{c} \rangle = \langle \bar{a}, (t\bar{b}) \times \bar{c} \rangle = \langle \bar{a}, \bar{b} \times t\bar{c} \rangle = t\langle \bar{a}, \bar{b} \times \bar{c} \rangle.$$

(The mixed product is linear in each of its three arguments).

(4) $$\langle \bar{u} + \bar{v}, \bar{b} \times \bar{c} \rangle = \langle \bar{u}, \bar{b} \times \bar{c} \rangle + \langle \bar{v}, \bar{b} \times \bar{c} \rangle.$$

$$(5) \qquad \langle \bar{a}\times\bar{b}, \bar{c}\times\bar{d}\rangle = \begin{vmatrix} \langle \bar{a},\bar{c}\rangle & \langle \bar{a},\bar{d}\rangle \\ \langle \bar{b},\bar{c}\rangle & \langle \bar{b},\bar{d}\rangle \end{vmatrix} \qquad \textit{(the Lagrange identity.)}$$

Proof. We will prove only (5) leaving the other proofs to the reader. Note that (1) is a straightforward computation using the formula for $\bar{b}\times\bar{c}$ and the known products of $\bar{i},\bar{j},\bar{k}$; (2), (3), (4) are easy consequences of (1). For (5) denote $\bar{w}=\bar{a}\times\bar{b}$. Then

$$\begin{aligned} \langle \bar{w}, \bar{c}\times\bar{d}\rangle &= \langle \bar{d}, \bar{w}\times\bar{c}\rangle = \\ &= -\langle \bar{d}, \bar{c}\times\bar{w}\rangle \end{aligned} \qquad \text{by (2)}$$

But

$$\bar{c}\times\bar{w} = \bar{c}\times(\bar{a}\times\bar{b}) = \begin{vmatrix} \bar{a} & \bar{b} \\ \langle \bar{c},\bar{a}\rangle & \langle \bar{c},\bar{b}\rangle \end{vmatrix}$$

by the formula for the double vector product. We replace and get

$$\begin{aligned} \langle \bar{w}, \bar{c}\times\bar{d}\rangle &= -\langle \bar{d}, \langle \bar{c},\bar{b}\rangle\bar{a} - \langle \bar{c},\bar{a}\rangle\bar{b}\rangle \\ &= -\langle \bar{c},\bar{b}\rangle\langle \bar{d},\bar{a}\rangle + \langle \bar{c},\bar{a}\rangle\langle \bar{d},\bar{b}\rangle \end{aligned} = \begin{vmatrix} \langle \bar{a},\bar{c}\rangle & \langle \bar{a},\bar{d}\rangle \\ \langle \bar{b},\bar{c}\rangle & \langle \bar{b},\bar{d}\rangle \end{vmatrix}.$$

A basis $\{\bar{a},\bar{b},\bar{c}\}$ of $\mathbf{V}_3$ is said to have *positive orientation* if $\langle \bar{a}, \bar{b}\times\bar{c}\rangle$ is strictly positive, and *negative orientation* if $\langle \bar{a}, \bar{b}\times\bar{c}\rangle$ is strictly negative.

The canonical basis $\{\bar{i},\bar{j},\bar{k}\}$ has positive orientation since $\langle \bar{i}, \bar{j}\times\bar{k}\rangle = 1$.

8 Problems

1. Let $A(1,2,0),\ B(-1,0,3),\ C(2,1,-1)$ be points in $\mathbf{E}_3$. Are these points collinear? Find the area of the triangle ABC, and the altitude from the base BC to the vertex A.

2. Given the points $A(1,1,-2),\ B(2,3,0),\ C(0,1,1),\ D(-1,2,-3)$, compute:
1) the mixed product $\langle\overline{AB},\overline{AC}\times\overline{AD}\rangle$; are the points coplanar?
2) the volume of the tetrahedron $ABCD$;
3) the altitude of the tetrahedron, from the base ACD to the vertex B.

3. Show that

$$\bar{a}\times(\bar{b}\times\bar{c})+\bar{b}\times(\bar{c}\times\bar{a})+\bar{c}\times(\bar{a}\times\bar{b})=\bar{0}$$

$$\langle\bar{a}\times(\bar{b}\times\bar{c}),\ (\bar{b}\times(\bar{c}\times\bar{a}))\times(\bar{c}\times(\bar{a}\times\bar{b}))\rangle\ >0.$$

4. Solve the equations: 1) $\langle\bar{a},\bar{x}\rangle=\alpha$, 2) $\bar{a}\times\bar{x}=\bar{b}$.

5. Find the volume of the parallelipiped constructed on some representatives with common origin of the vectors $\bar{a}=2\bar{u}-\bar{v}+\bar{w}$, $\bar{b}=\bar{u}-\bar{w}$, $\bar{c}=\bar{u}+\bar{w}$, where

$$\|\bar{u}\|=1,\ \|\bar{v}\|=2,\ \|\bar{w}\|=3,\ \angle(\bar{u},\bar{v})=\frac{\pi}{2},\ \angle(\bar{u},\bar{w})=\frac{\pi}{3},\ \angle(\bar{v},\bar{w})=\frac{\pi}{4}.$$

6. Consider the vectors $\bar{a}=\lambda\bar{i}+4\bar{j}+6\bar{k}$, $\bar{b}=\bar{i}+\lambda\bar{j}+3\bar{k}$, $\bar{c}=\lambda\bar{i}+4\bar{j}$. If the vectors are coplanar, determine the possible values of λ. In this case, decompose $\bar{a}$ along $\bar{b}$ and $\bar{c}$.

7. Show that if $\bar{a}\times\bar{b}+\bar{b}\times\bar{c}+\bar{c}\times\bar{a}=\bar{0}$, then the vectors $\bar{a},\bar{b},\bar{c}$ are coplanar.

8. Let $\bar{a},\bar{b},\bar{c}\in\mathbf{V}_3$ be noncoplanar vectors such that $\langle\bar{a},\bar{b}\rangle\neq 0$. Compute

$$E=\left\langle\frac{\bar{a}\times\bar{b}}{\langle\bar{a},\bar{b}\rangle},\frac{\bar{a}\times(\bar{b}\times\bar{c})}{\langle\bar{a},\bar{b}\times\bar{c}\rangle}\right\rangle+2.$$

9. Given the points $A(4,-2,2),\ B(3,1,1),\ C(4,2,0)$, determine the vertex D of the tetrahedron $ABCD$ such that $D\in Oz$ and the volume of $ABCD$ is 4. Determine also the altitude from the base ABC to D.

Chapter 6

Straight Lines and Planes in Space

1 Cartesian Frames

We mentioned that a fixed origin $O \in \mathbf{E}_3$ provides a natural one-to-one correspondence between $\mathbf{E}_3$ and $\mathbf{V}_3$, assigning to each point $M \in \mathbf{E}_3$ a unique vector $\bar{r} = \overline{OM} \in \mathbf{V}_3$, called the *position vector* of M. Each fixed basis of $\mathbf{V}_3$ determines a bijective correspondence between $\mathbf{V}_3$ and $\mathbf{R}^3$.

Throughout this chapter, $O \in \mathbf{E}_3$ will be a fixed origin and $\{\bar{i}, \bar{j}, \bar{k}\}$ a fixed orthonormal basis of $\mathbf{V}_3$. The set $\{O;\ \bar{i}, \bar{j}, \bar{k}\}$ is called a *Cartesian frame in* $\mathbf{E}_3$. The point O is called the *origin* of the frame, and $\{\bar{i}, \bar{j}, \bar{k}\}$ is the *basis of the frame*. The Euclidean coordinates of the position vector $\overline{OM}$ are also called *the Cartesian coordinates of the point* M with respect to the orthonormal frame $\{O;\ \bar{i}, \bar{j}, \bar{k}\}$. If $\overline{OM} = x\bar{i} + y\bar{j} + z\bar{k}$, we write $M(x, y, z)$.

The bijection between $\mathbf{E}_3$ and $\mathbf{R}^3$ determined by a fixed Cartesian frame is called *a Cartesian coordinate system in* $\mathbf{E}_3$.

The bijections mentioned above allow the identification of $\mathbf{E}_3$, $\mathbf{V}_3$ and $\mathbf{R}^3$.

The basis $\{\bar{i}, \bar{j}, \bar{k}\}$ will also be assumed to have positive (or right-handed) orientation in $\mathbf{V}_3$, i.e. $\langle \bar{i}, \bar{j} \times \bar{k} \rangle = 1$, or equivalently $\bar{i} \times \bar{j} = \bar{k}$.

The oriented lines Ox, Oy, Oz passing through O, of direction and sense given by $\bar{i}, \bar{j}, \bar{k}$ respectively are called *the Cartesian axes of the frame* $\{O;\ \bar{i}, \bar{j}, \bar{k}\}$. The Cartesian coordinates of the point M are the algebraic measures of the orthogonal projections of $\overline{OM}$ onto the coordinate axis (Fig. 20).

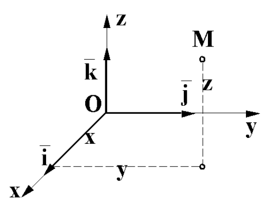

Fig. 20

The axes are characterized by the equations

$$Ox : \begin{cases} y = 0 \\ z = 0 \end{cases}, \qquad Oy : \begin{cases} z = 0 \\ x = 0 \end{cases}, \qquad Oz : \begin{cases} x = 0 \\ y = 0 \end{cases}.$$

Any two axes determine a plane called a *coordinate plane.* The coordinate planes xOy, yOz, zOx are characterized by the equations

$$xOy : \; z = 0, \qquad yOz : \; x = 0, \qquad zOx : \; y = 0.$$

2 Equations of Straight Lines in Space

A straight line in $\mathbf{E}_3$ may be determined by either:
– a point and a nonzero vector,
– two distinct points,
– the intersection of two planes, etc.

2.1. The Line Determined by a Point and a Nonzero Vector

The point $M_0(x_0, y_0, z_0)$ and $\bar{a} = l\bar{i} + m\bar{j} + n\bar{k} \in \mathbf{V}_3 \setminus \{\bar{0}\}$ determine a straight line D passing through M_0, and having the direction of $\bar{a}$ (Fig. 21).

Let $M(x, y, z) \in \mathbf{E}_3$, $\bar{r} = \overline{OM}$, $\bar{r}_0 = \overline{OM_0}$. Then $M \in D \Leftrightarrow \overline{M_0M}$ and $\bar{a}$ are collinear i.e.

(2.1) $$(\bar{r} - \bar{r}_0) \times \bar{a} = \bar{0} \quad \textit{(the vector equation of } D\textit{)}$$

The vector $\bar{a}$ is called a *director vector of* D. Any vector of the form $k\bar{a}$, $k \in \mathbf{R}\setminus\{\bar{0}\}$ is another director vector of D and may be used to obtain the equations of D, instead of using $\bar{a}$.

