Chapter 10
Spectral Theorems on Euclidean Spaces

In Chap. 7 we studied the operation of changing a basis for a real vector space. In particular, in the Theorem 7.9.6 and the Remark 7.9.7 there, we showed that any matrix giving a change of basis for the vector space $\mathbb{R}^n$ is an invertible $n \times n$ matrix, and noticed that any $n \times n$ invertible yields a change of basis for $\mathbb{R}^n$.

In this chapter we shall consider the endomorphisms of the euclidean space $E^n = (\mathbb{R}^n, \cdot)$, where the symbol $\cdot$ denotes the euclidean scalar product, that we have described in Chap. 3.

10.1 Orthogonal Matrices and Isometries

As we noticed, the natural notion of basis for a euclidean space is that of orthonormal one. This restricts the focus to matrices which gives a change of basis between orthonormal bases for E^n.

Definition 10.1.1 A square matrix $A \in \mathbb{R}^{n,n}$ is called *orthogonal* if its columns form an orthonormal basis $\mathcal{B}$ for E^n. In such a case $A = M^{\mathcal{E},\mathcal{B}}$, that is A is the matrix giving the change of basis from the canonical basis $\mathcal{E}$ to the basis $\mathcal{B}$.

It follow from this definition that an orthogonal matrix is invertible.

Exercise 10.1.2 The identity matrix I_n is clearly orthogonal for each E^n. Since the vectors

$$v_1 = \tfrac{1}{\sqrt{2}}(1, 1), \qquad v_2 = \tfrac{1}{\sqrt{2}}(1, -1)$$

© Springer International Publishing AG, part of Springer Nature 2018
G. Landi and A. Zampini, *Linear Algebra and Analytic Geometry for Physical Sciences*, Undergraduate Lecture Notes in Physics,
https://doi.org/10.1007/978-3-319-78361-1_10

form an orthonormal basis for E^2, the matrix

$$A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

is orthogonal.

Proposition 10.1.3 *A matrix A is orthogonal if and only if*

$$^tA\,A = I_n,$$

that is if and only if $A^{-1} = {}^tA$.

Proof With $(v_1, \dots, v_n)$ a system of vectors in E^n, we denote by $A = (v_1 \cdots v_n)$ the matrix with columns given by the given vectors, and by

$$^tA = \begin{pmatrix} ^tv_1 \\ \vdots \\ ^tv_n \end{pmatrix}$$

its transpose. We have the following equivalences. The matrix A is orthogonal (by definition) if and only if $(v_1, \dots, v_n)$ is an orthonormal basis for E^n, that is if and only if $v_i \cdot v_j = \delta_{ij}$ for any i, j. Recalling the representation of the row by column product of matrices, one has $v_i \cdot v_j = \delta_{ij}$ if and only if $({}^tAA)_{ij} = \delta_{ij}$ for any i, j, which amounts to say that ${}^tAA = I_n$. □

Exercise 10.1.4 For the matrix A considered in the Exercise 10.1.2 one has easily compute that $A = {}^tA$ and $A^2 = I_2$.

Exercise 10.1.5 The matrix

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

is not orthogonal, since

$$^tAA = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \neq I_2.$$

Proposition 10.1.6 *If A is orthogonal, then* $\det(A) = \pm 1$.

Proof This statement easily follows from the Binet Theorem 5.1.16: with ${}^tAA = I_n$, one has

$$\det({}^tA)\det(A) = \det(I_n) = 1,$$

and the Corollary 5.1.12, that is $\det({}^tA) = \det(A)$, which then implies $(\det(A))^2 = 1$. □

Remark 10.1.7 The converse to this statement does not hold. The matrix A from the Exercise 10.1.5 is not orthogonal, while $\det(A) = 1$.

Definition 10.1.8 An orthogonal matrix A with $\det(A) = 1$ is called *special orthogonal*.

Proposition 10.1.9 *The set* $\mathrm{O}(n)$ *of orthogonal matrices in* $\mathbb{R}^{n,n}$ *is a group, with respect to the usual matrix product. Its subset* $\mathrm{SO}(n) = \{A \in \mathrm{O}(n) : \det(A) = 1\}$ *is a subgroup of* $\mathrm{O}(n)$ *with respect to the same product.*

Proof We prove that $\mathrm{O}(n)$ is stable under the matrix product, has an identity element, and the inverse of an orthogonal matrix is orthogonal as well.

- The identity matrix I_n is orthogonal, as we already mentioned.
- If A and B are orthogonal, then we can write

$$\begin{aligned} {}^t(AB)AB &= {}^tB\,{}^tAAB \\ &= {}^tB\, I_n\, B \\ &= {}^tBB = I_n, \end{aligned}$$

  that is, AB is orthogonal.
- If A is orthogonal, ${}^tAA = I_n$, then

$$ {}^t(A^{-1})A^{-1} = (A\,{}^tA)^{-1} = I_n, $$

  that proves that A^{-1} is orthogonal.

From the Binet theorem it easily follows that the set of special orthogonal matrices is stable under the product, and the inverse of a special orthogonal matrix is special orthogonal. □

Definition 10.1.10 The group $\mathrm{O}(n)$ is called the *orthogonal group* of order n, its subset $\mathrm{SO}(n)$ is called the *special orthogonal group* of order n.

We know from the Definition 10.1.1 that a matrix is orthogonal if and only if it is the matrix of the change of basis between the canonical basis $\mathcal{E}$ (which is orthonormal) and a second orthonormal basis $\mathcal{B}$. A matrix A is then orthogonal if and only if $A^{-1} = {}^tA$ (Proposition 10.1.3).

The next theorem shows that we do not need the canonical basis. If one defines a matrix A to be orthogonal by the condition $A^{-1} = {}^tA$, then A is the matrix for a change between two orthonormal bases and viceversa, any matrix A giving the change between orthonormal bases satisfies the condition $A^{-1} = {}^tA$.

Theorem 10.1.11 *Let* $\mathcal{C}$ *be an orthonormal basis for the euclidean vector space* E^n, *with* $\mathcal{B}$ *another (arbitrary) basis for it. The matrix* $M^{\mathcal{C},\mathcal{B}}$ *of the change of basis from* $\mathcal{C}$ *to* $\mathcal{B}$ *is orthogonal if and only if also the basis* $\mathcal{B}$ *is orthonormal.*

Proof We start by noticing that, since $\mathcal{C}$ is an orthonormal basis, the matrix $M^{\mathcal{E},\mathcal{C}}$ giving the change of basis between the canonical basis $\mathcal{E}$ and $\mathcal{C}$ is orthogonal by the Definition 10.1.1. It follows that, being $\mathrm{O}(n)$ a group, the inverse $M^{\mathcal{C},\mathcal{E}} = (M^{\mathcal{E},\mathcal{C}})^{-1}$ is orthogonal. With $\mathcal{B}$ an arbitrary basis, from the Theorem 7.9.9 we can write

$$\begin{aligned} M^{\mathcal{C},\mathcal{B}} &= M^{\mathcal{C},\mathcal{E}}\, M^{\mathcal{E},\mathcal{E}}\, M^{\mathcal{E},\mathcal{B}} \\ &= M^{\mathcal{C},\mathcal{E}}\, I_n\, M^{\mathcal{E},\mathcal{B}} = M^{\mathcal{C},\mathcal{E}} M^{\mathcal{E},\mathcal{B}}. \end{aligned}$$

Firstly, let us assume $\mathcal{B}$ to be orthonormal. We have then that $M^{\mathcal{E},\mathcal{B}}$ is orthogonal; thus $M^{\mathcal{C},\mathcal{B}}$ is orthogonal since it is the product of orthogonal matrices.

