

# Chapter 9

## Endomorphisms and Diagonalization



Both in classical and quantum physics, and in several branches of mathematics, it is hard to overestimate the role that the notion of diagonal action of a linear map has. The aim of this chapter is to introduce this topic which will be crucial in all the following chapters.

### 9.1 Endomorphisms

**Definition 9.1.1** Let  $V$  be a real vector space. A linear map  $\phi : V \rightarrow V$  is called an *endomorphism* of  $V$ . The set of all endomorphisms of  $V$  is denoted  $\text{End}(V)$ . Non invertible endomorphisms are also called *singular* or *degenerate*.

As seen in Sect. 8.1, the set  $\text{End}(V)$  is a real vector space with  $\dim(\text{End}(V)) = n^2$  if  $\dim(V) = n$ .

The question we address now is whether there exists a class of bases of the vector space  $V$ , with respect to which a matrix  $M_\phi^{\mathcal{B},\mathcal{B}}$  has a particular (diagonal, say) form. We start with a definition.

**Definition 9.1.2** The matrices  $A, B \in \mathbb{R}^{n,n}$  are called *similar* if there exists a real vector space  $V$  and an endomorphism  $\phi \in \text{End}(V)$  such that  $A = M_\phi^{\mathcal{B},\mathcal{B}}$  and  $B = M_\phi^{\mathcal{C},\mathcal{C}}$ , where  $\mathcal{B}$  and  $\mathcal{C}$  are bases for  $V$ . We denote similar matrices by  $A \sim B$ .

Similarity between matrices can be described in a purely algebraic way.

**Proposition 9.1.3** *The matrices  $A, B \in \mathbb{R}^{n,n}$  are similar if and only if there exists an invertible matrix  $P \in \text{GL}(n)$ , such that  $P^{-1}AP = B$ .*

*Proof* Let us assume  $A \sim B$ : we then have a real vector space  $V$ , bases  $\mathcal{B}$  and  $\mathcal{C}$  for it and an endomorphism  $\phi \in \text{End}(V)$  such that  $A = M_{\phi}^{\mathcal{B},\mathcal{B}}$  e  $B = M_{\phi}^{\mathcal{C},\mathcal{C}}$ . From the Theorem 7.9.9 we have

$$B = M^{C,\mathcal{B}} A M^{\mathcal{B},C}.$$

Since the matrix  $M^{C,\mathcal{B}}$  is invertible, with  $(M^{C,\mathcal{B}})^{-1} = M^{\mathcal{B},C}$ , the claim follows with  $P = M^{\mathcal{B},C}$ .

Next, let us assume there exists a matrix  $P \in \text{GL}(n)$  such that  $P^{-1}AP = B$ . From the Theorem 7.9.6 and the Remark 7.9.7 we know that the invertible matrix  $P$  gives a change of basis in  $\mathbb{R}^n$ : there exists a basis  $\mathcal{C}$  for  $\mathbb{R}^n$  (the columns of  $P$ ), with  $P = M^{\mathcal{E},\mathcal{C}}$  and  $P^{-1} = M^{C,\mathcal{E}}$ . Let  $\phi = f_A^{\mathcal{E},\mathcal{E}}$  be the endomorphism in  $\mathbb{R}^n$  corresponding to the matrix  $A$  with respect to the canonical bases,  $A = M_{\phi}^{\mathcal{E},\mathcal{E}}$ . We then have

$$\begin{aligned} B &= P^{-1} A P \\ &= M^{C,\mathcal{E}} M_{\phi}^{\mathcal{E},\mathcal{E}} M^{\mathcal{E},C} \\ &= M_{\phi}^{C,C}. \end{aligned}$$

This shows that  $B$  corresponds to the endomorphism  $\phi$  with respect to the different basis  $\mathcal{C}$ , that is  $A$  and  $B$  are similar.  $\square$

*Remark 9.1.4* The similarity we have introduced is an equivalence relation in  $\mathbb{R}^{n,n}$ , since it is

- (a) reflexive, that is  $A \sim A$  since  $A = I_n A I_n$ ,
- (b) symmetric, that is  $A \sim B \Rightarrow B \sim A$  since

$$P^{-1} A P = B \Rightarrow P B P^{-1} = A,$$

- (c) transitive, that is  $A \sim B$  and  $B \sim C$  imply  $A \sim C$ , since  $P^{-1} A P = B$  and  $Q^{-1} B Q = C$  clearly imply  $Q^{-1} P^{-1} A P Q = (PQ)^{-1} A (PQ) = C$ .

If  $A \in \mathbb{R}^{n,n}$ , we denote its equivalence class by similarity as  $[A] = \{B \in \mathbb{R}^{n,n} : B \sim A\}$ .

**Proposition 9.1.5** *Let matrices  $A, B \in \mathbb{R}^{n,n}$  be similar. Then*

$$\det(B) = \det(A) \quad \text{and} \quad \text{tr}(B) = \text{tr}(A).$$

*Proof* From Proposition 9.1.3, we know there exists an invertible matrix  $P \in \text{GL}(n)$ , such that  $P^{-1}AP = B$ . From the Binet Theorem 5.1.16 and the Proposition 4.5.2 we can write

$$\begin{aligned} \det(B) &= \det(P^{-1}AP) \\ &= \det(P^{-1}) \det(A) \det(P) = \det(P^{-1}) \det(P) \det(A) \\ &= \det(A) \end{aligned}$$

and  $\text{tr}(B) = \text{tr}(P^{-1}AP) = \text{tr}(PP^{-1}A) = \text{tr}(A)$ .  $\square$

A natural question is whether, for a given  $A$ , the equivalence class  $[A]$  contains a diagonal element (equivalently, whether  $A$  is similar to a diagonal matrix).

**Definition 9.1.6** A matrix  $A \in \mathbb{R}^{n,n}$  is called *diagonalisable* if it is similar to a diagonal ( $\Delta$  say) matrix, that is if there is a diagonal matrix  $\Delta$  in the equivalence class  $[A]$ .

Such a definition has a counterpart in terms of endomorphisms.

**Definition 9.1.7** An endomorphism  $\phi \in \text{End}(V)$  is called *simple* if there exists a basis  $\mathcal{B}$  for  $V$  such that the matrix  $M_\phi^{\mathcal{B},\mathcal{B}}$  is diagonalisable.

We expect that for an endomorphism to be simple is an intrinsic property which does not depend on the basis with respect to which its corresponding matrix is given. The following proposition confirms this point.

**Proposition 9.1.8** Let  $V$  be a real vector space, with  $\phi \in \text{End}(V)$ . The following are equivalent:

- (i)  $\phi$  is simple, there is a basis  $\mathcal{B}$  for  $V$  such that  $M_\phi^{\mathcal{B},\mathcal{B}}$  is diagonalisable,
- (ii) there exists a basis  $\mathcal{C}$  for  $V$  such that  $M_\phi^{\mathcal{C},\mathcal{C}}$  is diagonal,
- (iii) given any basis  $\mathcal{D}$  for  $V$ , the matrix  $M_\phi^{\mathcal{D},\mathcal{D}}$  is diagonalisable.