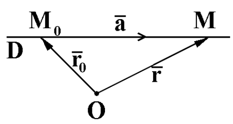

Fig. 21

The position vector $\bar{r} = \overline{OM}$ of an arbitrary point $M \in D$ is of the form

(2.2) $$\bar{r} = \bar{r}_0 + t\bar{a}, \qquad t \in \mathbf{R},$$

which is another equivalent condition to the collinearity of $\bar{r} - \bar{r}_0$ and $\bar{a}$. The vector equation of D is equivalent to

(2.3) $$\begin{cases} x = x_0 + lt \\ y = y_0 + mt \\ z = z_0 + nt \end{cases}, \qquad t \in \mathbf{R} \quad \textit{(the parametric equations of } D \textit{ in } \mathbf{R}^3\textit{)}$$

or to

(2.4) $$\frac{x - x_0}{l} = \frac{y - y_0}{m} = \frac{z - z_0}{n} \quad \textit{(the canonical Cartesian equations of } D \textit{ in } \mathbf{R}^3\textit{)},$$

with the convention that if a denominator is zero , then the corresponding numerator is zero too. More precisely,

(1) if $l = 0$, $mn \neq 0$, then

$$x = x_0, \quad \frac{y - y_0}{m} = \frac{z - z_0}{n}$$

and $D \parallel yOz$;

(2) if $l = m = 0$, then

$$x = x_0, \qquad y = y_0$$

and $D \parallel Oz$.

Note that $\bar{a} \neq \bar{0}$, thus at most two of the coordinates of $\bar{a}$ could possibly be zero.

2.2 The Line Determined by Two Distinct Points

Let $M_1(x_1, y_1, z_1) \neq M_2(x_2, y_2, z_2) \in \mathbf{E}_3$, and D =the straight line M_1M_2. Then $\bar{a} = \overline{M_1M_2} = (x_2 - x_1)\bar{i} + (y_2 - y_1)\bar{j} + (z_2 - z_1)\bar{k}$ is a director vector of D. In the previous equations we may replace $\bar{a}$ by $\overline{M_1M_2}$ and M_0 by M_1 (Fig. 22). For example the canonical Cartesian equations (2.4) become

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1}. \tag{2.5}$$

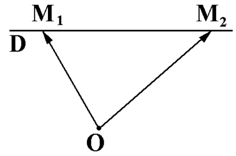

Fig. 22

3 Equations of Planes in Space

A plane in $\mathbf{E}_3$ may be determined by either:

- a point and a normal direction,
- three noncollinear points,
- a straight line and an exterior point,
- two concurrent straight lines,
- two parallel straight lines, etc.

3.1. The Plane Determined by a Point and a Nonzero Vector

The plane determined by $M_0(x_0, y_0, z_0)$ and a nonzero vector $\bar{n} = a\bar{i} + b\bar{j} + c\bar{k}$ is the unique plane P passing through M_0 such that the line D through M_0 of direction $\bar{n}$ is perpendicular onto P (Fig. 23).

$\bar{n}$ is called *a normal vector* to the plane P.

D is *normal line* of P.

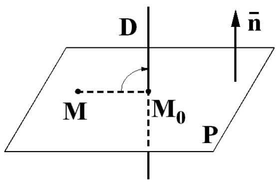

Fig. 23

Let $M(x, y, z)$. Then

$$M \in P \Leftrightarrow \overline{M_0M} \perp \bar{n} \Leftrightarrow \langle \overline{M_0, M}, \bar{n} \rangle = 0.$$

The last equality is called the *vector equation of the plane*. Replacing $\overline{M_0M}$ and $\bar{n}$ by their analytic expressions it follows that

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \tag{3.1}$$

(the Cartesian equation in $\mathbf{R}^3$ *of the plane passing through* M_0*, perpendicular to* $\bar{n}$*).*

If we denote $d = -(ax_0 + by_0 + cz_0)$, the previous equation becomes

$$ax + by + cz + d = 0 \quad (\textit{the general Cartesian equation of a plane}). \tag{3.2}$$

Conversely, any equation of the form (3.2) with $a^2 + b^2 + c^2 \neq 0$ represents a plane of normal vector $\bar{n} = a\bar{i} + b\bar{j} + c\bar{k}$.

3.2. Particular Planes

1) The coordinate planes and planes parallel to these planes are described by the following particular forms of (3.2). We start with the coordinate planes,

$$xOy : z = 0; \quad yOz : \ x = 0; \quad zOx : \ y = 0.$$

$\bar{k}$ is a normal vector for xOy and for any plan P parallel to xOy. Then (Fig. 24)

$$P : z = a, \quad a \in \mathbf{R} \setminus \{0\}, \quad \text{for } P \| xOy.$$

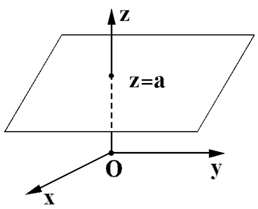

Fig. 24

Similarly, since $\bar{i}$ is normal to yOz and $\bar{j}$ is normal to zOx we obtain

$$P : x = a, \quad a \in \mathbf{R} \setminus \{0\}, \quad \text{for } P \| yOz$$

$$P : y = a, \quad a \in \mathbf{R} \setminus \{0\}, \quad \text{for } P \| zOx.$$

2) The equations of planes perpendicular to the coordinate planes are:

$$P : ax + by + d = 0 \quad a^2 + b^2 \neq 0, \quad \text{if} \quad P \perp xOy$$

$$P : by + cz + d = 0 \quad b^2 + c^2 \neq 0, \quad \text{if} \quad P \perp yOz$$

$$P : ax + cz + d = 0 \quad a^2 + c^2 \neq 0, \quad \text{if} \quad P \perp zOx.$$

3) The equation of a plane passing through $O(0,0,0)$ is

$$ax+by+cz=0.$$

4) The equations of planes containing one of the coordinate axes Oz, Ox, Oy are

$$ax+by=0,\ by+cz=0,\quad \text{and}\quad ax+cz=0 \text{ respectively.}$$

3.3. The Plane Determined by Three Noncollinear Points

Given the noncollinear points $M_i(x_i,y_i,z_i)$, $i=1,2,3$ there exists a unique plane P containing M_1,M_2,M_3 (Fig. 25). If $M(x,y,z)$ is another point, then

$$\begin{aligned} M\in P \quad &\Leftrightarrow \quad M,M_1,M_2,M_3 \quad \text{are coplanar in } \mathbf{E}_3\\ &\Leftrightarrow \quad \overline{M_1M},\overline{M_1M_2},\overline{M_1M_3} \quad \text{are coplanar in } \mathbf{V}_3\\ &\Leftrightarrow \quad \langle\overline{M_1M},\overline{M_1M_2}\times\overline{M_1M_3}\rangle=0 \quad \text{the vector equation of P.}\end{aligned}$$

The above vector equation is equivalent to the equation

$$\begin{vmatrix} x-x_1 & y-y_1 & z-z_1\\ x_2-x_1 & y_2-y_1 & z_2-z_1\\ x_3-x_1 & y_3-y_1 & z_3-z_1\end{vmatrix}=0 \tag{3.3}$$

in $\mathbf{R}^3$ or to

$$\begin{vmatrix} x & y & z & 1\\ x_1 & y_1 & z_1 & 1\\ x_2 & y_2 & z_2 & 1\\ x_3 & y_3 & z_3 & 1\end{vmatrix}=0 \tag{3.4}$$

(the equation of the plane determined by three points; coplanarity of four points).

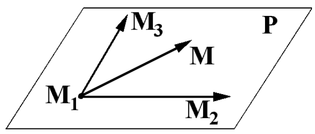

Fig. 25

A particular case of a plane determined by three points is $M_1(a,0,0)\in Ox$, $M_2(0,b,0)\in Oy$, $M_3(0,0,c)\in Oz$, a, b, $c\in\mathbf{R}^*$ (i.e. $0\notin P$ and M_1,M_2,M_3 are the intersections of P with the coordinate axes). Replacing in (3.3) or (3.4) we obtain

$$\frac{x}{a}+\frac{y}{b}+\frac{z}{c}-1=0 \quad \textit{(the intercept form of the equation of a plane, Fig. 26).}$$

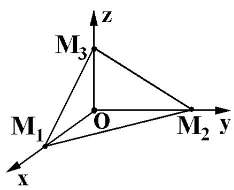

Fig. 26

3.4. The Plane Determined by a Point and Two Noncollinear Vectors

There exists a unique plane P passing through a given point $M_0(x_0, y_0, z_0)$ and containing lines of directions $\bar{v}_1 = l_1\bar{i}+m_1\bar{j}+n_1\bar{k}$ and $\bar{v}_2 = l_2\bar{i}+m_2\bar{j}+n_2\bar{k}$ respectively, where $\bar{v}_1, \bar{v}_2$ are noncollinear (Fig. 27).

Let M_1, M_2 such that $\bar{v}_1 = \overline{M_0M_1}$, $\bar{v}_2 = \overline{M_0M_2}$. Then P is determined by M_0, M_1, M_2. It follows that $M(x, y, z) \in P$ if and only if $\overline{M_0M}$, $\overline{M_0M_1}$, $\overline{M_0M_2}$ are coplanar, this is equivalent to $\overline{M_0M} = r\bar{v}_1 + s\bar{v}_2$ for some $r, s \in \mathbf{R}$. Identifying the coordinates we obtain

$$\text{(3.6)} \qquad \begin{cases} x = x_0 + rl_1 + sl_2 \\ y = y_0 + rm_1 + sm_2 \qquad r, s \in \mathbf{R} \\ z = z_0 + rn_1 + sn_2 \end{cases}$$

(the parametric equations of the plane in $\mathbf{R}^3$*).*

The coplanarity of $\overline{M_0M}$, $\overline{M_0M_1}$, $\overline{M_0M_2}$ is also equivalent to $\langle \overline{M_0M},\ \bar{v}_1 \times \bar{v}_2 \rangle = 0$, i.e.,

$$\text{(3.7)} \qquad \begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \end{vmatrix} = 0.$$

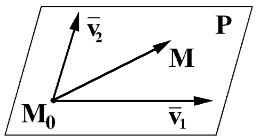

Fig. 27

4 The Intersection of Two Planes

It is known that any two distinct planes are either parallel or they have a common line.

Let $P_i : a_ix + b_iy + c_iz + d_i = 0$. $a_i^2 + b_i^2 + c_i^2 \neq 0$, $i = 1, 2$, $\bar{n}_i = a_i\bar{i} + b_i\bar{j} + c_i\bar{k}$, and consider the system

$$\text{(4.1)} \qquad \begin{cases} a_1x + b_1y + c_1z + d_1 = 0 \\ a_2x + b_2y + c_2z + d_2 = 0\,. \end{cases}$$

The planes P_1, P_2 are identical $\Leftrightarrow$ their equations are equivalent, i.e., $\dfrac{a_1}{a_2} = \dfrac{b_1}{b_2} = \dfrac{c_1}{c_2} = \dfrac{d_1}{d_2}$. This means that the system is compatible of rank 1. The planes are parallel $\Leftrightarrow$ $\bar{n}_1, \bar{n}_2$ are collinear but the system is not compatible, i.e.,

$$\frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2} \neq \frac{d_1}{d_2}.$$

The planes intersect along a straight line D $\Leftrightarrow$ the system (4.1) is compatible of rank 2. In this case the equations (4.1) represent the line $D = P_1 \cap P_2$.

The set of all planes passing through the straight line D is called the *pencil of planes determined by P_1 and P_2*. The line D is called the *axes of the pencil (Fig. 28).*

The equation of an arbitrary plane of the pencil is

$$(4.2) \quad r(a_1x + b_1y + c_1z + d_1) + s(a_2x + b_2y + c_2z + d_2) = 0, \ s, t \in \mathbf{R}, \ s^2 + t^2 \neq 0,$$

which also said to be the equation of the pencil.