Next, let us assume that $M^{\mathcal{C},\mathcal{B}}$ is orthogonal; from the chain relations displayed above we have

$$M^{\mathcal{E},\mathcal{B}} = (M^{\mathcal{C},\mathcal{E}})^{-1} M^{\mathcal{C},\mathcal{B}} = M^{\mathcal{E},\mathcal{C}} M^{\mathcal{C},\mathcal{B}}.$$

This matrix $M^{\mathcal{E},\mathcal{B}}$ is then orthogonal (being the product of orthogonal matrices), and therefore $\mathcal{B}$ is an orthonormal basis. □

We pass to endomorphisms corresponding to orthogonal matrices. We start by recalling, from the Definition 3.1.4, that a scalar product has a 'canonical' form when it is given with respect to orthonormal bases.

Remark 10.1.12 Let $\mathcal{C}$ be an orthonormal basis for the euclidean space E^n. If $v, w \in E^n$ are given by $v = (x_1, \ldots, x_n)_{\mathcal{C}}$ and $w = (y_1, \ldots, y_n)_{\mathcal{C}}$, one has that $v \cdot w = x_1 y_1 + \cdots + x_n y_n$. By denoting X and Y the one-column matrices whose entries are the components of v, w with respect to $\mathcal{C}$, that is

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

we can write

$$v \cdot w = x_1 y_1 + \cdots + x_n y_n = \begin{pmatrix} x_1 & \ldots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = {}^t X Y.$$

Theorem 10.1.13 *Let $\phi \in \mathrm{End}(E^n)$, with $\mathcal{E}$ the canonical basis of E^n. The following statements are equivalent:*

(i) The matrix $A = M_{\phi}^{\mathcal{E},\mathcal{E}}$ is orthogonal.
(ii) It holds that $\phi(v) \cdot \phi(w) = v \cdot w$ for any $v, w \in E^n$.
(iii) If $\mathcal{B} = (b_1, \ldots, b_n)$ is an orthonormal basis for E^n, then the set $\mathcal{B}' = (\phi(b_1), \ldots, \phi(b_n))$ is such.

Proof (i) $\Rightarrow$ (ii): by denoting $X = {}^t v$ and $Y = {}^t w$ we can write

$$v \cdot w = {}^t XY, \qquad \phi(v) \cdot \phi(w) = {}^t(AX)(AY) = {}^t X({}^t AA)Y,$$

and since A is orthogonal, ${}^t AA = I_n$, we conclude that $\phi(v) \cdot \phi(w) = v \cdot w$ for any $v, w \in E^n$.

(ii) $\Rightarrow$ (iii): let $A = M_\phi^{\mathcal{C},\mathcal{C}}$ be the matrix of the endomorphism ϕ with respect to the basis $\mathcal{C}$. We start by proving that A is invertible. By adopting the notation used above, we can represent the condition $\phi(v) \cdot \phi(w) = v \cdot w$ as ${}^t(AX)(AY) = {}^t XY$ for any $X, Y \in E^n$. It follows that ${}^t AA = I_n$, that is A is orthogonal, and then invertible. This means (see Theorem 7.8.4) that ϕ is an isomorphism, so it maps a basis for E^n into a basis for E^n. If $\mathcal{B}$ is an orthonormal basis, then we can write

$$\phi(b_i) \cdot \phi(b_j) = b_i \cdot b_j = \delta_{ij}$$

which proves that $\mathcal{B}'$ is an orthonormal basis.

(iii) $\Rightarrow$ (i): since $\mathcal{E}$, the canonical basis for E^n, is orthonormal, then $(\phi(e_1), \ldots, \phi(e_n))$ is orthonormal. Recall the Remark 7.1.10: the components with respect to $\mathcal{E}$ of the elements $\phi(e_i)$ are the column vectors of the matrix $M_\phi^{\mathcal{E},\mathcal{E}}$, thus $M_\phi^{\mathcal{E},\mathcal{E}}$ is orthogonal. □

We have seen that, if the action of $\phi \in \mathrm{End}(E^n)$ is represented with respect to the canonical basis by an orthogonal matrix, then ϕ is an isomorphism and preserves the scalar product, that is, for any $v, w \in E^n$ one has that,

$$v \cdot w = \phi(v) \cdot \phi(w).$$

The next result is therefore evident.

Corollary 10.1.14 *If* $\phi \in \mathrm{End}(E^n)$ *is an endomorphism of the euclidean space* E^n *whose corresponding matrix with respect to the canonical basis is orthogonal then* ϕ *preserves the norms, that is, for any* $v \in E^n$ *one has*

$$\|\phi(v)\| = \|v\|.$$

This is the reason why such an endomorphism is also called an isometry.

The analysis we developed so far allows us to introduce the following definition, which will be more extensively scrutinised when dealing with rotations maps.

Definition 10.1.15 If $\phi \in \mathrm{End}(E^n)$ takes the orthonormal basis $\mathcal{B} = (b_1, \ldots, b_n)$ to the orthonormal basis $\mathcal{B}' = (b'_1 = \phi(b_1), \ldots, b'_n = \phi(b_n))$ in E^n, we say that $\mathcal{B}$ and $\mathcal{B}'$ have the same *orientation* if the matrix representing the endomorphism ϕ is special orthogonal.

Remark 10.1.16 It is evident that this definition provides an equivalence relation within the collection of all orthonormal bases for E^n. The corresponding quotient can be labelled by the values of the determinant of the orthogonal map giving the change of basis, that is $\det\phi = \{\pm 1\}$. This is usually referred to by saying that the euclidean space E^n has two orientations.

10.2 Self-adjoint Endomorphisms

We need to introduce an important class of endomorphisms.

Definition 10.2.1 An endomorphism ϕ of the euclidean vector space E^n is called *self-adjoint* if

$$\phi(v)\cdot w = v\cdot\phi(w) \qquad \forall\, v, w \in E.$$

From the Proposition 9.2.11 we know that eigenvectors corresponding to distinct eigenvalues are linearly independent. When dealing with self-adjoint endomorphisms, a stronger property holds.

Proposition 10.2.2 *Let ϕ be a self-adjoint endomorphism of E^n, with $\lambda_1, \lambda_2 \in \mathbb{R}$ different eigenvalues for it. Any two corresponding eigenvectors, $0 \neq v_1 \in V_{\lambda_1}$ and $0 \neq v_2 \in V_{\lambda_2}$, are orthogonal.*

Proof Since ϕ is self-adjoint, one has $\phi(v_1)\cdot v_2 = v_1\cdot\phi(v_2)$ while, v_1 and v_2 being eigenvectors, one has $\phi(v_i) = \lambda_i v_i$ for $i = 1, 2$. We can then write

$$(\lambda_1 v_1)\cdot v_2 = v_1\cdot(\lambda_2 v_2)$$

which reads

$$\lambda_1(v_1\cdot v_2) = \lambda_2(v_1\cdot v_2) \qquad \Rightarrow \quad (\lambda_2-\lambda_1)(v_1\cdot v_2) = 0.$$

The assumption that the eigenvalues are different allows one to conclude that $v_1\cdot v_2 = 0$, that is v_1 is orthogonal to v_2. □

The self-adjointness of an endomorphism can be characterised in terms of properties of the matrices representing its action on E^n. We recall from the Definition 4.1.21 that a matrix $A = (a_{ij}) \in \mathbb{R}^{n,n}$ is called symmetric if ${}^tA = A$, that is if one has $a_{ij} = a_{ji}$, for any i, j.