*Proof* (i)  $\Rightarrow$  (ii): Since  $M_\phi^{\mathcal{B},\mathcal{B}}$  is similar to a diagonal matrix  $\Delta$ , from the proof of the Proposition 9.1.3 we know that there is a basis  $\mathcal{C}$  with respect to which  $\Delta = M_\phi^{\mathcal{C},\mathcal{C}}$  is diagonal.

(ii)  $\Rightarrow$  (iii): Let  $\mathcal{C}$  be a basis of  $V$  such that  $M_\phi^{\mathcal{C},\mathcal{C}} = \Delta$  is diagonal. For any basis  $\mathcal{D}$  we have then  $M_\phi^{\mathcal{D},\mathcal{D}} \sim \Delta$ , thus  $M_\phi^{\mathcal{D},\mathcal{D}}$  is diagonalisable.

(iii)  $\Rightarrow$  (i): obvious. □

## 9.2 Eigenvalues and Eigenvectors

*Remark 9.2.1* Let  $\phi : V \rightarrow V$  be a simple endomorphism, with  $\Delta = M_\phi^{\mathcal{C},\mathcal{C}}$  a diagonal matrix associated to  $\phi$ . It is then

$$\Delta = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

for scalars  $\lambda_j \in \mathbb{R}$ , with  $j = 1, \dots, n$ . By setting  $\mathcal{C} = (v_1, \dots, v_n)$ , we write then  $\phi(v_j) = \lambda_j v_j$ .

The vectors of the basis  $\mathcal{C}$  and the scalars  $\lambda_j$  plays a prominent role in the analysis of endomorphisms. This motivates the following definition.

**Definition 9.2.2** Let  $\phi \in \text{End}(V)$  with  $V$  a real vector space. If there exists a non zero vector  $v \in V$  and a scalar  $\lambda \in \mathbb{R}$ , such that

$$\phi(v) = \lambda v,$$

then  $\lambda$  is called an *eigenvalue* of  $\phi$  and  $v$  is called an *eigenvector* of  $\phi$  associated to  $\lambda$ . The *spectrum* of an endomorphism is the collection of its eigenvalues.

*Remark 9.2.3* Let  $\phi \in \text{End}(V)$  and  $\mathcal{C}$  be a basis of  $V$ . With the definition above, the content of the Remark 9.2.1 can be rephrased as follow:

- (a)  $M_{\phi}^{\mathcal{C},\mathcal{C}}$  is diagonal if and only if  $\mathcal{C}$  is a basis of eigenvectors for  $\phi$ ,
- (b)  $\phi$  is simple if and only if  $V$  has a basis of eigenvectors for  $\phi$  (from the Definition 9.1.7).

Notice that each eigenvector  $v$  for an endomorphism  $\phi$  is uniquely associated to an eigenvalue  $\lambda$  of  $\phi$ . On the other hand, more than one eigenvector can be associated to a given eigenvalue  $\lambda$ . It is indeed easy to see that, if  $v$  is associated to  $\lambda$ , also  $\alpha v$ , with  $\alpha \in \mathbb{R}$ , is associated to the same  $\lambda$  since  $\phi(\alpha v) = \alpha\phi(v) = \alpha(\lambda v) = \lambda(\alpha v)$ .

**Proposition 9.2.4** If  $V$  is a real vector space, and  $\phi \in \text{End}(V)$ , the set

$$V_{\lambda} = \{v \in V : \phi(v) = \lambda v\}$$

is a vector subspace in  $V$ .

*Proof* We explicitly check that  $V_{\lambda}$  is closed under linear combinations. With  $v_1, v_2 \in V_{\lambda}$  and  $a_1, a_2 \in \mathbb{R}$ , we can write

$$\phi(a_1 v_1 + a_2 v_2) = a_1 \phi(v_1) + a_2 \phi(v_2) = a_1 \lambda v_1 + a_2 \lambda v_2 = \lambda(a_1 v_1 + a_2 v_2),$$

showing that  $V_{\lambda}$  is a vector subspace of  $V$  □

**Definition 9.2.5** If  $\lambda \in \mathbb{R}$  is an eigenvalue of  $\phi \in \text{End}(V)$ , the space  $V_{\lambda}$  is called the *eigenspace* corresponding to  $\lambda$ .

*Remark 9.2.6* It is easy to see that if  $\lambda \in \mathbb{R}$  is not an eigenvalue for the endomorphism  $\phi$ , then the set  $V_{\lambda} = \{v \in V \mid \phi(v) = \lambda v\}$  contains only the zero vector. It is indeed clear that, if  $V_{\lambda}$  contains the zero vector only, then  $\lambda$  is not an eigenvalue for  $\phi$ . We have that  $\lambda \in \mathbb{R}$  is an eigenvalue for  $\phi$  if and only if  $\dim(V_{\lambda}) \geq 1$ .

**Exercise 9.2.7** Let  $\phi \in \text{End}(\mathbb{R}^2)$  be defined by  $\phi((x, y)) = (y, x)$ . Is  $\lambda = 2$  an eigenvalue for  $\phi$ ? The corresponding set  $V_2$  would then be

$$V_2 = \{v \in \mathbb{R}^2 : \phi(v) = 2v\} = \{(x, y) \in \mathbb{R}^2 : (y, x) = 2(x, y)\},$$

that is,  $V_2$  would be given by the solutions of the system

$$\begin{cases} y = 2x \\ x = 2y \end{cases} \Rightarrow \begin{cases} y = 2x \\ x = 4x \end{cases} \Rightarrow \begin{cases} x = 0 \\ y = 0 \end{cases}.$$

Since  $V_2 = \{(0, 0)\}$ , we conclude that  $\lambda = 2$  is *not* an eigenvalue for  $\phi$ .

**Exercise 9.2.8** The endomorphism  $\phi \in \text{End}(\mathbb{R}^2)$  given by  $\phi((x, y)) = (2x, 3y)$  is simple since the corresponding matrix with respect to the canonical basis  $\mathcal{E} = (e_1, e_2)$  is diagonal,

$$M_{\phi}^{\mathcal{E}, \mathcal{E}} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

Its eigenvalues are  $\lambda_1 = 2$  (with eigenvector  $e_1$ ) and  $\lambda_2 = 3$  (with eigenvector  $e_2$ ). The corresponding eigenspaces are then  $V_2 = \mathcal{L}(e_1)$  and  $V_3 = \mathcal{L}(e_2)$ .