The set of all planes parallel or coincident to a given plane P_1 is called a *pencil of parallel* planes, whose equations are of the form:

$$a_1x + b_1y + c_1z + \lambda = 0, \quad \lambda \in \mathbf{R}.$$

Note that from (4.1) one can deduce easily the equations of the line $D = P_1 \cap P_2$ in any of the forms studied previously, using the director vector $\bar{a} = \bar{n}_1 \times \bar{n}_2$ and $M_0(x_0, y_0, z_0)$, where (x_0, y_0, z_0) is a particular solution of the system (4.1)

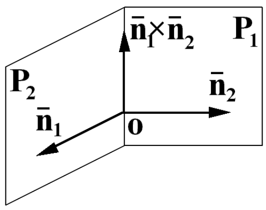

Fig. 28

5 Orientation of Straight Lines and Planes

5.1. Oriented Straight Lines

An *oriented straight line D* is a straight line together with a chosen sense of way along D. Choosing a sense on D is equivalent to choosing a director vector $\bar{a}$ of D. We say that the pair $(D, \bar{a})$ represents an oriented line.

Any line admits two orientations: the one given by $\bar{a}$ (or by any vector $k\bar{a}$, $k > 0$) and the orientation given by $-\bar{a}$ (or by any vector $k\bar{a}$, $k < 0$). If $\bar{a}$ is fixed, we agree to call the orientation given by $\bar{a}$ the *positive* sense on D, and the other the *negative sense* on D. For example the coordinate axes Ox, Oy, Oz are oriented by $\bar{i}, \bar{j}$ and $\bar{k}$ respectively.

The orientation given by $\bar{a}$ is uniquelly determined by the versor $\bar{e} = \dfrac{\bar{a}}{\|\bar{a}\|}$. Therefore an oriented straight line may be seen as a pair $(D, \bar{e})$, where D is a line and $\bar{e}$ one of the two versors of direction D (Fig. 29).

The angles $\alpha = \angle(\bar{e}, \bar{i})$, $\beta = \angle(\bar{e}, \bar{j})$, $\gamma = \angle(\bar{e}, \bar{k})$ are called the *director angles* of $(D, \bar{e})$, where

$$\bar{e} = \cos\alpha\bar{i} + \cos\beta\bar{j} + \cos\gamma\bar{k}.$$

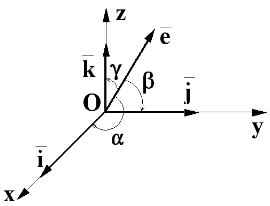

Fig. 29

The numbers $\cos\alpha, \cos\beta, \cos\gamma$ are called the *director cosines* of $(D, \bar{e})$. From $\|\bar{e}\| = 1$ follows

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1.$$

If $\bar{e} = \dfrac{\bar{a}}{\|\bar{a}\|}$, $\bar{a} = l\bar{i} + m\bar{j} + n\bar{k}$, then

$$\cos\alpha = \frac{l}{\sqrt{l^2+m^2+n^2}}, \quad \cos\beta = \frac{m}{\sqrt{l^2+m^2+n^2}}, \quad \cos\gamma = \frac{n}{\sqrt{l^2+m^2+n^2}}.$$

5.2. Orientation of a Plane. Semispaces

An *oriented plane* is a plane together with a choice of a normal vector i.e. a pair $(P, \bar{n})$. Intuitively, a plane has two faces; the face which corresponds to the sense of a fixed normal vector $\bar{n}$ is denoted by "+", and the opposite one denoted by "-". Of course, a plane may be oriented according to the right hand rule.

Any plane admits two orientations.

Let $\bar{n} = a\bar{i} + b\bar{j} + c\bar{k}$ be a fixed normal vector of $P : ax + by + cz + d = 0$. Denote $f(x, y, z) = ax + by + cz + d$. Then $\mathbf{E}_3 = S^+ \cup P \cup S^-$ where (Fig. 30)

$$S^+ = \{(x, y, z) | f(x, y, z) > 0\}, \quad S^- = \{(x, y, z) | f(x, y, z) < 0\}$$

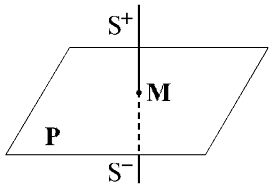

Fig. 30

6 Angles in Space

6.1. The Angle between Two Oriented Lines

If D_1, D_2 are two lines oriented by $\bar{a} = l_1\bar{i} + m_1\bar{j} + n_1\bar{k}$ and $\bar{b} = l_2\bar{i} + m_2\bar{j} + n_2\bar{k}$ respectively, the *angle of* $(D_1, \bar{a})$ and $(D_2, \bar{b})$ is the angle $\theta = \angle(\bar{a}, \bar{b})$, given by (Fig. 31)

$$\cos\theta = \frac{\langle \bar{a}, \bar{b} \rangle}{\|\bar{a}\| \cdot \|\bar{b}\|} = \frac{l_1 l_2 + m_1 m_2 + n_1 n_2}{\sqrt{l_1^2 + m_1^2 + n_1^2} \cdot \sqrt{l_2^2 + m_2^2 + n_2^2}}, \ \theta \in [0, \pi].$$

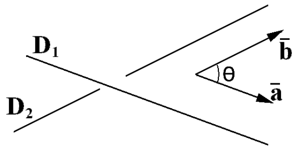

Fig. 31

Note that
(1) $D_1 \perp D_2 \Leftrightarrow \langle \bar{a}, \bar{b} \rangle = 0 \Leftrightarrow l_1 l_2 + m_1 m_2 + n_1 n_2 = 0$
(2) $D_1 \| D_2 \Leftrightarrow \bar{a} \times \bar{b} = 0, \quad D_1 \neq D_2$

$$\Leftrightarrow \frac{l_1}{l_2} = \frac{m_1}{m_2} = \frac{n_1}{n_2},\ D_1 \neq D_2.$$

6.2. The Angle between Two Oriented Planes

The dihedral angle of $(P_1, \bar{n}_1)$ and $(P_2, \bar{n}_2)$ is the angle of their normals $\bar{n}_1, \bar{n}_2$ (Fig. 32)

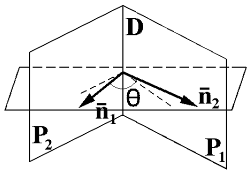

Fig. 32

Denote this angle by θ. Then

$$\cos\theta = \frac{\langle \bar{n}_1, \bar{n}_2 \rangle}{\|n_1\| \cdot \|n_2\|} = \frac{a_1 a_2 + b_1 b_2 + c_1 c_2}{\sqrt{a_1^2 + b_1^2 + c_1^2} \cdot \sqrt{a_2^2 + b_2^2 + c_2^2}},\ \theta \in [0, \pi],$$

where $\bar{n}_i = a_i \bar{i} + b_i \bar{j} + c_i \bar{k},\ i = 1, 2.$

6.3 The Angle between an Oriented Straight Line and an Oriented Plane

Let $(D, \bar{a})$ be an oriented line and $(P, \bar{n})$ an oriented plane (Fig. 33).

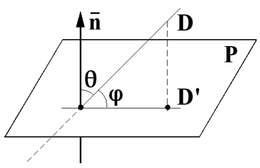

Fig. 33

The angle between $(D, \bar{a})$ and $(P, \bar{n})$ is the angle φ between $\bar{a}$ and the projection of $\bar{a}$ onto P. It turns out that $\theta + \varphi = 90°$ where $\theta = \angle(\bar{a}, \bar{n})$. Then

$$\sin\varphi = \cos\theta = \frac{\langle \bar{a}, \bar{n} \rangle}{\|\bar{a}\| \cdot \|\bar{n}\|} = \frac{al + bm + cn}{\sqrt{a^2 + b^2 + c^2} \cdot \sqrt{l^2 + m^2 + n^2}}.$$

$D \| P$ or $D \subset P \Leftrightarrow \langle \bar{a}, \bar{n} \rangle = 0,$

$D \perp P \qquad \Leftrightarrow \bar{a} \times \bar{n} = \bar{0} \ \Leftrightarrow \ \dfrac{a}{l} = \dfrac{b}{m} = \dfrac{c}{n}.$

7 Distances in Space

7.1. The Distance from a Point to a Straight Line

The distance from a point A to the line D passing through $M_0(x_0, y_0, z_0)$, and of director vector $\bar{a}$ is

$$d(A;D) = \frac{\|\bar{a} \times \overline{M_0A}\|}{\|\bar{a}\|}$$

since $d(A;D) = AA'$ is the height of a parallelogram constructed on $\bar{a}$ and $\overline{M_0A}$ (if $A \notin D$) corresponding to the basis of length $\|\bar{a}\|$. If $A \in D$, the formula is still true since $d(A;D) = 0$.

7.2. The Distance from a Point to a Plane

Let $M_0(x_0, y_0, z_0)$ and $P : ax + by + cz + d = 0$, $M_0 \notin P$ (Fig. 34).

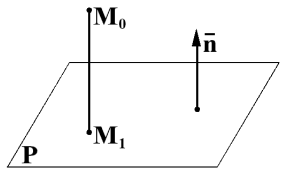

Fig. 34

If M_1 is the projection of M_0 onto P, then $d(M_0, P) = \|\overline{M_1M_0}\|$.

The collinearity of $\bar{n} = a\bar{i} + b\bar{j} + c\bar{k}$ and $\overline{M_1M_0}$ implies $\cos(\bar{n}, \overline{M_1M_0}) = \pm 1$. This means

$$\frac{|\langle \bar{n}, \overline{M_1M_0} \rangle|}{\|\bar{n}\| \cdot \|\overline{M_1M_0}\|} = 1.$$

Then

$$d(M_0, P) = \|\overline{M_1M_0}\| = \frac{|\langle \bar{n}, \overline{M_1M_0} \rangle|}{\|\bar{n}\|} = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

The formula still works when $M_0 \in P$, i.e., $d(M_0, P) = 0$.

7.3. The Common Perpendicular of two Noncoplanar Lines

Consider the straight lines D_1 and D_2 of director vectors $\bar{a}_1$ and $\bar{a}_2$ respectively. D_1 and D_2 are noncoplanar if and only if $D_1 \cap D_2 = \emptyset$ and $\bar{a}_1$, $\bar{a}_2$ are noncollinear. In this case there exists a unique straight line D such that $D \perp D_1$, $D \perp D_2$ and $D \cap D_1 \neq \emptyset$, $D \cap D_2 \neq \emptyset$. The line D is called *the common perpendicular of* D_1 *and* D_2 (Fig. 35).

Obviously, a director vector of D is $\bar{n} = \bar{a}_1 \times \bar{a}_2$.

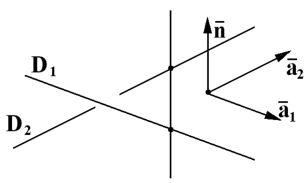

Fig. 35

In order to find out the equations of D we may consider the plane P_1 determined by D and D_1, and the plane P_2 determined by D and D_2; since $D = P_1 \cap P_2$, we are going to use the equations of P_1 and P_2 (see Section 4, (4.1)).

For $i = 1, 2$ let M_i be an arbitrary fixed point on D_i; it turns out that P_i is the plane determined by M_i and the noncollinear vectors $\bar{n}$ and $\bar{a}_i$. As we have seen at the end of Section 3, it follows that

$$P_i : \langle \overline{M_i M}, \bar{a}_i \times \bar{n} \rangle = 0, \qquad i = 1, 2,$$

where M is the current point of P_i.