Theorem 10.2.3 *Let $\phi \in \mathrm{End}(E^n)$ and $\mathcal{B}$ an orthonormal basis for E^n. The endomorphism ϕ is self-adjoint if and only if $M_\phi^{\mathcal{B},\mathcal{B}}$ is symmetric.*

Proof Using the usual notation, we set $A = (a_{ij}) = M_{\phi}^{\mathcal{B},\mathcal{B}}$ and X, Y be the columns giving the components with respect to $\mathcal{B}$ of the vectors v, w in E^n. From the Remark 10.1.12 we write

$$\begin{aligned} & \phi(v)\cdot w = {}^t(AX)Y = ({}^tX{}^tA)Y = {}^tX{}^tAY \\ \text{and} \quad & v\cdot\phi(w) = {}^tX(AY) = {}^tXAY. \end{aligned}$$

Let us assume A to be symmetric. From the relations above we conclude that $\phi(v)\cdot w = v\cdot\phi(w)$ for any $v, w \in E^n$, that is ϕ is self-adjoint.

If we assume ϕ to be self-adjoint, then we can equate

$${}^tX{}^tAY = {}^tXAY$$

for any X, Y in $\mathbb{R}^n$. If we let X and Y to range on the elements of the canonical basis $\mathcal{E} = (e_1, \ldots, e_n)$ in $\mathbb{R}^n$, such a condition is just the fact that $a_{ij} = a_{ji}$ for any i, j, that is A is symmetric. □

Exercise 10.2.4 The following matrix is symmetric:

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

Then the endomorphism $\phi \in \mathrm{End}(E^2)$ corresponding to A with respect to the canonical basis is self-adjoint. This can also be shown by a direct calculation: $\phi((x, y)) = (2x - y, -x + 3y)$; then

$$\begin{aligned} (a, b)\cdot\phi((x, y)) &= a(2x - y) + b(-x + 3y) \\ &= (2a - b)x + (-a + 3b)y \\ &= \phi((a, b))\cdot(x, y). \end{aligned}$$

Exercise 10.2.5 The following matrix is not symmetric

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

The corresponding (with respect to the canonical basis) endomorphism $\phi \in \mathrm{End}(E^2)$ is indeed not self-adjoint since for instance,

$$\begin{aligned} \phi(e_1)\cdot e_2 &= (1, -1)\cdot(0, 1) = -1, \\ e_1\cdot\phi(e_2) &= (1, 0)\cdot(1, 0) = 1. \end{aligned}$$

An important family of self-adjoint endomorphisms is illustrated in the following exercise.

Exercise 10.2.6 We know from Sect. 8.2 that, if $\mathcal{B} = (e_1, \ldots, e_n)$ is an orthonormal basis for E^n, then the action of an endomorphism ϕ whose associated matrix is $\Phi = M_{\phi}^{\mathcal{B},\mathcal{B}}$ can be written with the Dirac's notation as

$$\phi = \sum_{a,b=1}^{n} \Phi_{ab} |e_a\rangle\langle e_b|,$$

with $\Phi_{ab} = \langle e_a|\phi(e_b)\rangle$. Then, the endomorphism ϕ is self-adjoint if and only if $\Phi_{ab} = \Phi_{ba}$. Consider vectors $u = (u_1, \ldots, u_n)_{\mathcal{B}}$, $v = (v_1, \ldots, v_n)_{\mathcal{B}}$ in E^n, and define the operator $L = |u\rangle\langle v|$. We have

$$\langle e_a|Le_b\rangle = \langle e_a|u\rangle\langle v|e_b\rangle = u_a v_b,$$
$$\langle e_b|Le_a\rangle = \langle e_b|u\rangle\langle v|e_a\rangle = u_b v_a,$$

so we conclude that the operator $L = |u\rangle\langle v|$ is self-adjoint if and only if $u = v$.

Exercise 10.2.7 Let ϕ be a self-adjoint endomorphism of the euclidean space E^n, and let the basis $\mathcal{B} = (e_1, \ldots, e_n)$ made of orthonormal eigenvectors for ϕ with corresponding eigenvalues $(\lambda_1, \ldots, \lambda_n)$ (not necessarily all distinct). A direct computation shows that, in the Dirac's notation, the action of ϕ can be written as

$$\phi = \lambda_1 |e_1\rangle\langle e_1| + \cdots + \lambda_n |e_n\rangle\langle e_n|,$$

so that, for any $v \in E^n$, one writes

$$\phi(v) = \lambda_1 |e_1\rangle\langle e_1|v\rangle + \cdots + \lambda_n |e_n\rangle\langle e_n|v\rangle.$$

10.3 Orthogonal Projections

As we saw in Chap. 3, given any vector subspace $W \subset E^n$, with orthogonal complement $W^{\perp}$ we have a direct sum decomposition $E^n = W \oplus W^{\perp}$, so for any vector $v \in E^n$ we have (see the Proposition 3.2.5) a unique decomposition $v = v_W + v_{W^{\perp}}$. This suggests the following definition.

Definition 10.3.1 Given the (canonical) euclidean space E^n with $W \subset E^n$ a vector subspace and the orthogonal sum decomposition $v = v_W + v_{W^{\perp}}$, the map

$$P_W \,:\, E^n \to E^n, \qquad v \mapsto u_W$$

is linear, and it is called the *orthogonal projection onto the subspace* W. The dimension of W is called the *rank* of the orthogonal projection P_W.

If $W \subset E^n$ it is easy to see that $\mathrm{Im}(P_W) = W$ while $\ker(P_W) = W^{\perp}$. Moreover, since P_W acts as an identity operator on its range W, one also has $P_W^2 = P_W$. If

u, v are vectors in E^n, with orthogonal sum decomposition $u = u_W + u_{W^\perp}$ and $v = v_W + v_{W^\perp}$, we can explicitly compute

$$\begin{aligned} P_W(u) \cdot v &= u_W \cdot (v_W + v_{W^\perp}) \\ &= u_W \cdot v_W \qquad \text{and} \\ u \cdot P_W(v) &= (u_W + u_{W^\perp}) \cdot v_W \\ &= u_W \cdot v_W . \end{aligned}$$

This shows that orthogonal projectors are self-adjoint endomorphisms. To which extent can one reverse these computations, that is can one characterise, within all self-adjoint endomorphisms, the collection of orthogonal projectors? This is the content of the next proposition.

Proposition 10.3.2 *Given the euclidean vector space* E^n*, an endomorphism* $\phi \in \mathrm{End}(E^n)$ *is an orthogonal projection if and only if it is self-adjoint and satisfies the condition* $\phi^2 = \phi$.