**Exercise 9.2.9** We consider again the endomorphism  $\phi((x, y)) = (y, x)$  in  $\mathbb{R}^2$  given in the Exercise 9.2.7. We wonder whether it is simple. We start by noticing that its corresponding matrix with respect to the canonical basis is the following,

$$M_{\phi}^{\mathcal{E}, \mathcal{E}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

which is not diagonal. We look then for a basis (if it exists) with respect to which the matrix corresponding to  $\phi$  is diagonal. By recalling the Remark 9.2.3 we look for a basis of  $\mathbb{R}^2$  made up of eigenvectors for  $\phi$ . In order for  $v = (a, b)$  to be an eigenvector for  $\phi$ , there must exist a real scalar  $\lambda$  such that  $\phi((a, b)) = \lambda(a, b)$ ,

$$\begin{cases} b = \lambda a \\ a = \lambda b \end{cases}.$$

It follows that the eigenvalues, if they exist, must fulfill the condition  $\lambda^2 = 1$ . For  $\lambda = 1$  the corresponding eigenspace is

$$V_1 = \{(x, y) \in \mathbb{R}^2 : \phi((x, y)) = (x, y)\} = \{(x, x) \in \mathbb{R}^2\} = \mathcal{L}((1, 1)).$$

And for  $\lambda = -1$  the corresponding eigenspace is

$$V_{-1} = \{(x, y) \in \mathbb{R}^2 : \phi((x, y)) = -(x, y)\} = \{(x, -x) \in \mathbb{R}^2\} = \mathcal{L}((1, -1)).$$

Since the vectors  $(1, 1)$ ,  $(1, -1)$  form a basis  $\mathcal{B}$  for  $\mathbb{R}^2$  with respect to which the matrix of  $\phi$  is

$$M_{\phi}^{\mathcal{B}, \mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

we conclude that  $\phi$  is simple. We expect  $M_\phi^{\mathcal{B},\mathcal{B}} \sim M_\phi^{\mathcal{E},\mathcal{E}}$ , since they are associated to the same endomorphism; the algebraic proof of this claim is easy. By defining

$$P = M^{\mathcal{E},\mathcal{B}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

the matrix of the change of basis, we compute explicitly,

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

that is  $P^{-1}M_\phi^{\mathcal{E},\mathcal{E}}P = M_\phi^{\mathcal{B},\mathcal{B}}$  (see the Proposition 9.1.3).

Not any endomorphism is simple as the following exercise shows.

**Exercise 9.2.10** The endomorphism in  $\mathbb{R}^2$  defined as  $\phi((x, y)) = (-y, x)$  is not simple. For  $v = (a, b)$  to be an eigenvector,  $\phi((a, b)) = \lambda(a, b)$  it would be equivalent to  $(-b, a) = \lambda(a, b)$ , leading to  $\lambda^2 = -1$ . The only solution in  $\mathbb{R}$  is then  $a = b = 0$ , showing that  $\phi$  is not simple.

**Proposition 9.2.11** *Let  $V$  be a real vector space with  $\phi \in \text{End}(V)$ . If  $\lambda_1, \lambda_2$  are distinct eigenvalues, any two corresponding eigenvectors,  $0 \neq v_1 \in V_{\lambda_1}$  and  $0 \neq v_2 \in V_{\lambda_2}$ , are linearly independent. Also, the sum  $V_{\lambda_1} + V_{\lambda_2}$  is direct.*

*Proof* Let us assume that  $v_2 = \alpha v_1$ , with  $\mathbb{R} \ni \alpha \neq 0$ . By applying the linear map  $\phi$  to both members, we have  $\phi(v_2) = \alpha\phi(v_1)$ . Since  $v_1$  and  $v_2$  are eigenvectors with eigenvalues  $\lambda_1$  and  $\lambda_2$ ,

$$\begin{aligned} \phi(v_1) &= \lambda_1 v_1 \\ \phi(v_2) &= \lambda_2 v_2 \end{aligned}$$

and the relation  $\phi(v_2) = \alpha\phi(v_1)$ , using  $v_2 = \alpha v_1$  become

$$\lambda_2 v_2 = \alpha(\lambda_1 v_1) = \lambda_1(\alpha v_1) = \lambda_1 v_2,$$

that is

$$(\lambda_2 - \lambda_1)v_2 = 0_V.$$

Since  $\lambda_2 \neq \lambda_1$ , this would lead to the contradiction  $v_2 = 0_V$ . We therefore conclude that  $v_1$  and  $v_2$  are linearly independent.

For the last claim we use the Proposition 2.2.13 and show that  $V_{\lambda_1} \cap V_{\lambda_2} = \{0_V\}$ . If  $v \in V_{\lambda_1} \cap V_{\lambda_2}$ , we could write both  $\phi(v) = \lambda_1 v$  (since  $v \in V_{\lambda_1}$ ) and  $\phi(v) = \lambda_2 v$  (since  $v \in V_{\lambda_2}$ ): it would then be  $\lambda_1 v = \lambda_2 v$ , that is  $(\lambda_1 - \lambda_2)v = 0_V$ . From the hypothesis  $\lambda_1 \neq \lambda_2$ , we would get  $v = 0_V$ .  $\square$

The following proposition is proven along the same lines.

**Proposition 9.2.12** *Let  $V$  be a real vector space, with  $\phi \in \text{End}(V)$ . Let  $\lambda_1, \dots, \lambda_s \in \mathbb{R}$  be distinct eigenvalues of  $\phi$  with  $0_V \neq v_j \in V_{\lambda_j}$ ,  $j = 1, \dots, s$  corresponding eigenvectors. The set  $\{v_1, \dots, v_s\}$  is free, and the sum  $V_{\lambda_1} + \dots + V_{\lambda_s}$  is direct.*

**Corollary 9.2.13** *If  $\phi$  is an endomorphism of the real vector space  $V$ , with  $\dim(V) = n$ , then  $\phi$  has at most  $n$  distinct eigenvalues.*

*Proof* If  $\phi$  had  $s > n$  distinct eigenvalues, there would exist a set  $v_1, \dots, v_s$  of non zero corresponding eigenvectors. From the proposition above, such a system should be free, thus contradicting the fact that the dimension of  $V$  is  $n$ .  $\square$

*Remark 9.2.14* Let  $\phi$  and  $\psi$  be two commuting endomorphisms, that is they are such that  $\phi(\psi(w)) = \psi(\phi(w))$  for any  $v \in V$ . If  $v \in V_\lambda$  is an eigenvector for  $\phi$  corresponding to  $\lambda$ , it follows that

$$\phi(\psi(v)) = \psi(\phi(v)) = \lambda\psi(v).$$

Thus the endomorphism  $\psi$  maps any eigenspace  $V_\lambda$  of  $\phi$  into itself, and analogously  $\phi$  preserves any eigenspace  $V'_\lambda$  of  $\psi$ .

Finding the eigenspaces of an endomorphism amounts to compute suitable kernels. Let  $f : V \rightarrow W$  be a linear map between real vector spaces with bases  $\mathcal{B}$  and  $\mathcal{C}$ . We recall (see Proposition 7.5.1) that if  $A = M_f^{\mathcal{C}, \mathcal{B}}$  and  $\Sigma : AX = 0$  is the linear system associated to  $A$ , the map  $S_\Sigma \rightarrow \ker(f)$  given by

$$(x_1, \dots, x_n) \mapsto (x_1, \dots, x_n)_{\mathcal{B}}$$

is an isomorphism of vector spaces.