Consequently, *the equations of the common perpendicular* are

$$D : \begin{cases} \langle \overline{M_1 M},\, \bar{a}_1 \times \bar{n} \rangle = 0 \\ \langle \overline{M_2 M},\, \bar{a}_2 \times \bar{n} \rangle = 0 \end{cases},$$

where M is the current point of D.

7.4. The Distance between two Straight Lines

Let D_1, D_2 be two straight lines. It is known that the number $inf\{d(M,N) \mid M \in D_1,\, N \in D_2\}$ is called the *the distance between D_1 and D_2* and it is denoted by $d(D_1, D_2)$. It is also known that $d(D_1, D_2)$ may be computed as follows (Fig. 36):

1) if $D_1 \cap D_2 \neq \emptyset$, then $d(D_1, D_2) = 0$;

2) if $D_1 || D_2$, pick a point $M_1 \in D_1$; then $d(D_1, D_2) = d(M_1, D_2)$;

3) if D_1 and D_2 are noncoplanar, then $d(D_1, D_2) = d(A, B) = \|\overline{AB}\|$, where $\{A\} = D \cap D_1$, $\{B\} = D \cap D_2$, and D is the common perpendicular of D_1 and D_2.

Using the equations of D found previously we may compute the coordinates of A and B for the given lines D_1, D_2. However it is more convenient to regard $\|\overline{AB}\|$ as the height of a certain parallelepiped in the following way.

Let M_1, M_2 be arbitrary fixed points on D_1 and D_2 respectively. Consider the straight line D_2' passing through M_1, parallel to D_2. Denote by Q_1 the plane determined by D_1 and D_2'.

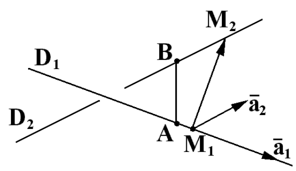

Fig. 36

Then Q_1 is the plane parallel to D_2, passing through D_1, and $\|\overline{AB}\|$ is the distance from B to Q_1. But $d(B, Q_1) = d(M_2, Q_1) = d(D_2, Q_1)$, since $D_2 \| Q_1$. On the other hand $d(M_2, Q_1)$ is the height of the parallelepiped constructed on the supports of the vectors $\overline{M_1 M_2}$, $\bar{a}_1$, $\bar{a}_2$, that corresponds to the basis determined by $\bar{a}_1$, $\bar{a}_2$. Then

$$d(D_1, D_2) = \frac{|\langle \overline{M_1 M_2},\, \bar{a}_1 \times \bar{a}_2 \rangle|}{\|\bar{a}_1 \times \bar{a}_2\|}.$$

8 Problems

1. Write the equations of a straight line passing through the point $A(1, 1, -2)$ and parallel to the straight line D.

1) $D : \dfrac{x-4}{2} = \dfrac{y+2}{3} = \dfrac{z+1}{4}$

2) $D : \begin{cases} x - y - 3z + 2 = 0 \\ 2x - y + 2z - 3 = 0 \end{cases}$

2. Compute the distance from the point $A(1, 1, 1)$ to the straight line D.

1) $D : \dfrac{x-1}{2} = \dfrac{y+1}{1} = \dfrac{z-1}{3}$

2) $D : \begin{cases} x - y + z = 0 \\ x + y - z = 0 \end{cases}$

3. Write the equation of the plane passing through $A(1, 1, -1)$ and perpendicular to D.

1) $D : \dfrac{x}{1} = \dfrac{y-1}{2} = \dfrac{z+1}{-1}$

2) $D : \begin{cases} x - y = 0 \\ x + 2y - z + 1 = 0 \end{cases}$

4. Write the equation of the plane which passes through the point $A(0, 1, -1)$ and through the straight line D given as follows:

1) $D : x = 4 + 2t,\ y = -2 + 3t,\ z = -1 + 4t;$

2) $D : \begin{cases} 2x - y + z + 1 = 0 \\ x + y + z = 0. \end{cases}$

5. Consider the planes

$$P : x - 2y + 2z - 7 = 0,\ Q : 2x - y - 2z + 1 = 0,\ R : 2x + 2y + z - 2 = 0.$$

1) Show that the planes are pairwise perpendicular.
2) Find the common point of these three planes.
3) Find the distance from $A(2, 4, 7)$ to the plane P.

6. Determine the projection of the straight line D onto the plane $P : 2x + 3y + 4z - 10 = 0$, where D is given by:

1) $D : \begin{cases} x - y - 3z + 2 = 0 \\ 2x - y + 2z - 3 = 0; \end{cases}$

2) $D : \dfrac{x-4}{2} = \dfrac{y+2}{3} = \dfrac{z+1}{4};$

3) $D : \dfrac{x-1}{3} = \dfrac{y+2}{-2} = \dfrac{z+1}{0}.$

7. Consider the straight lines D_1, D_2 of equations

$$D_1 : \frac{x}{1} = \frac{y}{2} = \frac{z}{3}, \quad D_2 : \frac{x-1}{2} = \frac{y-1}{-1} = \frac{z}{1}.$$

Prove that D_1, D_2 are noncoplanar, determine the equations of their common perpendicular, and the distance between D_1 and D_2.

8. Write down the equation of the plane P which contains the point $A(3, -1, 2)$ and the line

$$D : \begin{cases} 2x - y - 3z - 2 = 0 \\ x + 3y - z + 4 = 0. \end{cases}$$

Determine also the distance from A to D.

9. Consider the point $A(1, -2, 5)$ and the plane $Q : x + 2y + 2z + 1 = 0$. Find the projection of A onto Q, the symmetric of A w.r.t. Q, and the distance from A to Q.

10. Given the points $A(3, -1, 3)$, $B(5, 1, -1)$, $C(0, 4, -3)$, determine:
1) the parametric equations of the straight lines AB and AC
2) the angle between AB and AC
3) the distance from A to the straight line BC.

11. If the point $M(3, 4, 2)$ is the projection of the origin onto the plane P, determine the equation of P.

12. Determine the parameters λ, μ such that the planes

$$P : 2x - y + 3z - 1 = 0, \; Q : x + 2y - z + \mu = 0, \; R : x + \lambda y - 6z + 10 = 0$$

1) have exactly one common point;
2) have a common straight line;
3) intersect about three parallel and distinct straight lines.

13. Consider the points $A(1, 3, 2)$, $B(-1, 2, 1)$, $C(0, 1, -1)$, $D(2, 0, -1)$, and the plane $P : 2x + y - z - 1 = 0$. Which of the given points is on the same side as the origin, with respect to the plane P?

Chapter 7

Transformations of Coordinate Systems

Changes of Cartesian frames are related to isometries of $\mathbf{R}^3$. Using the group of isometries of $\mathbf{R}^3$ we may define in a natural way the congruence of space figures in $\mathbf{E}_3$. On the other hand, isometries may be described geometrically. It turns out that the fundamental isometries are: rotation, translation, symmetry w.r.t. a plane, symmetry w.r.t. a point, and translation. Any isometry is a composite of isometries listed above.

Rotations and symmetries are called *orthogonal transformations* and they actually correspond to linear orthogonal transformations on $\mathbf{V}_3 \simeq \mathbf{R}^3$.

As we have seen, any isometry is the product of a translation and an orthogonal transformation.

Let $\mathcal{I} = \mathcal{T} \circ \mathcal{R}$ be an isometry determined by the frames $\mathbf{F} = \{O, \bar{i}, \bar{j}, \bar{k}\}$ and $\mathbf{F}' = \{O'; \bar{i}', \bar{j}', \bar{k}'\}$. The isometry $\mathcal{I}$ is said to be *positive (displacement)* if the basis $\{\bar{i}, \bar{j}, \bar{k}\}$ has positive orientation, and *negative (antidisplacement)* otherwise.

Translations and rotations are positive isometries; symmetries are negative isometries.

1 Translations of Cartesian Frames

We say that the Cartesian frame $O'x'y'z'$ is obtained by the translation of the Cartesian frame $Oxyz$, if the axes of the new frame $O'x'y'z'$ are parallel and of the same sense as the axes of the initial frame (Fig. 37).

The translation $\mathcal{T}$ is described by:

$$O' = \mathcal{T}(O),\ \bar{i}' = \mathcal{T}(\bar{i}) = \bar{i},\ \bar{j}' = \mathcal{T}(\bar{j}) = \bar{j},\ \bar{k}' = \mathcal{T}(\bar{k}) = \bar{k}.$$

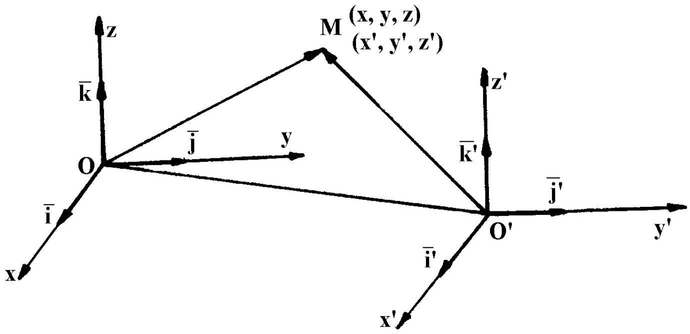

Fig. 37

If the coordinates of the new origin w.r.t. the initial frame are $O'(a, b, c)$, let us determine now the relationship between the coordinates x, y, z and x', y', z' of the same point M w.r.t each coordinate system. For, note that $\overline{OM} = \overline{OO'} + \overline{O'M}$. In terms of the basis $\{\bar{i}, \bar{j}, \bar{k}\}$, this relation becomes

$$x\bar{i} + y\bar{j} + z\bar{k} = a\bar{i} + b\bar{j} + c\bar{k} + x'\bar{i} + y'\bar{j} + z'\,\bar{k},$$

thus

$$x = x' + a,\ y = y' + b,\ z = z' + c.$$

In matrix form we obtain

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} + \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}.$$

Particular case. A translation in the plane xOy is described by

$$x = x' + a,\ y = y' + b.$$

2 Rotations of Cartesian Frames

Rotating the Cartesian frame $\{O; \bar{i}, \bar{j}, \bar{k}\}$ means passing from this frame to a new one $\{O;\ \bar{i}', \bar{j}', \bar{k}'\}$, keeping the same origin O and changing the orthonormal basis $\{\bar{i}, \bar{j}, \bar{k}\}$ of $\mathbf{V}_3$ into another orthonormal basis $\{\bar{i}', \bar{j}', \bar{k}'\}$, such that $\{\bar{i}', \bar{j}', \bar{k}'\}$ has positive orientation. If the coordinates of the new basis vectors w.r.t. the old basis are known (i.e. the matrix of change is known), then we can express a relationship between the coordinates x, y, z and x', y', z' of the same point M w.r.t each coordinate system. This is immediate since

$$\overline{OM} = x\bar{i} + y\bar{j} + z\bar{k} = x'\bar{i}' + y'\bar{j}' + z'\bar{k}'.$$