Proof We have already shown that the conditions are necessary for an endomorphism to be an orthogonal projection in E^n. Let us now assume that ϕ is a self-adjoint endomorphism fulfilling $\phi^2 = \phi$. For any choice of $u, v \in E^n$ we have

$$\begin{aligned} ((1-\phi)(u)) \cdot \phi(v) &= u \cdot \phi(v) - \phi(u) \cdot \phi(v) \\ &= u \cdot \phi(v) - u \cdot \phi^2(v) \\ &= u \cdot \phi(v) - u \cdot \phi(v) = 0 \end{aligned}$$

with the second line coming from the self-adjointness of ϕ and the third line from the condition $\phi^2 = \phi$. This shows that the vector subspace $\mathrm{Im}(1-\phi)$ is orthogonal to the vector subspace $\mathrm{Im}(\phi)$. We can then decompose any vector $y \in E^n$ as an orthogonal sum $y = y_{\mathrm{Im}(1-\phi)} + y_{\mathrm{Im}\phi} + \xi$, where ξ is an element in the vector subspace orthogonal to the sum $\mathrm{Im}(1-\phi) \oplus \mathrm{Im}(\phi)$. For any $u \in E^n$ and any such vector ξ we have

$$\phi(u) \cdot \xi = 0, \qquad ((1-\phi)(u)) \cdot \xi = 0.$$

These conditions give that $u \cdot \xi = 0$ for any $u \in E^n$, so we can conclude that $\xi = 0$. Thus we have the orthogonal vector space decomposition

$$E^n = \mathrm{Im}(1-\phi) \oplus \mathrm{Im}(\phi).$$

We show next that $\ker(\phi) = \mathrm{Im}(1-\phi)$. If $u \in \mathrm{Im}(1-\phi)$, we have $u = (1-\phi)v$ with $v \in E^n$, thus $\phi(u) = \phi(1-\phi)v = 0$, that is $\mathrm{Im}(1-\phi) \subseteq \ker(\phi)$. Conversely, if $u \in \ker(\phi)$, then $\phi(u) \cdot v = 0$ for any $v \in E^n$, and $u \cdot \phi(v) = 0$, since ϕ is self-adjoint, which gives $\ker(\phi) \subseteq (\mathrm{Im}(\phi))^\perp$ and $\ker(\phi) \subseteq \mathrm{Im}(1-\phi)$, from the decomposition of E^n above.

If $w \in \mathrm{Im}(\phi)$, then $w = \phi(x)$ for a given $x \in E^n$, thus $\phi(w) = \phi^2(x) = \phi(x) = w$. We have shown that we can identify $\phi = P_{\mathrm{Im}(\phi)}$. This concludes the proof. □

Exercise 10.3.3 Consider the three dimensional euclidean space E^3 with canonical basis and take $W = \mathcal{L}((1, 1, 1))$. Its orthogonal subspace is given by the vectors (x, y, z) whose components solve the linear equation $\Sigma : x + y + z = 0$, so we get $S_\Sigma = W^\perp = \mathcal{L}((1, -1, 0), (1, 0, -1))$. The vectors of the canonical basis when expressed with respect to the vectors $u_1 = (1, 1, 1)$ spanning W and $u_2 = (1, -1, 0),\ u_3 = (1, 0, -1)$ spanning $W^\perp$, are written as

$$
\begin{aligned}
e_1 &= \frac{1}{3}\,(u_1 + u_2 + u_3),\\
e_2 &= \frac{1}{3}\,(u_1 - 2u_2 + u_3),\\
e_3 &= \frac{1}{3}\,(u_1 + u_2 - 2u_3).
\end{aligned}
$$

Therefore,

$$
P_W(e_1) = \frac{1}{3}\,u_1, \qquad P_W(e_2) = \frac{1}{3}\,u_1, \qquad P_W(e_3) = \frac{1}{3}\,u_1,
$$

and

$$
P_{W^\perp}(e_1) = \frac{1}{3}\,(u_2 + u_3), \qquad P_{W^\perp}(e_2) = \frac{1}{3}\,(-2u_2 + u_3), \qquad P_{W^\perp}(e_3) = \frac{1}{3}\,(u_2 - 2u_3).
$$

Remark 10.3.4 Given an orthogonal space decomposition $E^n = W \oplus W^\perp$, the union of the basis $\mathcal{B}_W$ and $\mathcal{B}_{W^\perp}$ of W and $W^\perp$, is a basis $\mathcal{B}$ for W. It is easy to see that the matrix associated to the orthogonal projection operator P_W with respect to such a basis $\mathcal{B}$ has a block diagonal structure

$$
M_{P_W}^{\mathcal{B},\mathcal{B}} = \begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix},
$$

where the order of the diagonal identity block is the dimension of $W = \mathrm{Im}(P_W)$. This makes it evident that an orthogonal projection operator is diagonalisable: its spectrum contains the real eigenvalue $\lambda = 1$ with multiplicity equal to $m_{\lambda=1} = \dim(W)$ and the real eigenvalue $\lambda = 0$ with multiplicity equal to $m_{\lambda=0} = \dim(W^\perp)$.

It is clear that the rank of P_W (the dimension of W) is given by the trace $\mathrm{tr}(M_{P_W}^{\mathcal{B},\mathcal{B}})$ irrespectively of the basis chosen to represent the projection (see the Proposition 9.1.5) since as usual, for a change of basis with matrix $M^{\mathcal{B},\mathcal{C}}$, one has that $M_{P_W}^{\mathcal{C},\mathcal{C}} = M^{\mathcal{C},\mathcal{B}} M_{P_W}^{\mathcal{B},\mathcal{B}} M^{\mathcal{B},\mathcal{C}}$, with $M^{\mathcal{B},\mathcal{C}} = (M^{\mathcal{C},\mathcal{B}})^{-1}$.

Exercise 10.3.5 The matrix

$$M = \begin{pmatrix} a & \sqrt{a-a^2} \\ \sqrt{a-a^2} & 1-a \end{pmatrix}$$

is symmetric and satisfies $M^2 = M$ for any $a \in (0, 1]$. With respect to an orthonormal basis (e_1, e_2) for E^2, it is then associated to an orthogonal projection with rank given by $\mathrm{tr}(M) = 1$. In order to determine its range, we diagonalise M. Its characteristic polynomial is

$$p_M(T) = \det(M - TI_2) = T^2 - T$$

and the eigenvalues are then $\lambda = 0$ and $\lambda = 1$. Since they are both simple, the matrix M is diagonalisable. The eigenspace $V_{\lambda=1}$ corresponding to the range of the orthogonal projection is one dimensional and given as the solution (x, y) of the system

$$\begin{pmatrix} a-1 & \sqrt{a-a^2} \\ \sqrt{a-a^2} & -a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

that is $(x, \sqrt{\frac{1-a}{a}}x)$ with $x \in \mathbb{R}$. This means that the range of the projection is given by $\mathcal{L}((1, \sqrt{\frac{1-a}{a}}))$.

We leave as an exercise to show that M is the most general rank 1 orthogonal projection in E^2.

Exercise 10.3.6 We know from Exercise 10.2.6 that the operator $L = |u\rangle\langle u|$ is self-adjoint. We compute

$$L^2 = |u\rangle\langle u|u\rangle\langle u| = \|u\|^2 L.$$

Thus such an operator L is an orthogonal projection if and only if $\|u\| = 1$. It is then the rank one orthogonal projection $L = |u\rangle\langle u| = P_{\mathcal{L}(u)}$.