**Lemma 9.2.15** *If  $V$  is a real vector space with basis  $\mathcal{B}$ , let  $\phi \in \text{End}(V)$  and  $\lambda \in \mathbb{R}$ . Then*

$$V_\lambda = \ker(\phi - \lambda \text{id}_V) \cong S_{\Sigma_\lambda},$$

where  $S_{\Sigma_\lambda}$  is the space of the solutions of the linear homogeneous system

$$S_{\Sigma_\lambda} : \left( M_\phi^{\mathcal{B}, \mathcal{B}} - \lambda I_n \right) X = 0.$$

*Proof* From the Definition 9.2.4 we write

$$\begin{aligned} V_\lambda &= \{v \in V : \phi(v) = \lambda v\} \\ &= \{v \in V : \phi(v) - \lambda v = 0_V\} \\ &= \ker(\phi - \lambda \text{id}_V). \end{aligned}$$

Such a kernel is isomorphic (as recalled above) to the space of solutions of the linear system given by the matrix  $M_{\phi - \lambda \text{id}_V}^{\mathcal{B}, \mathcal{B}}$ , where  $\mathcal{B}$  is an arbitrary basis of  $V$ . We conclude by noticing that  $M_{\phi - \lambda \text{id}_V}^{\mathcal{B}, \mathcal{B}} = M_\phi^{\mathcal{B}, \mathcal{B}} - \lambda I_n$ .  $\square$

**Proposition 9.2.16** *Let  $\phi \in \text{End}(V)$  be an endomorphism of the real vector space  $V$ , with  $\dim(V) = n$ , and let  $\lambda \in \mathbb{R}$ . The following are equivalent:*

- (i)  $\lambda$  is an eigenvalue for  $\phi$ ,
- (ii)  $\dim(V_\lambda) \geq 1$ ,
- (iii)  $\det(M_\phi^{\mathcal{B},\mathcal{B}} - \lambda I_n) = 0$  for any basis  $\mathcal{B}$  in  $V$ .

*Proof* (i)  $\Leftrightarrow$  (ii) is the content of the Remark 9.2.6;

(ii)  $\Leftrightarrow$  (iii). Let  $\mathcal{B}$  be an arbitrary basis of  $V$ , and consider the linear system

$$S_{\Sigma_\lambda} : \begin{pmatrix} M_\phi^{\mathcal{B},\mathcal{B}} - \lambda I_n \\ X = 0. \end{pmatrix}$$

We have

$$\begin{aligned} \dim(V_\lambda) &= \dim(S_{\Sigma_\lambda}) \\ &= n - \text{rk} \left( M_\phi^{\mathcal{B},\mathcal{B}} - \lambda I_n \right); \end{aligned}$$

the first and the second equality follow from Definition 6.2.1 and Theorem 6.4.3 respectively. From Proposition 5.3.1 we finally write

$$\dim(V_\lambda) \geq 1 \quad \Leftrightarrow \quad \text{rk} \left( M_\phi^{\mathcal{B},\mathcal{B}} - \lambda I_n \right) < n \quad \Leftrightarrow \quad \det \left( M_\phi^{\mathcal{B},\mathcal{B}} - \lambda I_n \right) = 0,$$

which concludes the proof.  $\square$

This proposition shows that the computation of an eigenspace reduces to finding the kernel of a linear map, a computation which has been described in the Proposition 7.5.1.

### 9.3 The Characteristic Polynomial of an Endomorphism

In this section we describe how to compute the eigenvalues of an endomorphism. These will be the roots of a canonical polynomial associate with the endomorphism.

**Definition 9.3.1** Given a square matrix  $A \in \mathbb{R}^{n,n}$ , the expression

$$p_A(T) = \det(A - T I_n)$$

is a polynomial of order  $n$  in  $T$  with real coefficients. Such a polynomial is called the *characteristic polynomial* of the matrix  $A$ .

**Exercise 9.3.2** If  $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$  is a square  $2 \times 2$  matrix, then

$$\begin{aligned} p_A(T) &= \begin{vmatrix} a_{11} - T & a_{12} \\ a_{21} & a_{22} - T \end{vmatrix} \\ &= T^2 - (a_{11} + a_{22})T + (a_{11}a_{22} - a_{12}a_{21}) \\ &= T^2 - (\text{tr}(A))T + (\det(A)). \end{aligned}$$

If  $\lambda_1$  and  $\lambda_2$  are the zeros (the roots) of the polynomial  $p_A(T)$ , with elementary algebra we write

$$p_A(T) = T^2 - (\lambda_1 + \lambda_2)T + \lambda_1\lambda_2$$

thus obtaining

$$\lambda_1 + \lambda_2 = a_{11} + a_{22} = \text{tr}(A), \quad \lambda_1\lambda_2 = (a_{11}a_{22} - a_{12}a_{21}) = \det(A).$$

**Proposition 9.3.3** *Let  $V$  be a real vector space with  $\dim(V) = n$ , and let  $\phi \in \text{End}(V)$ . For any choice of bases  $\mathcal{B}$  and  $\mathcal{C}$  in  $V$ , with corresponding matrices  $A = M_{\phi}^{\mathcal{B},\mathcal{B}}$  and  $B = M_{\phi}^{\mathcal{C},\mathcal{C}}$ , it is*

$$p_A(T) = p_B(T).$$

*Proof* We know that  $B = P^{-1}AP$ , with  $P = M^{\mathcal{B},\mathcal{C}}$  the matrix of change of basis. So we write

$$B - TI_n = P^{-1}AP - P^{-1}(TI_n)P = P^{-1}(A - TI_n)P.$$

From the Binet Theorem 5.1.16 we have then

$$\begin{aligned} \det(B - TI_n) &= \det(P^{-1}(A - TI_n)P) = \det(P^{-1})\det(A - TI_n)\det(P) \\ &= \det(A - TI_n), \end{aligned}$$

which yields a proof of the claim, since  $\det(P^{-1})\det(P) = \det(I_n) = 1$ .  $\square$

Given a matrix  $A \in \mathbb{R}^{n,n}$ , an explicit computation of  $\det(A - TI_n)$  shows that

$$p_A(T) = (-1)^n T^n + (-1)^{n-1} \text{tr}(A) T^{n-1} + \dots + \det(A).$$

The case  $n = 2$  is the Exercise 9.3.2.

Given  $\phi \in \text{End}(V)$ , the Proposition 9.3.3 shows that the characteristic polynomial of the matrix associated to  $\phi$  does not depend on the given basis of  $V$ .

**Definition 9.3.4** For any matrix  $A$  associated to the endomorphism  $\phi \in \text{End}(V)$ , the polynomial  $p_{\phi}(T) = p_A(T)$  is called the *characteristic polynomial* of  $\phi$ .

From the Proposition 9.2.16 and the Definition 9.3.4 we have the following result.