Denote the matrix of change by $R = [a_{ij}] \in M_{3,3}(\mathbf{R})$. Then

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}. \tag{2.1}$$

Since both bases are orthonormal, the linear transformation sending $\{\bar{i},\bar{j},\bar{k}\}$ to $\{\bar{i}',\bar{j}',\bar{k}'\}$ is an orthogonal transformation $\mathcal{R}$ of $\mathbf{V}_3$ whose associated matrix w.r.t. $\{\bar{i},\bar{j},\bar{k}\}$ is R. The entries of R can be expressed in terms of inner products of the bases vectors (see Chapter 1, section 7) as follows

$$\mathcal{R}(\bar{i}) = \bar{i}' = \langle \bar{i}', \bar{i}\rangle \bar{i} + \langle \bar{i}', \bar{j}\rangle \bar{j} + \langle \bar{i}', \bar{k}\rangle \bar{k}$$

$$\mathcal{R}(\bar{j}) = \bar{j}' = \langle \bar{j}', i\rangle \bar{i} + \langle \bar{j}', \bar{j}\rangle \bar{j} + \langle \bar{j}', \bar{k}\rangle \bar{k}$$

$$\mathcal{R}(\bar{k}) = \bar{k}' = \langle \bar{k}', \bar{i}\rangle \bar{i} + \langle \bar{k}', \bar{j}\rangle \bar{j} + \langle \bar{k}', \bar{k}\rangle \bar{k}.$$

Then

$$a_{11} = \langle \bar{i}', \bar{i}\rangle,\ a_{21} = \langle \bar{i}', \bar{j}\rangle,\ a_{31} = \langle \bar{i}', \bar{k}\rangle,$$
$$a_{12} = \langle \bar{j}', \bar{i}\rangle,\ a_{22} = \langle \bar{j}', \bar{j}\rangle,\ a_{32} = \langle \bar{j}', \bar{k}\rangle,$$
$$a_{13} = \langle \bar{k}', \bar{i}\rangle,\ a_{23} = \langle \bar{k}', \bar{j}\rangle,\ a_{33} = \langle \bar{k}', \bar{k}\rangle.$$

Since $\bar{i}', \bar{j}', \bar{k}'$ are pairwise orthonormal unit vectors, their coordinates represent director cosines, which implies that R is an orthogonal matrix, i.e. $R\,{}^tR = I = {}^tRR$, or $R^{-1} = {}^tR$. (As we mentioned above, it is a general fact that a linear transformation of a Euclidean space, which sends an orthonormal basis to another orthonormal basis, is an orthogonal transformation, therefore its associated matrix w.r.t. any orthonormal basis is an orthogonal matrix.) An equivalent form of (2.1) is

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = {}^tR \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

The positive orientation of $\{\bar{i},\bar{j},\bar{k}\}$ means $\langle \bar{i}', \bar{j}', \bar{k}'\rangle = 1$. But $\langle \bar{i}', \bar{j}', \bar{k}'\rangle = det\, R$, thus $\mathcal{R}$ is a rotation.

If we do not impose the positive orientation of the new frame, then the determinant of the matrix of change can be -1. This is the case of a symmetry, or rotation followed by a symmetry.

Particular cases.

1) Rotation about the Oz axis. Denote by θ the rotation angle (Fig. 38).

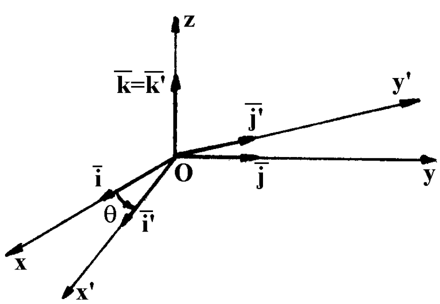

Fig. 38

From the figure above it follows that

$$\mathcal{R}(\bar{i}) = \bar{i}' = \bar{i}\cos\theta + \bar{j}\sin\theta$$

$$\mathcal{R}(\bar{j}) = \bar{j}' = -\bar{i}\sin\theta + \bar{j}\cos\theta$$

$$\mathcal{R}(\bar{k}) = \bar{k}' = \bar{k}.$$

Therefore $\mathcal{R}$ is described by

$$\mathcal{R} : \begin{cases} x = x'\cos\theta - y'\sin\theta \\ y = x'\sin\theta + y'\cos\theta \\ z = z'. \end{cases}$$

Obviously, the determinant of the associated matrix is $+1$, thus $\mathcal{R}$ is a positive isometry. In particular, a rotation in the xOy plane, of angle θ is described by

$$\mathcal{R} : \begin{cases} x = x'\cos\theta - y'\sin\theta \\ y = x'\sin\theta + y'\cos\theta. \end{cases}$$

The composite of a translation and a rotation is called a *roto-translation* (Fig. 39). Any roto-translation in the xOy plane is characterized by

$$\mathcal{R} : \begin{cases} x = x'\cos\theta - y'\sin\theta + a \\ y = x'\sin\theta + y'\cos\theta + b. \end{cases}$$

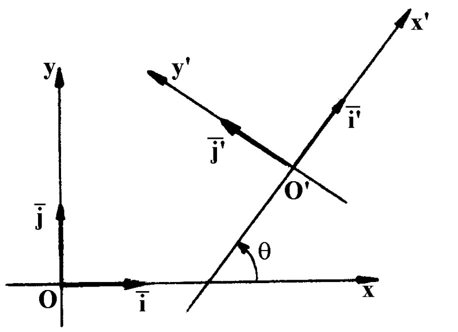

Fig. 39

2) Symmetry with respect to a plane. Consider the Cartesian frame $\{O;\bar{i},\bar{j},\bar{k}\}$ and $\mathcal{S}$ the symmetry w.r.t. the plane $(O;\bar{i},\bar{j})$. Then

$$\mathcal{S}(\bar{i}) = \bar{i}' = \bar{i};\ \mathcal{S}(\bar{j}) = \bar{j}' = \bar{j};\ \mathcal{S}(\bar{k}) = \bar{k}' = -\bar{k}.$$

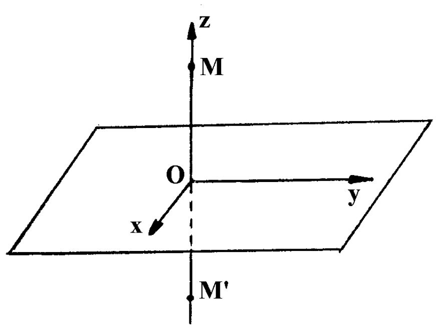

Fig. 40

By $x\bar{i} + y\bar{j} + z\bar{k} = x'\bar{i} + y'\bar{j} + z'\,\bar{k}$, it follows

$$\mathcal{S} : x = x',\ y = y',\ z = -z'$$

or in matrix form

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}.$$

The determinant of $\mathcal{S}$ is -1, thus $\mathcal{S}$ is a negative isometry.

3 Cylindrical Coordinates

Let $Oxyz$ be a fixed Cartesian system of coordinates in $\mathbf{E}_3$. Any point M is uniquelly determined by its Cartesian coordinates (x, y, z). Denote $\mathbf{E}_3^* = \mathbf{E}_3 \setminus Oz$. The point M is also characterized by the ordered triplet (ρ, θ, z), where ρ is the distance from the origin to the projection M' of M onto the xOy plane, $\rho = \sqrt{x^2 + y^2}$, and θ is the angle of the semilines Ox and OM (Fig. 41).

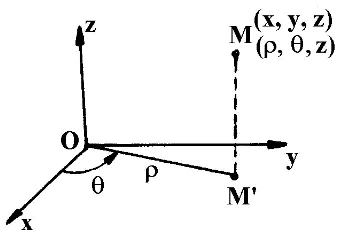

Fig. 41

The numbers ρ, θ, z are called the *cylindrical coordinates* of the point M. The cylindrical coordinates and the Cartesian coordinates of M are related by

$$\begin{cases} x = \rho \cos\theta \\ y = \rho \sin\theta \\ z = z. \end{cases}$$

If we impose $\rho > 0$, $\theta \in [0, 2\pi)$, then the above relations give a one-to-one correspondence between $\mathbf{E}_3 \setminus Oz$ and $(0, \infty) \times [0, 2\pi) \times \mathbf{R}$.

Coordinate sufaces

$\rho = \rho_0$: circular cylinder with generator lines parallel to Oz.
$\theta = \theta_0$: semiplane bounded by Oz.
$z = z_0$: plane parallel to xOy, without the point $(0, 0, z_0)$.

Coordinate curves

$\theta = \theta_0$, $z = z_0$: semiline parallel to xOy, with the origin on Oz.
$\rho = \rho_0$, $z = z_0$: circle whose center is on Oz, contained in a plane parallel to xOy.
$\theta = \theta_0$, $\rho = \rho_0$: line parallel to Oz.

The coordinate curves of different types are orthogonal, so the coordinate surfaces of different types are orthogonal too.

Consider the point $M(\rho, \theta, z)$. The unit vectors $\bar{e}_\rho, \bar{e}_\theta, \bar{e}_z$ tangent to the coordinate curves passing through M are pairwise orthogonal. The moving orthonormal frame $\{M(\rho, \theta, z);\ \bar{e_\rho}, \bar{e}_\theta, \bar{e}_z\}$ is called *cylindrical frame* (Fig. 42)

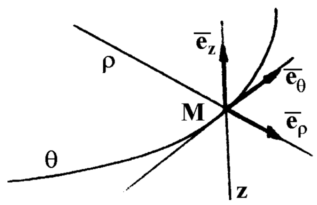

Fig. 42

The change from the Cartesian frame $\{O;\ \bar{i}, \bar{j}, \bar{k}\}$ to the cylindrical frame $\{M(\rho, \theta, z);\ \bar{e_\rho}, \bar{e}_\theta, \bar{e}_z\}$ is described by

$$\begin{cases} \bar{e}_\rho = \cos\theta\,\bar{i} + \sin\theta\,\bar{j} \\ \bar{e}_\theta = -\sin\theta\,\bar{i} + \cos\theta\,\bar{j} \\ \bar{e}_z = \bar{k}. \end{cases}$$

These formulas are based on the rule which gives the components of a vector w.r.t. an orthonormal basis as projections of that vector onto the basis vectors. For example

$$\bar{e}_\rho = \langle \bar{e}_\rho, \bar{i} \rangle \bar{i} + \langle \bar{e}_\rho, \bar{j} \rangle \bar{j} + \langle \bar{e}_\rho, \bar{k} \rangle \bar{k} = \cos\,\theta\,\bar{i} + \sin\theta\,\bar{j}.$$

4 Spherical Coordinates

Sometimes it is convenient to characterize the position of a point $M \in \mathbf{E}_3^* = \mathbf{E}_3 \setminus Oz$ using the ordered triplet (r, φ, θ), where r represents the distance $d(O, M)$, θ is the angle of the semilines Ox and OM', and φ is the angle between Oz and OM. M' denotes the projection of M onto xOy (Fig. 43).

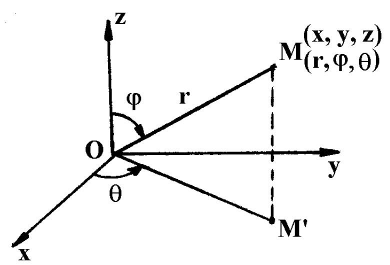

Fig. 43

The numbers r, φ, θ are called the *spherical coordinates* of the point M. The Cartesian and the spherical coordinates of M are related by

$$\begin{cases} x = r\, sin\, \varphi\, cos\, \theta \\ y = r\, sin\, \varphi\, sin\, \theta \\ z = z\, cos\, \varphi. \end{cases}$$

These formulas provide a one-to-one correspondence between the sets $\mathbf{E}_3 \setminus Oz$ and $(0, \infty) \times (0, \pi) \times [0, 2\pi)$.

Coordinate sufaces

$r = r_0$: sphere of radius r, center at O, without the north and the south poles.

$\theta = \theta_0$: semiplane bounded by Oz.

$\varphi = \varphi_0$: semicone without vertex (the origin).

Coordinate curves

$\theta = \theta_0$, $\varphi = \varphi_0$: semiline with origin at O.