Let us assume that W_1 and W_2 are two orthogonal subspaces (to be definite we take $W_2 \subseteq W_1^\perp$). By using for instance the Remark 10.3.4 it is not difficult to show that $P_{W_1} P_{W_2} = P_{W_2} P_{W_1} = 0$. As a consequence,

$$(P_{W_1} + P_{W_2})(P_{W_1} + P_{W_2}) = P_{W_1}^2 + P_{W_2}^2 + P_{W_1} P_{W_2} + P_{W_2} P_{W_1} = (P_{W_1} + P_{W_2}).$$

Since the sum of two self-adjoint endomorphisms is self-adjoint, we can conclude (from Proposition 10.3.2) that the sum $P_{W_1} + P_{W_2}$ is an orthogonal projector, with $P_{W_1} + P_{W_2} = P_{W_1 \oplus W_2}$. This means that with two orthogonal subspaces, the sum of the corresponding orthogonal projectors is the orthogonal projection onto the direct sum of the given subspaces.

These results can be extended. If the euclidean space has a finer orthogonal decomposition, that is there are mutually orthogonal subspaces $\{W_a\}_{a=1,\dots,k}$ with $E^n = W_1 \oplus \cdots \oplus W_k$, then we have a corresponding set of orthogonal projectors P_{W_a}. We omit the proof of the following proposition, which we shall use later on in the chapter.

Proposition 10.3.7 *If $E^n = W_1 \oplus \cdots \oplus W_k$ with mutually orthogonal subspaces W_a, $a = 1, \dots, k$, then the following hold:*

(a) For any $a, b = 1, \dots, k$, one has

$$P_{W_a} P_{W_b} = \delta_{ab} P_{W_a}.$$

(b) If $\widetilde{W} = W_{a_1} \oplus \cdots \oplus W_{a_s}$ is the vector subspace given by the direct sum of the orthogonal subspaces $\{W_{a_j}\}$ with a_j any subset of $(1, \dots, k)$ without repetition, then the sum $\widetilde{P} = P_{W_{a_1}} + \dots + P_{W_{a_s}}$ is the orthogonal projection operator $\widetilde{P} = P_{\widetilde{W}}$.

(c) For any $v \in E^n$, one has

$$v = (P_{W_1} + \cdots + P_{W_k})(v).$$

Notice that point (c) shows that the identity operator acting on E^n can be *decomposed* as the sum of *all* the orthogonal projectors corresponding to *any* orthogonal subspace decomposition of E^n.

Remark 10.3.8 All we have described for the euclidean space E^n can be naturally extended to the hermitian space $(\mathbb{C}^n, \cdot)$ introduced in Sect. 3.4. If for example $(e_1, \dots, e_n)$ gives a hermitian orthonormal basis for H^n, the orthogonal projection onto $W_a = \mathcal{L}(e_a)$ can be written in the Dirac's notation (see the Exercise 10.3.6) as

$$P_{W_a} = |e_a\rangle\langle e_a|,$$

while the orthogonal projection onto $\widetilde{W} = W_{a_1} \oplus \cdots \oplus W_{a_s}$ (point b) of the Proposition 10.3.7) as

$$P_{\widetilde{W}} = |e_{a_1}\rangle\langle e_{a_1}| + \cdots + |e_{a_s}\rangle\langle e_{a_s}|.$$

The decomposition of the identity operator can be now written as

$$\mathrm{id}_{H^n} = |e_1\rangle\langle e_1| + \cdots + |e_n\rangle\langle e_n|.$$

Thus, any vector $v \in H^n$ can be decomposed as

$$v = |v\rangle = |e_1\rangle\langle e_1|v\rangle + \cdots + |e_n\rangle\langle e_n|v\rangle.$$

10.4 The Diagonalization of Self-adjoint Endomorphisms

The following theorem is a central result for the diagonalization of real symmetric matrices.

Theorem 10.4.1 *Let $A \in \mathbb{R}^{n,n}$ be symmetric, ${}^tA = A$. Then, any root of its characteristic polynomial $p_A(T)$ is real.*

Proof Let us assume λ to be a root of $p_A(T)$. Since $p_A(T)$ has real coefficients, its roots are in general complex (see the fundamental theorem of algebra, Theorem A.5.7). We therefore think of A as the matrix associate to an endomorphism

$$\phi : \mathbb{C}^n \longrightarrow \mathbb{C}^n,$$

with $M_{\phi}^{\mathcal{E},\mathcal{E}} = A$ with respect to the canonical basis $\mathcal{E}$ for $\mathbb{C}^n$ as a complex vector space. Let v be a non zero eigenvector for ϕ, that is

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

By denoting with X the column of the components of $v = (x_1, \dots, x_n)$ with respect to $\mathcal{E}$, we write

$${}^tX = {}^t(x_1, \dots, x_n), \qquad AX = \lambda X.$$

Under complex conjugation, with $\bar{A} = A$ since A has real entries, we get

$${}^t\bar{X} = {}^t(\bar{x}_1, \dots, \bar{x}_n), \qquad A\bar{X} = \bar{\lambda}\bar{X}.$$

From these relations we can write the scalar ${}^t\bar{X}AX$ in the following two ways,

$$\begin{aligned} &{}^t\bar{X}AX = {}^t\bar{X}(AX) = {}^t\bar{X}(\lambda X) = \lambda\,({}^t\bar{X}X) \\ \text{and}\quad &{}^t\bar{X}AX = ({}^t\bar{X}A)X = {}^t(A\bar{X})X = {}^t(\bar{\lambda}\bar{X})X = \bar{\lambda}\,({}^t\bar{X}X). \end{aligned}$$

By equating them, we have

$$(\lambda - \bar{\lambda})\,({}^t\bar{X}X) = 0.$$

The quantity ${}^t\bar{X}X = \bar{x}_1x_1 + \bar{x}_2x_2 + \cdots + \bar{x}_nx_n$ is a positive real number, since $v \neq 0_{\mathbb{C}^n}$; we can then conclude $\lambda = \bar{\lambda}$, that is $\lambda \in \mathbb{R}$. □

Example 10.4.2 The aim of this example is threefold, namely

- it provides an *ad hoc* proof of the Theorem 10.4.1 for symmetric 2×2 matrices;
- it provides a direct proof for the Proposition 10.2.2 for symmetric 2×2 matrices;
- it shows that, if ϕ is a self-adjoint endomorphism in E^2, then E^2 has an orthonormal basis made up of eigenvectors for ϕ. This result anticipates the general result which will be proven in the Theorem 10.4.5.

We consider then a symmetric matrix $A \in \mathbb{R}^{2,2}$,

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}.$$

Its characteristic polynomial $p_A(T) = \det(A - T\,I_2)$ is then

$$p_A(T) = T^2 - (a_{11} + a_{22})T + a_{11}a_{22} - a_{12}^2.$$

The discriminant of this degree 2 characteristic polynomial $p_A(T)$ is not negative:

$$\begin{aligned} \Delta &= (a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}^2) \\ &= (a_{11} - a_{22})^2 + 4a_{12}^2 \geq 0 \end{aligned}$$

being the sum of two square terms; therefore the roots λ_1, λ_2 of $p_A(T)$ are both real.

We prove next that A is diagonalisable, and that the matrix P giving the change of basis is orthogonal. We consider the endomorphism ϕ corresponding to $A = M_{\phi}^{\mathcal{E},\mathcal{E}}$, for the canonical basis $\mathcal{E}$ for E^2, and compute the eigenspaces V_{λ_1} and V_{λ_2}.