**Corollary 9.3.5** *The eigenvalues of the endomorphism  $\phi \in \text{End}(V)$  (the spectrum of  $\phi$ ) are the real roots of the characteristic polynomial  $p_{\phi}(T)$ .*

**Exercise 9.3.6** Let  $\phi \in \text{End}(\mathbb{R}^2)$  be associated to the matrix

$$M_{\phi}^{\mathcal{E},\mathcal{E}} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Since  $p_{\phi}(T) = T^2 + 1$ , the endomorphism has no (real) eigenvalues.

**Definition 9.3.7** Let  $p(X)$  be a polynomial with real coefficients, and let  $\alpha$  be one of its real root. From the fundamental theorem of algebra (see the Proposition A.5.7) we know that then  $(X - \alpha)$  is a divisor for  $p(X)$ , and that we have the decomposition

$$p(X) = (X - \alpha)^{m(\alpha)} \cdot q(X)$$

where  $q(X)$  is not divisible by  $(X - \alpha)$  and  $1 \leq m(\alpha)$  is an integer depending on  $\alpha$ . Such an integer is called the multiplicity of  $\alpha$ .

**Exercise 9.3.8** Let  $p(X) = (X - 2)(X - 3)(X^2 + 1)$ . Its real roots are 2 (with multiplicity  $m(2) = 1$ , since  $(X - 3)(X^2 + 1)$  cannot be divided by 2) and 3 (with multiplicity  $m(3) = 1$ ). Clearly the polynomial  $p(X)$  has also two imaginary roots, given by  $\pm i$ .

**Proposition 9.3.9** Let  $V$  be a real vector space with  $\phi \in \text{End}(V)$ . If  $\lambda$  is an eigenvalue for  $\phi$  with multiplicity  $m(\lambda)$  and eigenspace  $V_\lambda$ , it holds that

$$1 \leq \dim(V_\lambda) \leq m(\lambda).$$

*Proof* Let  $r = \dim(V_\lambda)$  and  $\mathcal{C}$  be a basis of  $V_\lambda$ . We complete  $\mathcal{C}$  to a basis  $\mathcal{B}$  for  $V$ . We then have  $\mathcal{B} = (v_1, \dots, v_r, v_{r+1}, \dots, v_n)$ , where the first elements  $v_1, \dots, v_r \in V_\lambda$  are eigenvectors for  $\lambda$ . The matrix  $M_\phi^{\mathcal{B}, \mathcal{B}}$  has the following block form,

$$A = M_\phi^{\mathcal{B}, \mathcal{B}} = \begin{pmatrix} \lambda & 0 & \dots & 0 & a_{1,r+1} & \dots & a_{1,n} \\ 0 & \lambda & \dots & 0 & a_{2,r+1} & \dots & a_{2,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda & a_{r,r+1} & \dots & a_{r,n} \\ 0 & 0 & \dots & 0 & a_{r+1,r+1} & \dots & a_{r+1,n} \\ 0 & 0 & \dots & 0 & a_{r+2,r+1} & \dots & a_{r+2,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & a_{n,r+1} & \dots & a_{n,n} \end{pmatrix}.$$

If  $\det(A - T I_n)$  is computed by the Laplace theorem (with respect to the first row, say), we have

$$p_\phi(T) = \det(A - T I_n) = (\lambda - T)^r g(T),$$

where  $g(T)$  is the characteristic polynomial of the lower diagonal  $(n - r) \times (n - r)$  square block of  $A$ . We can then conclude that  $r \leq m(\lambda)$ .  $\square$

**Definition 9.3.10** The integer  $\dim(V_\lambda)$  is called the *geometric multiplicity* of the eigenvalue  $\lambda$ , while  $m(\lambda)$  is called the *algebraic multiplicity* of the eigenvalue  $\lambda$ .

*Remark 9.3.11* Let  $\phi \in \text{End}(V)$ .

(a) If  $\lambda = 0$  is an eigenvalue for  $\phi$ , the corresponding eigenspace  $V_0$  is  $\ker(\phi)$ .

- (b) If  $\lambda \neq 0$  is an eigenvalue for  $\phi$ , then  $V_\lambda \subseteq \text{Im}(\phi)$ :  
 let us indeed consider  $0_V \neq v \in V_\lambda$  with  $\phi(v) = \lambda v$ . Since  $\lambda \neq 0$ , we divide by  $\lambda$  and write

$$v = \lambda^{-1}\phi(v) = \phi(\lambda^{-1}v) \in \text{Im}(\phi).$$

- (c) If  $\lambda_1 \neq \lambda_2 \neq \dots \neq \lambda_s$  are distinct non zero eigenvalues for  $\phi$ , from the Proposition 9.2.12 we have the direct sum of corresponding eigenspaces and

$$V_{\lambda_1} \oplus \dots \oplus V_{\lambda_s} \subseteq \text{Im}(\phi).$$

**Exercise 9.3.12** Let  $\phi \in \text{End}(\mathbb{R}^4)$  be given by

$$\phi((x, y, z, t)) = (2x + 4y, x + 2y, -z - 2t, z + t).$$

The corresponding matrix with respect to the canonical basis  $\mathcal{E}_4$  is

$$A = M_\phi^{\mathcal{E}, \mathcal{E}} = \begin{pmatrix} 2 & 4 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Its characteristic polynomial reads

$$\begin{aligned} p_\phi(T) &= p_A(T) = \det(A - TI_4) \\ &= \begin{vmatrix} 2-T & 4 & 0 & 0 \\ 1 & 2-T & 0 & 0 \\ 0 & 0 & -1-T & -2 \\ 0 & 0 & 1 & 1-T \end{vmatrix} \\ &= \begin{vmatrix} 2-T & 4 \\ 1 & 2-T \end{vmatrix} \begin{vmatrix} -1-T & -2 \\ 1 & 1-T \end{vmatrix} \\ &= T(T-4)(T^2+1). \end{aligned}$$

The eigenvalues (the real roots of such a polynomial) of  $\phi$  are  $\lambda = 0, 4$ . It is easy to compute that

$$\begin{aligned} V_0 &= \ker(\phi) = \mathcal{L}((-2, 1, 0, 0)), \\ V_4 &= \ker(\phi - 4I_4) = \mathcal{L}((2, 1, 0, 0)). \end{aligned}$$

This shows that  $V_4$  is the only eigenspace corresponding to a non zero eigenvalue for  $\phi$ .

From the Theorem 7.6.4 we know that  $\dim \operatorname{Im}(\phi) = 4 - \dim \ker(\phi) = 3$ , with a basis of the image of  $\phi$  given by 3 linearly independent columns in  $A$ . It is immediate to notice that the second column is a multiple of the first one, so we have

$$\operatorname{Im}(\phi) = \mathcal{L}((2, 1, 0, 0), (0, 0, -1, 1), (0, 0, -2, 1)).$$

It is evident that  $V_4 \subset \operatorname{Im}(\phi)$ , as shown in general in the Remark 9.3.11.