$r = r_0$, $\varphi = \varphi_0$: circle whose center is on Oz, contained in a plane parallel to xOy.

$\theta = \theta_0$, $r = r_0$: open semicircle.

The coordinate curves of different types are orthogonal, so the coordinate surfaces of different types are orthogonal too.

Consider the point $M(r, \theta, \varphi)$. The unit vectors $\bar{e}_r, \bar{e}_\varphi, \bar{e}_\theta$ tangent to the coordinate curves passing through M are pairwise orthogonal. The moving orthonormal frame $\{M(r, \varphi, \theta);\ \bar{e_r}, \bar{e}_\varphi, \bar{e}_\theta\}$ is called *spherical frame* (Fig.44).

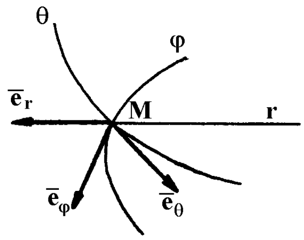

Fig. 44

The change from the Cartesian frame $\{O; \bar{i}, \bar{j}, \bar{k}\}$ to the spherical frame $\{M(r,\varphi,\theta); \bar{e_r}, \bar{e}_\varphi, \bar{e}_\theta\}$ is described by

$$\begin{cases} \bar{e}_r = sin\,\varphi\, cos\,\theta\,\bar{i} + sin\,\varphi\, sin\,\theta\,\bar{j} + cos\,\varphi\,\bar{k} \\ \bar{e}_\varphi = cos\,\varphi\, cos\,\theta\,\bar{i} + cos\,\varphi\, sin\,\theta\,\bar{j} - sin\,\varphi\,\bar{k} \\ \bar{e}_\theta = -sin\,\theta\,\bar{i} + cos\,\theta\,\bar{j}. \end{cases}$$

5 Problems

1. If the spherical coordinates of a point are $r = 5$, $\theta = 60^0$, $\varphi = 45^0$, determine its cylindrical coordinates and its Cartesian coordinates.

2. Express the cylindrical coordinates of an arbitrary point in terms of its spherical coordinates.

3. Express the spherical coordinates of an arbitrary point in terms of its cylindrical coordinates.

4. Determine the change of coordinates formulas for the following transformation: $O = O'$; the Ox' axis is contained in the xOy plane, and the angle (as oriented straight lines) between Ox' and Ox is acute; the Oz' axis has the same direction and the same sense as the vector $\bar{v} = 2\bar{i} + \bar{j} + 2\bar{k}$; the Oy' axis is such that the new Cartesian system $Ox'y'z'$ has positive orientation.

5. Consider the Cartesian frame $Oxyz$, and the the points $A(3,0,0)$, $B(0,2,0)$, $C(0,0,6)$. After a rotation we obtain the Cartesian frame $Ox'y'z'$ such that: Oz' has the direction and sense of the height OO' of the tetrahedron $OABC$; Oy' is parallel to $O'A'$, where A' is the orthogonal projection of A onto (ABC), and the Ox' axis is such that the the system $O'x'y'z'$ has positive orientation.

Write the matrix of the rotation and determine its invariant direction (its real 1-dimensional eigenspace).

6. Let D be a straight line of equations

$$D : \frac{x-2}{2} = \frac{y+1}{1} = \frac{z-3}{-2}.$$

Let $\bar{i}'$ be the unit director vector of D (choose a sense on D). Let $\bar{j}'$ be a versor contained in yOz and perpendicular on D, and the versor $\bar{k}'$ such that $\{\bar{i}', \bar{j}', \bar{k}'\}$ is an orthonormal basis.

Find out the change of basis formulas and compare the orientations of the two frames.

7. Consider three points given by their cylindrical coordinates:

$$A(5, \frac{\pi}{3}, 4),\ B(7, \frac{4\pi}{3}, -2),\ C(2, \frac{5\pi}{6}, -1).$$

Show that A and B belong to a plane passing through Oz; determine the Cartesian coordinates of A and C, and the distance $d(A, C)$.

8. Write the following equations in spherical coordinates.

$$(x^2 + y^2 + z^2)^2 = 3(x^2 + y^2); \quad (x^2 + y^2 + z^2)^2(x^2 + y^2) = 4x^2y^2.$$

Exam Samples

I.

1. Eigenvalues and eigenvectors; characteristic polynomial.

2. Determine the equation of the plane equidistant from

$$D_1 : \frac{x-1}{1} = \frac{y+1}{-1} = \frac{z}{1}, \quad D_2 : \frac{x}{2} = \frac{y}{-1} = \frac{z-1}{1}.$$

3. Find the canonical form of $x_1x_2 + x_2x_3 + x_3x_1$.

II.

1. Bases and dimension.

2. Prove the identity

$$\langle \bar{a} \times \bar{b}, \bar{c} \times \bar{d} \rangle = \begin{vmatrix} \langle \bar{a}, \bar{c} \rangle & \langle \bar{a}, \bar{d} \rangle \\ \langle \bar{b}, \bar{c} \rangle & \langle \bar{b}, \bar{d} \rangle \end{vmatrix}.$$

3. Determine the canonical form of the matrix

$$\begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}.$$

III.

1. Scalar product and vector product of free vectors.

2. Choose the basis $\{1, x, x^2, x^3\}$ in the vector space $\mathbf{V}$of all real polynomials of degree ≤ 3. Let D denote the differentation operator and $\mathcal{T} : \mathbf{V} \to \mathbf{V}$, $\mathcal{T}(p)(x) = xp'(x)$. Determine the matrix of each of the following linear transformations: $D, \mathcal{T}, D\mathcal{T}, \mathcal{T}D$. Find the eigenvalues and eigenvectors of $\mathcal{T}$.

3. Let $\mathbf{V}$ be the real vector space of all polynomials of order ≤ 2 and

$$\mathcal{A}(x, y) = \int_0^1 \int_0^1 x(t)y(s)dtds.$$

Show that $\mathcal{A}$ is a bilinear form which is symmetric and positive semidefinite. Find the matrix of $\mathcal{A}$ with respect to the basis $\{1, t, t^2\}$.

IV.

1. Euclidean vector spaces.

2. Show that if the vector $\bar{b}$ is perpendicular to $\bar{c}$, then $(\bar{a} \times \bar{b}) \times \bar{c} = \bar{b}\langle \bar{a}, \bar{c}\rangle$. Give sufficient conditions for the orthogonality of $(\bar{a} \times \bar{b}) \times \bar{c}$ and $\bar{c}$.

3. Let $\mathbf{V}$ be the vector space of all real polynomials $p(x)$ of degree $\leq n$. Define $\mathcal{T} : \mathbf{V} \to \mathbf{V}$, $\mathcal{T}(p) = q$, $q(t) = p(t+1)$. Determine the eigenvalues and the eigenvectors of $\mathcal{T}$.

V.

1. Linear transformations: general properties, kernel and image.

2. Show that the area of a figure F contained in the plane $P : z = px + qy + l$ and the area of its projection $\bar{F}$ onto the xOy plane are related by

$$S(F) = \sqrt{1 + p^2 + q^2}\, S(\bar{F}).$$

3. Show that $\langle f, g\rangle = \int_a^b f(t)g(t)dt$ is a scalar product on the real vector of all continuous functions on $[a, b]$.

VI.

1. Scalar product and mixed product of free vectors.

2. Let $\mathbf{V}$ and $\mathbf{W}$ be vector spaces, each of dimension 2, and $\{e_1, e_2\}$ a basis in $\mathbf{V}$and $\{f_1, f_2\}$ a basis in $\mathbf{W}$. Let $\mathcal{T} : \mathbf{V} \to \mathbf{W}$, $\mathcal{T}(e_1 + e_2) = 3f_1 + 9f_2$, $\mathcal{T}(3e_1 + 2e_2) = 7f_1 + 23f_2$, be a linear transformation. Compute $\mathcal{T}(e_2 - e_1)$ and determine the nullity and rank of $\mathcal{T}$. Determine $\mathcal{T}^{-1}$.

3. Determine an orthogonal matrix C which reduces the quadratic form $Q(x) = 2x_1^2 + 4x_1x_2 + 5x_2^2$ to a diagonal form.

VII.

1. The matrix of a linear transformation. Particular endomorphisms.

2. Find the equation of the plane P passing through the point (1,1,1), which is perpendicular to the planes $Q_1 : x + y + z + 1 = 0$, $Q_2 : x - y + z = 0$. Using the normal vectors of P, Q_1, Q_2 and the Gram-Schmidt procedure, find an orthonormal basis of $\mathbf{R}^3$.

3. Determine the canonical form of

$$Q(x) = 3x_1^2 - 5x_2^2 - 7x_3^2 - 8x_1x_2 + 8x_2x_3.$$

VIII.

1. Free vectors: collinearity, coplanarity.

2. Let $\mathbf{V}$ be the vector space of all continuous functions on $(-\infty, \infty)$ and such that the integral $\int\limits_{-\infty}^{x} f(t)dt$ exists for all x. Define $\mathcal{T} : \mathbf{V} \to \mathbf{V}$, $g = \mathcal{T}(f)$, $g(x) = \int\limits_{-\infty}^{x} f(t)dt$. Prove that every positive real number λ is an eigenvalue for $\mathcal{T}$ and determine the eigenfunctions corresponding to λ.

3. Let $Q : \mathbf{R}^3 \to \mathbf{R}$, $Q(x) = x_1x_2 + x_2x_3 + x_3x_1$. Determine the canonical form of Q and the image of

$$D : \frac{x_1 - 1}{1} = \frac{x_2}{-1} = \frac{x_3}{2}$$

by Q.

IX.

1. Vectors spaces. Vector subspaces.

2. Let $\mathcal{T} : \mathbf{C}^3 \to \mathbf{C}^3$ be represented by the matrix

$$\begin{bmatrix} a & \frac{i}{\sqrt{4}} & \frac{-1+2i}{\sqrt{12}} \\ b & \frac{1+i}{\sqrt{4}} & \frac{1-i}{\sqrt{12}} \\ c & \frac{-1}{\sqrt{4}} & \frac{2-i}{\sqrt{12}} \end{bmatrix}.$$

Determine a, b and c so that $\mathcal{T}$ is unitary (for $\mathbf{C}^3$ with the usual inner product).

3. Find the projection of

$$D : \frac{x-1}{2} = \frac{y}{1} = \frac{z-2}{-2}$$

onto

$$P : x + y + z = 0$$

and the symmetric of P with respect to D.

X.

1. The spectrum of endomorphisms on Euclidean spaces.

2. Determine the relation between $a, b, c, d, \alpha, \beta, \gamma, \delta, \lambda, \mu$ such that the three planes $ax + by + cz + d = 0$, $\alpha x + \beta y + \gamma z + \delta = 0$, $\lambda(ax + by + cz) + \mu(\alpha x + \beta y + \gamma z) = 0$ have no points in common.

3. Let $\{x, y\}$ be a linearly independent set in a Euclidean space $\mathbf{V}$.
Define $f : \mathbf{R} \to \mathbf{R}$ by $f(\alpha) = \|x - \alpha y\|$.

(a) Where does f take on its minimum value ?
(b) Give a geometric interpretation of (a). What happens if $f : \mathbf{C} \to \mathbf{R}$?

XI.