- If $\Delta = 0$, then $a_{11} = a_{22}$ and $a_{12} = 0$. The matrix A is already diagonal, so we may take $P = I_2$. There is only one eigenvalue $\lambda_1 = a_{11} = a_{22}$. Its algebraic multiplicity is 2 and its geometric multiplicity is 2, with corresponding eigenspace $V_{\lambda_1} = E^2$.
- If $\Delta > 0$ the characteristic polynomial has two simple roots $\lambda_1 \neq \lambda_2$ with corresponding one dimensional orthogonal (from the Proposition 10.2.2) eigenspaces V_{λ_1} and V_{λ_2}. The change of basis matrix P, whose columns are the normalised eigenvectors
$$\frac{v_1}{\|v_1\|} \quad \text{and} \quad \frac{v_2}{\|v_2\|},$$
is then orthogonal by construction. We notice that P can be always chosen to be an element in SO(2), since a permutation of its columns changes the sign of its determinant, and is compatible with the permutation of the eigenvalue λ_1, λ_2 in the diagonal matrix.

In order to explicitly compute the matrix P we see that the eigenspace V_{λ_i} for any $i = 1, 2$ is given by the solutions of the linear homogeneous system associated to the matrix

$$A - \lambda_i I_2 = \begin{pmatrix} a_{11} - \lambda_i & a_{12} \\ a_{12} & a_{22} - \lambda_i \end{pmatrix}.$$

Since we already know that $\dim(V_{\lambda_i}) = 1$, such a linear system is equivalent to a single linear equation. We can write

$$\begin{aligned} V_{\lambda_i} &= \{(x, y) \,:\, (a_{11} - \lambda_i)x + a_{12}y = 0\} \\ &= \mathcal{L}((-a_{12}, a_{11} - \lambda_i)) = \mathcal{L}(v_i), \end{aligned}$$

where we set

$$v_1 = (-a_{12}, a_{11} - \lambda_1), \qquad v_2 = (-a_{12}, a_{11} - \lambda_2).$$

For the scalar product,

$$v_1 \cdot v_2 = a_{12}^2 + a_{11}^2 - (\lambda_1 + \lambda_2)a_{11} + \lambda_1\lambda_2 = 0$$

since one has

$$\lambda_1 + \lambda_2 = a_{11} + a_{22}, \qquad \lambda_1\lambda_2 = a_{11}a_{22} - a_{12}^2.$$

Exercise 10.4.3 We consider again the symmetric matrix

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

from the Exercise 10.2.4. Its characteristic polynomial is

$$p_A(T) = \det(A - TI_2) = p_A(T) = T^2 - 5T + 5,$$

with roots

$$\lambda_\pm = \frac{1}{2}(5 \pm \sqrt{5}).$$

The corresponding eigenspaces $V_\pm$ are the solutions of the homogeneous linear systems associated to the matrices

$$A - \lambda_\pm I_2 = \frac{1}{2}\begin{pmatrix} (-1 \mp \sqrt{5}) & -2 \\ -2 & (1 \pm \sqrt{5}) \end{pmatrix}.$$

one has $\dim(V_\pm) = 1$, so each system is equivalent to a single linear equation, that is

$$V_\pm = \mathcal{L}((-2, 1 \pm \sqrt{5}) = \mathcal{L}(v_\pm),$$

where

$$v_+ = (-2, 1 + \sqrt{5}), \qquad v_- = (-2, 1 - \sqrt{5}),$$

and one computes that

$$v_+ \cdot v_- = 4 - 4 = 0,$$

that is the eigenspaces are orthogonal. The elements

$$u_1 = \frac{v_+}{\|v_+\|} \quad \text{and} \quad u_2 = \frac{v_-}{\|v_-\|}$$

form an orthonormal basis for E^2 of eigenvectors for the endomorphism ϕ_A.

We present now the fundamental result of this chapter, that is the *spectral theorem* for self-adjoint endomorphisms and for symmetric matrices. Towards this, it is worth mentioning that the whole theory, presented in this chapter for the euclidean space E^n, can be naturally formulated for any finite dimensional real vector space equipped with a scalar product (see Chap. 3).

Definition 10.4.4 Let $\phi : V \to V$ be an endomorphism of the real vector space V, and let $\widetilde{V} \subset V$ be a vector subspace in V. If the image of $\widetilde{V}$ for ϕ is a subset of the same $\widetilde{V}$ (that is, $\phi(\widetilde{V}) \subseteq \widetilde{V}$), there is a well defined endomorphism $\phi_{\widetilde{V}} : \widetilde{V} \to \widetilde{V}$ given by

$$\phi_{\widetilde{V}}(v) = \phi(v), \qquad \text{for all} \quad v \in \widetilde{V}$$

(clearly a linear map). The endomorphism $\phi_{\widetilde{V}}$ acts in the same way as the endomorphism ϕ, but on a restricted domain. This is why $\phi_{\widetilde{V}}$ is called the *restriction* to $\widetilde{V}$ of ϕ.

Proposition 10.4.5 (Spectral theorem for endomorphisms) *Let* $(V, \cdot)$ *be a real vector space equipped with a scalar product, and let* $\phi \in \mathrm{End}(V)$. *The endomorphism* ϕ *is self-adjoint if and only if* V *has an orthonormal basis of eigenvectors for* ϕ.

Proof Let us assume the orthonormal basis $\mathcal{C}$ for V is made of eigenvectors for ϕ. This implies that $M_{\phi}^{\mathcal{C},\mathcal{C}}$ is diagonal and therefore symmetric. From the Theorem 10.2.3 we conclude that ϕ is self-adjoint.

The proof of the converse is by induction on $n = \dim(V)$. For $n = 2$ the statement is true, as we explicitly proved in the Example 10.4.2. Let us then assume it to be true for any $(n-1)$-dimensional vector space. Then, let us consider a real n-dimensional vector space $(V, \cdot)$ equipped with a scalar product, and let ϕ be a self-adjoint endomorphism on V. With $\mathcal{B}$ an orthonormal basis for V (remember from the Theorem 3.3.9 that such a basis always exists V finite dimensional), the matrix $A = M_{\phi}^{\mathcal{B},\mathcal{B}}$ is symmetric (from the Theorem 10.2.3) and thus any root of the characteristic polynomial $p_A(T)$ is real. Denote by λ one such an eigenvalue for ϕ, with v_1 a corresponding eigenvector that we can assume of norm 1.

Then, let us consider the orthogonal complement to the vector line spanned by v_1,

$$\widetilde{V} = (\mathcal{L}(v_1))^{\perp}.$$

In order to meaningfully define the restriction to $\widetilde{V}$ of ϕ, we have to verify that for any $v \in \widetilde{V}$ one has $\phi(v) \in \widetilde{V}$, that is, we have to prove the implication

$$v \cdot v_1 = 0 \quad \Rightarrow \quad \phi(v) \cdot v_1 = 0.$$

By recalling that ϕ is self-adjoint and $\phi(v_1) = \lambda v_1$ we can write

$$\begin{aligned}\phi(v) \cdot v_1 &= v \cdot \phi(v_1) = v \cdot (\lambda v_1) \\ &= \lambda\,(v \cdot v_1) = 0.\end{aligned}$$

This proves that ϕ can be restricted to a $\phi_{\widetilde{V}} : \widetilde{V} \to \widetilde{V}$, clearly self-adjoint. Since $\dim(\widetilde{V}) = n - 1$, by the inductive assumption there exist $n - 1$ elements $(v_2, \dots, v_n)$ of eigenvectors for $\phi_{\widetilde{V}}$ making up an orthonormal basis for $\widetilde{V}$. Since $\phi_{\widetilde{V}}$ is a restriction of ϕ, the elements $(v_2, \dots, v_n)$ are eigenvectors for ϕ as well, and orthogonal to v_1 as they all belong to $\widetilde{V}$. Then the elements $(v_1, v_2, \dots, v_n)$ are orthonormal and eigenvectors for ϕ. Being $n = \dim(V)$, they are an orthonormal basis for V. □

10.5 The Diagonalization of Symmetric Matrices

There is a counterpart of Proposition 10.4.5 for symmetric matrices.