**Exercise 9.3.13** We consider the endomorphism in  $\mathbb{R}^4$  given by

$$\phi((x, y, z, t)) = (2x + 4y, x + 2y, -z, z + t),$$

whose corresponding matrix with respect to the canonical basis  $\mathcal{E}_4$  is

$$A = M_{\phi}^{\mathcal{E}, \mathcal{E}} = \begin{pmatrix} 2 & 4 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

The characteristic polynomial reads

$$\begin{aligned} p_{\phi}(T) &= p_A(T) = \det(A - TI_4) \\ &= \begin{vmatrix} 2-T & 4 & 0 & 0 \\ 1 & 2-T & 0 & 0 \\ 0 & 0 & -1-T & 0 \\ 0 & 0 & 1 & 1-T \end{vmatrix} \\ &= T(T-4)(T+1)(T-1). \end{aligned}$$

The eigenvalues are given by  $\lambda = 0, 4, -1, 1$ . The corresponding eigenspaces are

$$\begin{aligned} V_0 &= \ker(\phi) = \mathcal{L}((-2, 1, 0, 0)), \\ V_4 &= \ker(\phi - 4I_4) = \mathcal{L}((2, 1, 0, 0)), \\ V_{-1} &= \ker(\phi + I_4) = \mathcal{L}((0, 0, -2, 1)), \\ V_1 &= \ker(\phi - I_4) = \mathcal{L}((0, 0, 0, 1)), \end{aligned}$$

with

$$\operatorname{Im}(\phi) = V_{-1} \oplus V_1 \oplus V_4.$$

The characteristic polynomial  $p_{\phi}(T)$  of an endomorphism over a real vector space has real coefficients. If  $\lambda_1, \dots, \lambda_s$  are its non zero real distinct roots (that is, the eigenvalues of  $\phi$ ), we can write

$$p_{\phi}(T) = (T - \lambda_1)^{m_1} \cdots (T - \lambda_p)^{m_s} \cdot q(T),$$

where  $m_j$ ,  $j = 1, \dots, s$  are the algebraic multiplicities and  $q(T)$  has no real roots. We have then

$$\deg(p_\phi(T)) \geq m_1 + \dots + m_s.$$

This proves the following proposition.

**Proposition 9.3.14** *Let  $V$  be a real vector space with  $\dim(V) = n$ , and let  $\phi \in \text{End}(V)$ . By denoting  $\lambda_1, \dots, \lambda_s$  the distinct eigenvalues of  $\phi$  with corresponding algebraic multiplicities  $m_1, \dots, m_s$ , one has*

$$m_1 + \dots + m_s \leq n,$$

with the equality holding if and only if every root in  $p_\phi(T)$  is real. □

## 9.4 Diagonalisation of an Endomorphism

In this section we describe conditions under which an endomorphism is simple. As we have seen, this problem is equivalent to study conditions under which a square matrix is diagonalisable. The first theorem we prove characterises simple endomorphisms.

**Theorem 9.4.1** *Let  $V$  be a real  $n$ -dimensional vector space, with  $\phi \in \text{End}(V)$ . If  $\lambda_1, \dots, \lambda_s$  are the different roots of  $p_\phi(T)$  with multiplicities  $m_1, \dots, m_s$ , the following claims are equivalent:*

- (a)  $\phi$  is a simple endomorphism,
- (b)  $V$  has a basis of eigenvectors for  $\phi$ ,
- (c)  $\lambda_i \in \mathbb{R}$  for any  $i = 1, \dots, s$ , with  $V = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_s}$ ,
- (d)  $\lambda_i \in \mathbb{R}$  and  $m_i = \dim(V_{\lambda_i})$  for any  $i = 1, \dots, s$ .

When  $\phi$  is simple, each basis of  $V$  of eigenvectors for  $\phi$  contains  $m_i$  eigenvectors for each distinct eigenvalues  $\lambda_i$ , for  $i = 1, \dots, s$ .

*Proof* • (a)  $\Leftrightarrow$  (b): this has been shown in the Remark 9.2.3.

- (b)  $\Rightarrow$  (c): let  $\mathcal{B} = (v_1, \dots, v_n)$  be a basis of  $V$  of eigenvectors for  $\phi$ . Any vector  $v_i$  belongs to one of the eigenspaces, so we can write

$$V = \mathcal{L}(v_1, \dots, v_n) \subseteq V_{\lambda_1} + \dots + V_{\lambda_s},$$

while the opposite inclusion is obvious. Since the sum of eigenspaces corresponding to distinct eigenvalues is direct (see the Proposition 9.2.12), we have  $V = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_s}$ .

- (c)  $\Rightarrow$  (b): let  $\mathcal{B}_i$  be a basis of  $V_{\lambda_i}$  for any  $i$ . Since  $V$  is the direct sum of all the eigenspaces  $V_{\lambda_i}$ , the set  $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_s$  is a basis of  $V$  made by eigenvectors for  $\phi$ .

- (c)  $\Rightarrow$  (d): from the Grassmann Theorem 2.5.8, we have

$$\begin{aligned} n &= \dim(V) = \dim(V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s}) \\ &= \dim(V_{\lambda_1}) + \cdots + \dim(V_{\lambda_s}) \\ &\leq m_1 + \cdots + m_s \\ &\leq n, \end{aligned}$$

where the inequalities follow from the Propositions 9.3.9 and 9.3.14. We can then conclude that  $\dim(V_{\lambda_i}) = m(\lambda_i)$  for any  $i$ .

- (d)  $\Rightarrow$  (c): from the hypothesis  $m_i = \dim(V_{\lambda_i})$  for any  $i = 1, \dots, s$ , and the Proposition 9.3.14 we have

$$n = m_1 + \cdots + m_s = \dim(V_{\lambda_1}) + \cdots + \dim(V_{\lambda_s}).$$

We have then  $n = \dim(V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s})$  and this equality amounts to prove the claim, since  $V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s}$  has dimension  $n$  and therefore coincides with  $V$ .  $\square$

**Corollary 9.4.2** *If  $\lambda_i \in \mathbb{R}$  and  $m(\lambda_i) = 1$  for any  $i = 1, \dots, n$ , then  $\phi$  is simple.*

*Proof* It is immediate, by recalling the Proposition 9.3.9 and (d) in the Theorem 9.4.1.  $\square$

**Exercise 9.4.3** Let  $\phi$  be the endomorphism in  $\mathbb{R}^2$  whose corresponding matrix with respect to the canonical basis is the matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

It is  $p_A(T) = (1 - T)^2$ : such a polynomial has only one root  $\lambda = 1$  with algebraic multiplicity  $m = 2$ . It is indeed easy to compute that  $V_1 = \mathcal{L}((1, 0))$ , so the geometric multiplicity is 1. This proves that the matrix  $A$  is not diagonalisable, the corresponding endomorphism is not simple.