1. Diagonal form of an endomorphism.

2. Let $P : x + y - 2z = 0$. Find the equation of the plane Q symmetric to P about xOy (about the origin O).

3. Let $\mathcal{A} : \mathbf{R}^4 \times \mathbf{R}^4 \to \mathbf{R}$,

$$\mathcal{A}(x, y) = x_1y_2 - x_2y_1 + x_1y_3 - x_3y_1 + x_1y_4 - x_4y_1 + x_4y_4.$$

Find the matrix of $\mathcal{A}$ with respect to the basis $f_1 = (1, 1, 0, 0)$, $f_2 = (0, 1, 1, 0)$, $f_3 = (0, 1, 0, 1)$, $f_4 = (1, 0, 0, 1)$.

XII.

1. Basic facts on straight lines and planes.

2. Let P_4 denote the space of all polynomials in t of degree at most 3. Find the matrix representation of $\mathcal{T} : P_4 \to P_4$, $y = \mathcal{T}x$, $y(t) = \dfrac{d}{dt}\left((t^2 - 1)\dfrac{dx}{dt}\right)$ with respect to the basis $B = \left\{1, t, \dfrac{3}{2}t^2 - \dfrac{1}{2}, \dfrac{5}{2}t^3 - \dfrac{3}{2}t\right\}$.

3. Let $\mathbf{V}$ consist of all infinite sequences $x = \{x_n\}$ of real numbers for which the series $\sum x_n^2$ converges. Define $\langle x, y\rangle = \sum x_n y_n$. Prove that this series converges absolutely and $\langle x, y\rangle$ is a scalar product. Compute $\langle x, y\rangle$ if $x_n = \dfrac{1}{n}$ and $y_n = \dfrac{1}{n+1}$.

XIII.

1. Quadratic forms.

2. Let $\bar{a}$, $\bar{b}$, $\bar{c} \in \mathbf{V}_3$ such that $\bar{b}$ perpendicular to $\bar{c}$. Show that $(\bar{a} \times \bar{b}) \times \bar{c} = \bar{b}\langle \bar{a}, \bar{c}\rangle$.

3. Let $\mathbf{V} = C[0, T]$ and define $(\mathcal{P}x)(t) = x(0)(1 - t)$, for $0 \le t \le T$. Show that $\mathcal{P}$ is a projection and determine the range of $\mathcal{P}$.

XIV.

1. Polar, cylindrical and spherical coordinates.

2. Consider the linear transformation $\mathcal{T} : \mathbf{V} \to \mathbf{V}$ given by $y = \mathcal{T}x$, where $y(t) = \displaystyle\int_0^{2\pi} x(0) 4 \cos 2(t - s) ds$, and $\mathbf{V} = L\{1, \cos s, \cos 2s, \sin s, \sin 2s\}$.
(a) Express $\mathcal{T}$ as a matrix.
(b) Is $\mathcal{T}$ one-to-one ?

(c) Does it map $\mathbf{V}$ onto itself ?

3. Let $\mathbf{V}$ be the vector space of all real-valued functions defined on the real line. Which of the subsets $\{1, e^{ax}, xe^{ax}\}$, $\{1, \cos 2x, \sin^2 x\}$ is linearly independent in $\mathbf{V}$? Compute the dimension of the subspace spanned by each subset.

XV.

1. Equations of a straight line.

2. Let $\mathbf{V} = L_2[-\pi, \pi]$, and $V_1 = L(A_1)$, $V_2 = L(A_2)$, where $A_1 = \{1, \cos t, \cos 2t, \ldots\}$, $A_2 = \{\sin t, \sin 2t, \ldots\}$. Show that A_1, A_2 are linearly independent, and the sum of V_1 and V_2 is direct.

3. Let $\mathcal{T} : \mathbf{V} \to \mathbf{V}$, $g = \mathcal{T}(f)$, $g(x) = \displaystyle\int_{-\pi}^{\pi} (1 + \cos(x - t))f(t)dt$. Find a basis for $\mathcal{T}(\mathbf{V})$.

(a) Determine the kernel of $\mathcal{T}$.
(b) Find the eigenvalues of $\mathcal{T}$.

XVI.

1. Equation of a plane in space.

2. Let $\mathbf{V}$ be the real vector space of all functions of the form $x(t) = a\cos(\omega t + \phi)$, where ω is fixed. Show that $B = \{\cos \omega t, \sin \omega t\}$ is a basis of $\mathbf{V}$. Give examples of linear transformations on $\mathbf{V}$.

3. Let $\mathbf{V}$ be the real Euclidean space of real polynomial functions on [-1,1]. Determine which of the following linear transformations is symmetric or skew symmetric: $\mathcal{T}(f)(x) = f(-x)$, $\mathcal{T}(f)(x) = f(x) + f(-x)$, $\mathcal{T}(f)(x) = f(x) - f(-x)$.

XVII.

1. Transformations of Cartesian frames.

2. Let $\mathbf{V}$ denote the real vector space of all polynomials in t, of degree at most four, and define $\mathcal{T} : \mathbf{V} \to \mathbf{V}$ by $\mathcal{T} = D^2 + 2D + I$, where $Dx = \dfrac{dx}{dt}$ (differential operator).

(a) Represent $\mathcal{T}$ by a matrix T w.r.t. the basis $\{1, t, t^2, t^3, t^4\}$.
(b) Represent $\mathcal{T}^2$ by a matrix.

3. Let $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in \mathbf{R}^n$. Find out whether $\langle x, y\rangle = \displaystyle\sum_{i=1}^{n} x_i \sum_{y=1}^{n} y_j$, $\langle x, y\rangle = \left|\displaystyle\sum_{i=1}^{n} x_i y_i\right|$ are scalar products or not. When $\langle x, y\rangle$ is not a scalar product, point out the axioms which are not fulfilled.

XVIII.

1. Jordan form of an endomorphism.

2. What is the locus of points with the property that the ratio of distances to two given planes is constant?

3. In the linear space of all real polynomials, with scalar product $\langle x, y\rangle = \int_1^0 x(t)y(t)dt$, let $x_n(t) = t^n$ for $n = 0, 1, 2, \ldots$. Prove that the functions $y_0(t) = 1$, $y_1(t) = \sqrt{3}(2t-1)$, $y_2(t) = \sqrt{5}(6t^2 - 6t + 1)$ form an orthonormal set spanning the same subspace as $\{x_0, x_1, x_2\}$.

XIX.

1. Free vectors: addition, multiplication of a vector by a scalar.

2. Let $\mathbf{V}$ be the vector space of all continuous functions defined on $(-\infty, \infty)$, and such that the integral $\int_{-\infty}^{x} tf(t)dt$ exists for all real numbers x. Define $\mathcal{T} : \mathbf{V} \to \mathbf{V}$, $g = \mathcal{T}(f)$, $g(x) = \int_{-\infty}^{x} tf(t)dt$. Prove that every negative λ is a proper value for $\mathcal{T}$ and determine the eigenvectors corresponding to λ.

3. Determine the signature of the quadratic form

$$Q(x) = 3x_1^2 - 5x_2^2 + x_3^2 - 2x_1x_2 + x_2x_3.$$

XX.

1. Linear transformations on Euclidean spaces.

2. Show that the locus of points equidistant from three pairwise non-parallel planes is a straight line.

3. Show that the quadratic forms

$$Q(x) = 3x_1^2 + 4x_2^2 + 5x_3^2 + 4x_1x_2 - 4x_2x_3,$$

$$Q(x) = 2x_1^2 + 5x_2^2 + 5x_3^2 + 4x_1x_2 - 4x_1x_3 - 8x_2x_3$$

are positive definite.

XXI.

1. Isometries.

2. Write the equations of the straight line D passing through the point (1,1,1) and parallel to the planes $P : x - y + z = 0$, $Q : x + 2y - z = 0$. Find the points of P which are at the distance 2 from D.

3. Define $\mathcal{T} : \mathbf{C}^2 \to \mathbf{C}^2$ by

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

where a, b, c, d are complex numbers. Determine necessary and sufficient conditions on a, b, c, d such that $\mathcal{T}$ is self-adjoint, or unitary.

XXII.

1. Polynomials of matrices. Functions of matrices.

2. Find out the angles between the coordinate axes and the plane $P : x+y+2z = 0$. Determine the symmetric of P with respect to the line $D : \frac{x}{1} = \frac{y-1}{-1} = \frac{z}{1}$.

3. Let x, y be vectors in a Euclidean space $\mathbf{V}$and assume that

$$\|\lambda x + (1-\lambda)y\| = \|x\|, \quad \forall \lambda \in [0,1].$$

Show that $x = y$.

XXIII.

1. Orthogonality. The Gram-Schmidt orthogonalization procedure.

2. Prove that: if the vectors $\bar{a}$ and $\bar{b}$ are perpendicular to the vector $\bar{c}$, then $(\bar{a} \times \bar{b}) \times \bar{c} = 0$. Give sufficient conditions for the collinearity of $(\bar{a} \times \bar{b}) \times \bar{c}$ and $\bar{c}$.

3. Find the canonical form of $\mathcal{T} : \mathbf{R}^3 \to \mathbf{R}^3$,

$$T = \begin{bmatrix} 7 & 4 & -1 \\ 4 & 7 & -1 \\ -4 & -4 & 4 \end{bmatrix}.$$

XXIV.

1. Linear dependence and independence.

2. Find the angles between the coordinate planes and the straight line

$$D : \frac{x-1}{1} = \frac{y}{-2} = \frac{z+2}{-1}.$$

Determine the equations of the symmetric of D with respect to the plane

$$P : x + y + z = 0.$$

3. Determine the canonical form of $\mathcal{T} : \mathbf{R}^3 \to \mathbf{R}^3$ given by

$$T = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ -1 & -3 & 3 \end{bmatrix}.$$

XXV.

1. Bilinear forms.

2. Prove that the straight lines joining the mid-points of the opposite edges of a tetrahedron intersect at one point. Express the coordinates of this point in terms of the coordinates of the vertices of the tetrahedron.

3. Show that the set of all functions $x_n(t) = e^{int}$, $n = 0, 1, 2, \ldots$ is linearly independent in $L_2[0, 2\pi]$.

XXVI.

1. Eigenvalues and eigenvectors of an endomorphism.

2. Given the point $A(0, 1, 2)$ and the line

$$D : \begin{cases} x + y = 0 \\ x + z + 1 = 0, \end{cases}$$

detemine the symmetric of A with respect to D and the symmetric of D w.r.t. A.

3. Let $Q : \mathbf{R}^3 \to \mathbf{R}$,

$$Q(x) = 4x_1^2 + x_2^2 + 9x_3^2 + 4x_1x_2 - 12x_1x_3 - 6x_2x_3.$$

Reduce Q to the canonical expression using the eigenvalues method; determine the corresponding basis.

Use another method of reduction to the canonical form and check the inertia law.

XXVII.

1. The common perpendicular of two noncoplanar straight lines. The distance between two noncoplanar straight lines.

2. Consider the real vector space $\mathbf{V} = C^\infty(\mathbf{R})$ and $D : \mathbf{V} \to \mathbf{V}$, $D(f) = f'$. Find the eigenvalues and the eigenvectors of D. Is $\lambda = 2$ an eigenvalue? What is the dimension of $S(2)$?

3. Let $\mathcal{T} : \mathbf{R}_2[X] \to \mathbf{R}_2[X]$ be a linear map such that

$$\mathcal{T}(1 + X) = X^2, \ \mathcal{T}(X + X^2) = 1 - X, \ \mathcal{T}(1 + X + X^2) = X + X^2.$$

Find the matrix associated to $\mathcal{T}$ w.r.t. canonical basis of $\mathbf{R}_2[X]$. Determine the dimension and a basis for each of $Ker(\mathcal{T})$ and $Im(\mathcal{T})$.