Proposition 10.5.1 (Spectral theorem for symmetric matrices) *Let $A \in \mathbb{R}^{n,n}$ be symmetric. There exists an orthogonal matrix P such that ${}^tP A P$ is diagonal. This result is often referred to by saying that* symmetric matrices *are* orthogonally diagonalisable.

Proof Let us consider the endomorphism $\phi = f_A^{\mathcal{E},\mathcal{E}} : E^n \to E^n$, which is self-adjoint since A is symmetric and $\mathcal{E}$ is the canonical basis (see the Theorem 10.2.3). From the Proposition 10.4.5, the space E^n has an orthonormal basis $\mathcal{C}$ of eigenvectors for ϕ, so the matrix $M_\phi^{\mathcal{C},\mathcal{C}}$ is diagonal. From Theorem 7.9.9 we can write

$$M_\phi^{\mathcal{C},\mathcal{C}} = M^{\mathcal{C},\mathcal{E}}\, M_\phi^{\mathcal{E},\mathcal{E}}\, M^{\mathcal{E},\mathcal{C}}.$$

Since $M_\phi^{\mathcal{E},\mathcal{E}} = A$, by setting $P = M^{\mathcal{C},\mathcal{E}}$ we have that $P^{-1}AP$ is diagonal. The columns of the matrix P are given by the components with respect to $\mathcal{E}$ of the elements in $\mathcal{C}$, so P is orthogonal since $\mathcal{C}$ is orthonormal. □

Remark 10.5.2 The orthogonal matrix P can always be chosen in $\mathrm{SO}(n)$, since, as already mentioned, the sign of its determinant changes under a permutation of two columns.

Exercise 10.5.3 Consider $\phi \in \mathrm{End}(\mathbb{R}^4)$ given by

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

Its corresponding matrix with respect to the canonical basis $\mathcal{E}$ in $\mathbb{R}^4$ is given by

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

Being A symmetric and $\mathcal{E}$ orthonormal, than ϕ is self-adjoint. Its characteristic polynomial is

$$\begin{aligned} p_\phi(T) = p_A(T) &= \det(A - T\, I_4) \\ &= \begin{vmatrix} 1-T & 1 \\ 1 & 1-T \end{vmatrix} \begin{vmatrix} -1-T & 1 \\ 1 & -1-T \end{vmatrix} \\ &= T^2(T-2)(T+2). \end{aligned}$$

The eigenvalues are then $\lambda_1 = 0$ with (algebraic) multiplicity $m(0) = 2$, $\lambda_2 = -2$ with $m(-2) = 1$ and $\lambda_2 = 2$ with $m(2) = 1$. The corresponding eigenspaces are computed to be

$$\begin{aligned} V_0 &= \ker(\phi) = \mathcal{L}((1,-1,0,0),(0,0,1,1)), \\ V_{-2} &= \ker(\phi - 2I_4) = \mathcal{L}((1,1,0,0)), \\ V_2 &= \ker(\phi - I_4) = \mathcal{L}((0,0,1,-1)) \end{aligned}$$

and as we expect, these three eigenspaces are mutually orthogonal, with the two basis vectors spanning V_0 orthogonal as well. In order to write the matrix P which diagonalises A one just needs to normalise such a system of four basis eigenvectors. We have

$$P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \end{pmatrix}, \qquad {}^tPAP = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix},$$

where we have chosen an ordering for the eigenvalues that gives $\det(P) = 1$.

Corollary 10.5.4 *Let $\phi \in \mathrm{End}(E^n)$. If the endomorphism ϕ is self-adjoint then it is simple.*

Proof The proof is immediate. From the Proposition 10.4.5 we know that the self-adjointness of ϕ implies that E^n has an orthonormal basis of eigenvectors for ϕ. From the Remark 9.2.3 we conclude that ϕ is simple. □

Exercise 10.5.5 The converse of the previous corollary does not hold in general. Consider for example the endomorphism ϕ in E^2 whose matrix with respect to the canonical basis $\mathcal{E}$ is

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

An easy calculation gives for the eigenvalues $\lambda_1 = 1$ e $\lambda_2 = -1$ and ϕ is (see the Corollary 9.4.2) therefore simple. But ϕ is not self-adjoint, since

$$\begin{aligned} \phi(e_1) \cdot e_2 &= (1, 0) \cdot (0, 1) = 0, \\ e_1 \cdot \phi(e_2) &= (1, 0) \cdot (1, -1) = 1, \end{aligned}$$

or simply because A is not symmetric. The eigenspaces are given by

$$V_1 = \mathcal{L}((1, 0)), \qquad V_{-1} = \mathcal{L}((1, -2)),$$

and they are not orthogonal. As a further remark, notice that the diagonalising matrix

$$P = \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix}$$

is not orthogonal.

What we have shown in the previous exercise is a general property characterising self-adjoint endomorphisms within the class of simple endomorphisms, as the next theorem shows.

Theorem 10.5.6 *Let $\phi \in \mathrm{End}(E^n)$ be simple, with $V_{\lambda_1}, \dots, V_{\lambda_s}$ the corresponding eigenspaces. Then ϕ is self-adjoint if and only if $V_{\lambda_i} \perp V_{\lambda_j}$ for any $i \neq j$.*

Proof That the eigenspaces corresponding to distinct eigenvalues are orthogonal for a self-adjoint endomorphism comes directly from the Proposition 10.2.2.

Conversely, let us assume that ϕ is simple, so that $E^n = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s}$. The union of the bases given by applying the Gram-Schmidt orthogonalisation procedure to an arbitrary basis for each V_{λ_j}, yield an orthonormal basis for E^n, which is clearly made of eigenvectors for ϕ. The statement then follows from the Proposition 10.4.5. □

Exercise 10.5.7 The aim of this exercise is to define (if possible) a self-adjoint endomorphism $\phi : E^3 \to E^3$ such that $\ker(\phi) = \mathcal{L}((1, 2, 1))$ and $\lambda_1 = 1, \ \lambda_2 = 2$ are eigenvalues of ϕ.