**Proposition 9.4.4** *Let  $\phi \in \text{End}(V)$  be a simple endomorphism and  $\mathcal{C}$  be a basis of  $V$  such that  $\Delta = M_{\phi}^{\mathcal{C}, \mathcal{C}}$ . Then,*

- the eigenvalues  $\lambda_1, \dots, \lambda_s$  for  $\phi$ , counted with their multiplicities  $m(\lambda_1), \dots, m(\lambda_s)$ , are the diagonal elements for  $\Delta$ ;*
- the diagonal matrix  $\Delta$  is uniquely determined up to permutations of the eigenvalues (such a permutation corresponds to a permutation in the ordering of the basis elements in  $\mathcal{C}$ ).*

*Proof* (a) From the Remark 9.2.1 we know that the diagonal elements in  $\Delta = M_{\phi}^{\mathcal{C}, \mathcal{C}} \in \mathbb{R}^{n, n}$  are given by the eigenvalues  $\lambda_1, \dots, \lambda_s$ : each eigenvalue  $\lambda_i$  must be counted as many times as the geometric multiplicity of the eigenvector  $v_i$ ,

since  $\mathcal{C}$  is a basis of eigenvectors. From the claim (d) in the Theorem 9.4.1, the geometric multiplicity of each eigenvalue coincides with its algebraic multiplicity.

(a) This is obvious.  $\square$

**Proposition 9.4.5** *Let  $\phi$  be a simple endomorphism on  $V$ , with  $\mathcal{B}$  an arbitrary basis of  $V$ . By setting  $A = M_{\phi}^{\mathcal{B},\mathcal{B}}$ , let  $P$  be a matrix such that*

$$P^{-1} A P = \Delta.$$

*Then the columns in  $P$  are the components, with respect to  $\mathcal{B}$ , of a basis of  $V$  made by eigenvectors for  $\phi$ .*

*Proof* Let  $\mathcal{C}$  be a basis of  $V$  such that  $\Delta = M_{\phi}^{\mathcal{C},\mathcal{C}}$ . From the Remark 9.2.3 the basis  $\mathcal{C}$  is made by eigenvectors for  $\phi$ . The claim follows by setting  $P = M^{\mathcal{B},\mathcal{C}}$ , that is the matrix of the change of basis.  $\square$

**Definition 9.4.6** Given a matrix  $A \in \mathbb{R}^{n,n}$ , its *diagonalisation* consists of determining, (if they exist) a diagonal matrix  $\Delta \sim A$  and an invertible matrix  $P \in \text{GL}(n)$  such that  $P^{-1} A P = \Delta$ .

The following remark gives a resumé of the steps needed for the diagonalisation of a given matrix.

**Remark 9.4.7 (An algorithm for the diagonalisation)** Let  $A \in \mathbb{R}^{n,n}$  be a square matrix. In order to diagonalise it:

- (1) Write the characteristic polynomial  $p_A(T)$  of  $A$  and find its roots  $\lambda_1, \dots, \lambda_s$  with the corresponding algebraic multiplicities  $m_1, \dots, m_s$ .
- (2) If one of the roots  $\lambda_i \notin \mathbb{R}$ , then  $A$  is *not* diagonalisable.
- (3) If  $\lambda_i \in \mathbb{R}$  for any  $i = 1, \dots, s$ , compute the geometric multiplicities

$$\dim(V_{\lambda_i}) = n - \text{rk}(A - \lambda_i I_n).$$

If there is an eigenvalue  $\lambda_i$  such that  $m_i \neq \dim(V_{\lambda_i})$ , then  $A$  is *not* diagonalisable.

- (4) if  $\lambda_i \in \mathbb{R}$  and  $m(\lambda)_i = \dim(V_{\lambda_i})$  for any  $i = 1, \dots, s$ , then  $A$  is diagonalisable. In such a case,  $A$  is similar to a diagonal matrix  $\Delta$ : the eigenvalues  $\lambda_i$ , counted with their multiplicities, give the diagonal elements for  $\Delta$ .
- (5) it is  $\Delta = M_{\phi}^{\mathcal{B},\mathcal{B}}$ , where  $\mathcal{B}$  is a basis of  $V$  given by eigenvectors for the endomorphism corresponding to the matrix  $A$ . By defining  $P = M^{\mathcal{E},\mathcal{B}}$ , it is  $\Delta = P^{-1} A P$ . Since  $V$  is the direct sum of the eigenspaces for  $A$  (see Theorem 9.4.1), it follows that  $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_s$ , with  $\mathcal{B}_i$  a basis of  $V_{\lambda_i}$  for any  $i = 1, \dots, s$ . (The spaces  $V_{\lambda_i}$  can be obtained explicitly as in the Lemma 9.2.15.)

**Exercise 9.4.8** We study whether the matrix

$$A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix}$$

is diagonalisable. Its characteristic polynomial is

$$\begin{aligned} p_A(T) &= \det(A - TI_3) \\ &= \begin{vmatrix} 3-T & 1 & 1 \\ 1 & -T & 2 \\ 1 & 2 & -T \end{vmatrix} \\ &= -T^3 + 3T^2 + 6T - 8 = (T-1)(T-4)(T+2). \end{aligned}$$

Its eigenvalues are found to be  $\lambda_1 = 1$ ,  $\lambda_2 = 4$ ,  $\lambda_3 = -2$ . Since each root of the characteristic polynomial has algebraic multiplicity  $m = 1$ , from the Corollary 9.4.2 the matrix  $A$  is diagonalisable, and indeed similar to

$$\Delta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

We compute a basis  $\mathcal{B}$  for  $\mathbb{R}^3$  of eigenvectors for  $A$ . We know that  $V_1 = \ker(A - I_3)$ , so  $V_1$  is the space of the solutions of the homogeneous linear system  $(A - I_3)X = 0$  associated to the matrix

$$A - I_3 = \begin{pmatrix} 2 & 1 & 1 \\ 1 & -1 & 2 \\ 1 & 2 & -1 \end{pmatrix},$$

which is reduced to

$$\begin{pmatrix} 2 & 1 & 1 \\ 3 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}.$$

The solution of such a linear system are given by  $(x, y, z) = (x, -x, -x)$ , thus  $V_1 = \mathcal{L}((-1, 1, 1))$ . Along the same lines we compute

$$\begin{aligned} V_4 &= \ker(A - 4I_3) = \mathcal{L}((2, 1, 1)), \\ V_{-2} &= \ker(A + 2I_3) = \mathcal{L}((0, -1, 1)). \end{aligned}$$

We have then  $\mathcal{B} = ((-1, 1, 1), (2, 1, 1), (0, -1, 1))$  and

$$P = M^{\mathcal{E}, \mathcal{B}} = \begin{pmatrix} -1 & 2 & 0 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}.$$

It is easy to compute that  $P^{-1} A P = \Delta$ .