XXVIII.

1. Endomorphisms of Euclidean vector spaces.

2. Consider $v = (14, -3, -6) \in \mathbf{R}^3$ and $S = \{v_1, v_2, v_3\}$, where $v_1 = (-3, 0, 7)$, $v_2 = (1, 4, 3)$, $v_3 = (2, 2, -2)$. Determine the orthogonal projection w of v on $Span\, S$ and the vector $w^\perp$.

3. Let $\mathcal{T} : \mathbf{R}^3 \to \mathbf{R}^3$ be an endomorphism, and

$$T = \begin{bmatrix} 4 & 6 & 0 \\ -3 & -5 & 0 \\ -3 & -6 & 1 \end{bmatrix}$$

the matrix of $\mathcal{T}$ w.r.t. the canonical basis. Find the diagonal form of $\mathcal{T}$ and the corresponding basis. Is $\mathcal{T}$ injective? Is $\mathcal{T}$ surjective?

XXIX.

1. The vector space of free vectors: mixed product - properties, geometric interpretation.

2. Let $\mathcal{T} : \mathbf{R}^3 \to \mathbf{R}^3$ be an endomorphism such that $\mathcal{T}(1,0,0) = (3,1,0)$, $\mathcal{T}(0,0,1) = (1,2,0)$, $Ker\,\mathcal{T} = \{(\alpha, 2\alpha, 3\alpha) \mid \alpha \in \mathbf{R}\}$. Determine a basis for $Ker\,\mathcal{T}$ and $Im\,\mathcal{T}$ respectively; determine also the matrix of $\mathcal{T}$ w.r.t. the canonical basis.

3. Given

$$A = \begin{bmatrix} 1 & -1 & -2 \\ -1 & 3 & -1 \\ -2 & -1 & 1 \end{bmatrix},$$

find the canonical expression of the quadratic form $Q : \mathbf{R}^3 \to \mathbf{R}$, $Q(x) = {}^tXAX$ and the correspoding change of coordinates.

XXX.

1. Matrix polynomials.

2. Consider the vectors $v_1 = (3,2,-1)$, $v_2 = (1,-2,1)$, $v_3 = (1,0,2)$. Prove that there exists a unique linear form $f : \mathbf{R}^3 \to \mathbf{R}$ such that $f(v_1) = 5$, $f(v_2) = -3$, $f(v_3) = 6$. Write $f(x)$ in terms of the components of x, for x arbitrary in $\mathbf{R}^3$, and determine an orthonormal basis of $Ker\,f$. Is $Im\,f$ a proper subspace of $\mathbf{R}$?

3. Let $M(1,-2,5)$ and $Q : x + 2y + 2z + 1 = 0$. Determine the projection of the point M onto the plane Q, the symmetric of M w.r.t. Q, the distance from M to Q, and the symmetric of Q w.r.t. M.

XXXI.

1. Bilinear forms.

2. Use the Cayley-Hamilton theorem to determine the inverse of the matrix

$$A = \begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}$$

and the value of the matrix polynomial $R(A) = A^3 + 3A^2 - 8A - 27I$.

3. Let $V = C^0[0, 2\pi]$, $S = \{f_0, f_1, \ldots, f_k \ldots\} \subset V$, where $f_0(x) = 1$, $f_{2n-1}(x) = \cos nx$, $f_{2n}(x) = \sin nx$, $n \in \mathbf{N}^*$, $x \in [0, \pi]$.

Prove that S is linearly independent and determine an orthonormal basis of $Span\, S$.

XXXII.

1. Quadratic forms.

2. If $\bar{a}, \bar{b}, \bar{c} \in \mathbf{V}_3$ are noncoplanar vectors such that $\langle \bar{a}, \bar{b} \rangle \neq 0$, show that

$$E = \left\langle \frac{\bar{a} \times \bar{b}}{\langle \bar{a}, \bar{b} \rangle}, \frac{\bar{a} \times (\bar{b} \times \bar{c})}{\langle \bar{a}, \bar{b} \times \bar{c} \rangle} \right\rangle$$

does not depend on $\bar{a}, \bar{b}, \bar{c}$.

3. Determine the eigenvalues and the eigenvectors of the linear transformation $\mathcal{T} : \mathbf{R}^3 \to \mathbf{R}^3$, $\mathcal{T}(x) = (x_1 - 2x_2 - x_3, -x_1 + x_2 + x_3, x_1 - x_3)$.

(a) Is $\mathcal{T}$ diagonalizable? (b) Find orthonormal bases in $Ker\mathcal{T}$ and $Im\mathcal{T}$.

XXXIII.

1. Kernel and image of a linear transformation.

2. Given the vectors

$$\bar{a} = \bar{i} - \alpha\bar{j} + 3\bar{k},\ \bar{b} = \alpha\bar{i} - \bar{j} + \bar{k},\ \bar{c} = 3\bar{i} + \bar{j} - \bar{k},$$

find the value of $\alpha \in \mathbf{R}$ such that $\bar{a}, \bar{b}, \bar{c}$ are coplanar.

For $\alpha = 2$, determine the altitude of the parallelipiped constructed on some representatives with common origin of the vectors $\bar{a}, \bar{b}, \bar{c}$, corresponding to the base whose sides are the representatives of $\bar{a}$ and $\bar{b}$.

3. Let

$$T = \begin{bmatrix} 4 & 2 & -6 \\ 2 & 1 & -3 \\ -6 & -3 & 9 \end{bmatrix}$$

be the matrix associated to the endomorphism $\mathcal{T} : \mathbf{R}^3 \to \mathbf{R}^3$ w.r.t. the canonical basis of $\mathbf{R}^3$.

(a) Find out whether $\mathcal{T}$ admits a Jordan form or not.

(b) Compute $\mathcal{T}^n$, $n \in \mathbf{N}^*$.

(c) Is $\mathcal{T}$ an injective endomorphism?

XXXIV.

1. Collinearity and coplanarity in $\mathbf{V}_3$.

2. (a) Write the analytic expression of the bilinear form $\mathcal{A} : \mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}$, if the associated matrix w.r.t. the canonical basis is

$$A = \begin{bmatrix} 1 & 2 & -1 \\ -1 & 2 & 3 \\ 3 & 0 & 1 \end{bmatrix}.$$

(b) Is $\mathcal{A}$ a symmetric bilinear form?
(c) Let $Q : \mathbf{R}^3 \to \mathbf{R}$ be a quadratic form,

$$Q(x) = x_1^2 + 8x_2^2 + x_3^2 + 16x_1x_2 + 4x_1x_3 + 4x_2x_3.$$

Determine the canonical form of Q and the corresponding basis, using Jacobi's method.

3. Prove that the straight lines

$$D_1 : \frac{x}{1} = \frac{y}{2} = \frac{z}{3},$$

$$D_2 : \frac{x-1}{2} = \frac{y-1}{-1} = \frac{z}{1}$$

are noncoplanar, determine the equations of their common perpendicular and the distance between D_1 and D_2.

XXXV.

1. Linear dependence. Linear independence.

2. Consider the straight line D and the plane P of equations

$$D : \frac{x}{-1} = \frac{y}{2} = \frac{z}{1}$$

$$P : x - 2y + z = 0.$$

(a) Find the projection of D onto P.
(b) If D and P are viewed as vector subspaces of $\mathbf{R}^3$, determine $D + P$, $D \cap P$ and an orthonormal basis of P.

3. Given the matrix

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix},$$

use the Cayley - Hamilton theorem in order to:
(a) write A^{-1} as a polynomial of degree 2 in A;
(b) show that

$$A^n = \begin{bmatrix} 1 & 0 & (3^n - 1)/2 \\ 0 & 2^n & 0 \\ 0 & 0 & 3^n \end{bmatrix}.$$

Bibliography

[1] M.Artin, *Algebra*, Prentice-Hall, 1991.

[2] Gh.Atanasiu, Gh.Munteanu, M.Postolache, *Algebră liniară, geometrie analitică şi diferenţială* (in Romanian), Editura All, Bucureşti, 1994.

[3] V. Balan, *Algebră liniară, geometrie analitică şi diferenţială* (in Romanian), Universitatea Politehnica Bucureşti, 1998.

[4] R.M.Bowen, C.C.Wang, *Introduction to Vectors and Tensors*, vol. 1-2, Plenum Press, New York, 1976.

[5] V.Brînzănescu, O.Stănăşilă, *Matematici speciale* (in Romanian), Editura All, Bucureşti, 1994.

[6] J.Dieudonne, *Linear Algebra and Geometry*, Paris, Hermann, 1969.

[7] O.Dogaru, M.Doroftei, *Algebră liniară* (in Romanian), Geometry Balkan Press, Bucureşti, 1998.

[8] Gh.Th.Gheorghiu, *Algebră liniară, geometrie analitică şi diferenţială şi programare* (in Romanian), Editura Didactică şi Pedagogică, Bucureşti, 1977.

[9] S.Ianuş, *Curs de geometrie diferentială* (in Romanian), Universitatea Bucureşti, 1981.

[10] I.D.Ion, R.Nicolae, *Algebra* (in Romanian), Editura Didactică şi Pedagogică, Bucureşti, 1981.

[11] N.Jacobson, *Lectures in Abstract Algebra, II - Linear algebra*, Springer-Verlag, 1975.

[12] W.Klingenberg, *Lineare Algebra und Geometrie*, Springer-Verlag, Berlin, 1990.

[13] S. Lang, *Algebra*, Addison-Wesley, 1984.

[14] A.V.Pogorelov, *Analytic Geometry*, Mir Publishers, Moscow, 1961.

[15] C.Radu, *Algebră liniară, geometrie analitică şi diferenţială* (in Romanian), Editura All, Bucureşti, 1996.

[16] C.Radu, C.Drăguşin, L.Drăguşin, *Aplicaţii de algebră, geometrie şi matematici speciale* (in Romanian), Editura Didactică şi Pedagogică, Bucureşti, 1991.

[17] L.Smith, *Linear Algebra*, Springer-Verlag, 1978.

[18] N.Soare, *Curs de geometrie* (in Romanian), Universitatea Bucureşti, 1996.

[19] C.Udrişte, *Linear Algebra*, University Politehnica of Bucharest, 1991-1992.

[20] C.Udrişte, *Problems in Algebra, Geometry and Differential Equations I, II*, University Politehnica of Bucharest, 1992.

[21] C.Udrişte, *Aplicaţii de algebră, geometrie şi ecuaţii difereţiale* (in Romanian), Editura Didactică şi Pedagogică, Bucureşti, 1993.

[22] C.Udrişte, *Algebră liniară, Geometrie analitică* (in Romanian), Geometry Balkan Press, Bucureşti, 1996.

[23] C.Udrişte, O.Dogaru, *Algebră liniară, Geometrie analitică* (in Romanian), Universitatea Politehnica Bucureşti, 1991.

[24] C.Udrişte, C.Radu, C.Dicu, O.Mălăncioiu, *Probleme de algebră, geometrie şi ecuaţii diferenţiale* (in Romanian), Editura Didactică şi Pedagogică, Bucureşti, 1981.

[25] C.Udrişte, C.Radu, C.Dicu, O.Mălăncioiu, *Algebră, geometrie şi ecuaţii diferenţiale* (in Romanian), Editura Didactică şi Pedagogică, Bucureşti, 1982.