Since $\ker(\phi) \neq \{(0, 0, 0)\}$, then $\lambda_3 = 0$ is the third eigenvalue for ϕ, with $\ker(\phi) = V_0$. Thus ϕ is simple since it has three distinct eigenvalues, with $E^3 = V_1 \oplus V_2 \oplus V_0$. In order for ϕ to be self-adjoint, we have to impose that $V_{\lambda_i} \perp V_{\lambda_j}$, for all $i \neq j$. In particular, one has

$$(\ker(\phi))^{\perp} = (V_0)^{\perp} = V_1 \oplus V_2.$$

We compute

$$\begin{aligned}(\ker(\phi))^{\perp} &= (\mathcal{L}((1,2,1))^{\perp} \\ &= \{(\alpha,\beta,-\alpha-2\beta)\ :\ \alpha,\beta\in\mathbb{R}\} \\ &= \mathcal{L}((1,0,-1),(a,b,c))\end{aligned}$$

where we impose that (a,b,c) belongs to $\mathcal{L}((1,2,1))^{\perp}$ and is orthogonal to $(1,0,-1)$. By setting

$$\begin{cases}(1,2,1)\cdot(a,b,c)=0\\ (1,0,-1)\cdot(a,b,c)=0\end{cases},$$

we have $(a,b,c)=(1,-1,1)$, so we select

$$V_1=\mathcal{L}((1,0,-1)),\qquad V_2=\mathcal{L}((1,-1,1)).$$

Having a simple ϕ with mutually orthogonal eigenspaces, the endomorphism ϕ self-adjoint. To get a matrix representing ϕ we can choose the basis in E^3

$$\mathcal{B}\ =\ ((1,0,-1),(1,-1,1),(1,2,1)),$$

thus obtaining

$$M_{\phi}^{\mathcal{B},\mathcal{B}}\ =\ \begin{pmatrix}1&0&0\\0&2&0\\0&0&0\end{pmatrix}.$$

By defining $e_1=(1,0,-1)$, $e_2=(1,-1,1)$ we can write, in the Dirac's notation,

$$\phi=|e_1\rangle\langle e_1|+2|e_2\rangle\langle e_2|.$$

Exercise 10.5.8 This exercise defines a simple, but not self-adjoint, endomorphism $\phi\ :\ E^3\ \to\ E^3$ such that $\ker(\phi)=\mathcal{L}((1,-1,1))$ and $\mathrm{Im}(\phi)=(\ker(\phi))^{\perp}$.

We know that ϕ has the eigenvalue $\lambda_1=0$ with $V_0\ =\ \ker(\phi)$. For ϕ to be simple, the algebraic multiplicity of the eigenvalue λ_1 must be 1, and there have to be two additional eigenvalues λ_2 and λ_3 with either $\lambda_2=\lambda_3$ or $\lambda_2\neq\lambda_3$. If $\lambda_2=\lambda_3$, one has then

$$V_{\lambda_2}=\mathrm{Im}(f)=(\ker(f))^{\perp}=(V_{\lambda_1})^{\perp}.$$

In such a case, ϕ would be a simple endomorphism with mutually orthogonal eigenspaces for distinct eigenvalues. This would imply ϕ to be self-adjoint. Thus to satisfy the conditions we require for ϕ we need $\lambda_2\neq\lambda_3$. In such a case, one has $V_{\lambda_2}\oplus V_{\lambda_3}=\mathrm{Im}(\phi)$ and also clearly $V_{\lambda_i}\perp V_0$ for $i=2,3$. In order for ϕ to be simple but not self-adjoint, we select the eigenspaces V_{λ_2} and V_{λ_3} to be not mutually orthogonal subspaces in $\mathrm{Im}(f)$. Since

$$\begin{aligned}\mathrm{Im}(f) &= (\mathcal{L}((1,-1,1)))^{\perp} \\ &= \{(x,y,z)\ :\ x-y+z=0\} \\ &= \mathcal{L}((1,1,0),(0,1,1))\end{aligned}$$

we can choose

$$V_{\lambda_2} = \mathcal{L}((1,1,0)), \qquad V_{\lambda_3} = \mathcal{L}((0,1,1)).$$

If we set $\mathcal{B} = \big((1,-1,1),(1,1,0),(0,1,1)\big)$ (clearly not an orthonormal basis for E^3), we have

$$M_{\phi}^{\mathcal{B},\mathcal{B}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

Exercise 10.5.9 Consider the endomorphism $\phi\ :\ E^3 \to E^3$ whose corresponding matrix with respect to the basis $\mathcal{B} = \big(v_1 = (1,1,0), v_2 = (1,-1,0), v_3 = (0,0,-1)\big)$ is

$$M_{\phi}^{\mathcal{B},\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

With $\mathcal{E}$ as usual the canonical basis for E^3, in this exercise we would like to determine:

(1) an orthonormal basis $\mathcal{C}$ for E^3 given by eigenvectors for ϕ,
(2) the orthogonal matrix $M^{\mathcal{E},\mathcal{C}}$,
(3) the matrix $M^{\mathcal{C},\mathcal{E}}$,
(4) the matrix $M_{\phi}^{\mathcal{E},\mathcal{E}}$,
(5) the eigenvalues of ϕ with their corresponding multiplicities.

(1) We start by noticing that, since $M_{\phi}^{\mathcal{B},\mathcal{B}}$ is diagonal, the basis $\mathcal{B}$ is given by eigenvectors of ϕ, as the action of ϕ on the basis vectors in $\mathcal{B}$ can be clearly written as $\phi(v_1) = v_1$, $\phi(v_2) = 2v_2$, $\phi(v_3) = 3v_3$. The basis $\mathcal{B}$ is indeed orthogonal, but not orthonormal, and for an orthonormal basis $\mathcal{C}$ of eigenvectors for ϕ we just need to normalize, that is to consider

$$u_1 = \frac{v_1}{\|v_1\|}, \qquad u_2 = \frac{v_2}{\|v_2\|}, \qquad u_3 = \frac{v_3}{\|v_3\|}$$

just obtaining $\mathcal{C} = (\frac{1}{\sqrt{2}}(1,1,0), \frac{1}{\sqrt{2}}(1,-1,0), (0,0,-1))$. While the existence of such a basis $\mathcal{C}$ implies that ϕ is self-adjoint, the self-adjointness of ϕ could not be derived from the matrix $M_{\phi}^{\mathcal{B},\mathcal{B}}$, which is symmetric with respect to a basis $\mathcal{B}$ which is not orthonormal.

(2) From its definition, the columns of $M^{\mathcal{E},\mathcal{C}}$ are given by the components with respect to $\mathcal{E}$ of the vectors in $\mathcal{C}$. We then have

$$M^{\mathcal{E},\mathcal{C}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -\sqrt{2} \end{pmatrix}.$$

(3) We know that $M^{\mathcal{C},\mathcal{E}} = \left(M^{\mathcal{E},\mathcal{C}}\right)^{-1}$. Since the matrix above is orthogonal, we have

$$M^{\mathcal{C},\mathcal{E}} = {}^{t}\left(M^{\mathcal{E},\mathcal{C}}\right) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -\sqrt{2} \end{pmatrix}.$$

(4) From the Theorem 7.9.9 we have

$$M_{\phi}^{\mathcal{E},\mathcal{E}} = M^{\mathcal{E},\mathcal{C}}\, M_{\phi}^{\mathcal{C},\mathcal{C}}\, M^{\mathcal{C},\mathcal{E}}.$$

Since $M_{\phi}^{\mathcal{C},\mathcal{C}} = M_{\phi}^{\mathcal{B},\mathcal{B}}$, the matrix $M_{\phi}^{\mathcal{E},\mathcal{E}}$ can be now directly computed.

(5) Clearly, from $M_{\phi}^{\mathcal{B},\mathcal{B}}$ the eigenvalues for ϕ are all simple and given by $\lambda = 1, 2, 3$.