**Proposition 9.4.9** *Let  $A \in \mathbb{R}^{n,n}$  be diagonalisable, with eigenvalues  $\lambda_1, \dots, \lambda_s$  and corresponding multiplicities  $m_1, \dots, m_s$ . Then*

$$\det(A) = \lambda_1^{m_1} \cdot \lambda_2^{m_2} \cdot \dots \cdot \lambda_s^{m_s},$$

$$\operatorname{tr}(A) = m_1 \lambda_1 + m_2 \lambda_2 + \dots + m_s \lambda_s.$$

*Proof* Since  $A$  is diagonalisable, there exists an invertible  $n$ -dimensional matrix  $P$  such that  $\Delta = P^{-1} A P$ . The matrix  $\Delta$  is diagonal, and its diagonal elements are (see the Proposition 9.4.4) the eigenvalues of  $A$  counted with their multiplicities. Then, from the Proposition 9.1.5 on has,

$$\det(A) = \det(P^{-1} A P) = \det(\Delta) = \lambda_1^{m_1} \cdot \lambda_2^{m_2} \cdot \dots \cdot \lambda_s^{m_s}$$

and

$$\operatorname{tr}(A) = \operatorname{tr}(P^{-1} A P) = \operatorname{tr}(\Delta) = m_1 \lambda_1 + m_2 \lambda_2 + \dots + m_s \lambda_s.$$

□

## 9.5 The Jordan Normal Form

In this section we briefly describe the notion of *Jordan normal form* of a matrix. As we have described before in this chapter, a square matrix is not necessarily diagonalisable, that is it is not necessarily similar to a diagonal matrix. It is nonetheless possible to prove that any square matrix is similar to a triangular matrix  $J$  which is not far from being diagonal. Such a matrix  $J$  is diagonal if and only if  $A$  is diagonalisable; if not it has a ‘standard’ block structure.

An example of a so called *Jordan block* is the non diagonalisable matrix  $A$  in Exercise 9.4.3. We denote it by

$$J_2(1) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

A Jordan block of order  $k$  is a  $k$ -dimensional upper triangular square matrix of the form

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix},$$

where the diagonal terms are given by a scalar  $\lambda \in \mathbb{R}$ , the  $(J_k(\lambda))_{j,j+1}$  entries are 1 and the remaining entries are zero. It is immediate to show that the characteristic polynomial of such a matrix is given by

$$p_{J_k(\lambda)}(T) = (T - \lambda)^k,$$

and the parameter  $\lambda$  is the unique eigenvalue with algebraic multiplicity  $m_\lambda = k$ . The corresponding eigenspace is

$$V_\lambda = \ker(J_k(\lambda) - \lambda I_n) = \mathcal{L}((1, 0, \dots, 0)),$$

with geometric multiplicity  $\dim(V_\lambda) = 1$ . Thus, if  $k > 1$ , a Jordan block is not diagonalisable.

A matrix  $J$  is said to be in (*canonical* or *normal*) *Jordan form* if it has a block diagonal form

$$J = \begin{pmatrix} J_{k_1}(\lambda_1) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & J_{k_2}(\lambda_2) & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & J_{k_s}(\lambda_s) \end{pmatrix},$$

where each  $J_{k_j}(\lambda_j)$  is a Jordan block of order  $k_j$  and eigenvalue  $\lambda_j$ , for  $j = 1, \dots, s$ .

Notice that nothing prevents from having the same eigenvalue in different Jordan blocks, that is  $\lambda_j = \lambda_l$  even with  $k_j \neq k_l$ . Since each Jordan block  $J_{k_j}(\lambda_j)$  provides a one dimensional eigenspace for  $\lambda_j$ , the geometric multiplicity of  $\lambda_j$  coincides with the number of Jordan blocks with eigenvalue  $\lambda_j$ . The algebraic multiplicity of  $\lambda_j$  coincides indeed with the sum of the orders of the Jordan blocks having the same eigenvalue  $\lambda_j$ .

**Theorem 9.5.1** (Jordan) *Let  $A \in \mathbb{R}^{n \times n}$  such that its characteristic polynomial has only real roots (such roots are all the eigenvalues for  $A$ ). Then,*

- (i) *the matrix  $A$  is similar to a Jordan matrix,*
- (ii) *two Jordan matrices  $J$  and  $J'$  are similar if and only if one is mapped into the other under a block permutation.*

We omit a complete proof of this theorem, and we limit ourselves to briefly introduce the notion of *generalised* eigenvector of a matrix  $A$ . We recall that, when  $A$  is not diagonalisable, the set of eigenvectors for  $A$  is not enough for a basis of  $\mathbb{R}^n$ . The columns of the invertible matrix  $P$  that realises the similarity between  $A$  and the Jordan form  $J$  (such that  $P^{-1}AP = J$ ) are the components with respect to the canonical basis  $\mathcal{E}_n$  of the so called generalised eigenvectors for  $A$ .

Given an eigenvalue  $\lambda$  for  $A$  with algebraic multiplicity  $m_\lambda \geq 1$ , a corresponding generalised eigenvector is a non zero vector  $v$  that solves the linear homogeneous system

$$(A - \lambda I_n)^m v = 0_{\mathbb{R}^n}.$$

It is possible to show that such a system has  $m$  solutions  $v_j$  (with  $v_j = 1, \dots, m$ ) which can be obtained by recursion,

$$\begin{aligned} (A - \lambda I_n)v_1 &= 0_{\mathbb{R}^n}, \\ (A - \lambda I_n)v_k &= v_{k-1}, \quad k = 2, \dots, m. \end{aligned}$$

The elements  $v_j$  span the *generalised eigenspace*  $V_\lambda$  for  $A$  corresponding to the eigenvalue  $\lambda$ . The generalised eigenvectors satisfy the condition

$$(A - \lambda I_n)^k v_k = 0_{\mathbb{R}^n} \quad \text{for any } k = 1, 2, \dots, m.$$

Since the characteristic polynomial of  $A$  has in general complex roots, we end by noticing that a more natural version of the Jordan theorem is valid on  $\mathbb{C}$ .

**Exercise 9.5.2** We consider the matrix

$$A = \begin{pmatrix} 5 & 4 & 2 & 1 \\ 0 & 1 & -1 & -1 \\ -1 & -1 & 3 & 0 \\ 1 & 1 & -1 & 2 \end{pmatrix}.$$

Its characteristic polynomial is computed to be  $p_A(T) = (T - 1)(T - 2)(T - 4)^2$ , so its eigenvalues are  $\lambda = 1, 2, 4, 4$ . Since the algebraic multiplicity of the eigenvalues  $\lambda = 1$  and  $\lambda = 2$  is 1, their geometric multiplicity is also 1. An explicit computation shows that

$$\dim(\ker(A - 4I_4)) = 1.$$

We have then that  $A$  is not diagonalisable, and that the eigenvalue  $\lambda = 4$  corresponds to a Jordan block. A canonical form for the matrix  $A$  is then given by

$$J = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

**Exercise 9.5.3** The matrices

$$J = \begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}, \quad J' = \begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

have the same characteristic polynomial, the same determinant, and the same trace. They are however *not* similar, since they are in Jordan form, and there is no block permutation under which  $J$  is mapped into  $J'$ .