Chapter 12
Spectral Theorems on Hermitian Spaces

In this chapter we shall extend to the complex case some of the notions and results of Chap. 10 on euclidean spaces, with emphasis on spectral theorems for a natural class of endomorphisms.

12.1 The Adjoint Endomorphism

Consider the vector space $\mathbb{C}^n$ and its dual space $\mathbb{C}^{n*}$, as defined in Sect. 8.1. The duality between $\mathbb{C}^n$ and $\mathbb{C}^{n*}$ allows one to define, for any endomorphism ϕ of $\mathbb{C}^n$, its adjoint.

Definition 12.1.1 Given $\phi : \mathbb{C}^n \to \mathbb{C}^n$, the map $\phi^\dagger : \omega \in \mathbb{C}^{n*} \mapsto \phi^\dagger(\omega) \in \mathbb{C}^{n*}$ defined by

$$(\phi^\dagger(\omega))(v) = \omega(\phi(v)) \tag{12.1}$$

for any $\omega \in \mathbb{C}^{n*}$ and any $v \in \mathbb{C}^n$ is called the *adjoint* to ϕ.

Remark 12.1.2 From the linearity of ϕ and ω it follows that $\phi^\dagger$ is linear, so $\phi^\dagger \in \mathrm{End}(\mathbb{C}^{n*})$.

Example 12.1.3 Let $\mathcal{B} = (b_1, b_2)$ be a basis for $\mathbb{C}^2$, with $\mathcal{B}^* = (\beta_1, \beta_2)$ its dual basis for $\mathbb{C}^{2*}$. If ϕ is the endomorphism given by

$$\begin{aligned} \phi : b_1 &\mapsto kb_1 + b_2 \\ \phi : b_2 &\mapsto b_2, \end{aligned}$$

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_12

with $k \in \mathbb{C}$, we see from the definition of adjoint that

$$\begin{aligned} \phi^\dagger(\beta_1) \,:\, b_1 &\mapsto \beta_1(\phi(b_1)) = k \\ \phi^\dagger(\beta_1) \,:\, b_2 &\mapsto \beta_1(\phi(b_2)) = 0 \\ \phi^\dagger(\beta_2) \,:\, b_1 &\mapsto \beta_2(\phi(b_1)) = 1 \\ \phi^\dagger(\beta_2) \,:\, b_2 &\mapsto \beta_2(\phi(b_2)) = 1. \end{aligned}$$

The (linear) action of the adjoint map $\phi^\dagger$ to ϕ is then

$$\begin{aligned} \phi^\dagger \,:\, \beta_1 &\mapsto k\beta_1 \\ \phi^\dagger \,:\, \beta_2 &\mapsto \beta_1 + \beta_2. \end{aligned}$$

Consider now the canonical hermitian space $H^n = (\mathbb{C}^n, \cdot)$, that is the vector space $\mathbb{C}^n$ with the canonical hermitian product (see Sect. 3.4). As described in Sect. 8.2, the hermitian product allows one to identify $\mathbb{C}^{n*}$ with $\mathbb{C}^n$. Under such identification, the defining relation for $\phi^\dagger$ can be written as

$$(\phi^\dagger u) \cdot v = u \cdot (\phi\, v) \qquad \text{or equivalently} \qquad \langle \phi^\dagger(u) \mid v \rangle = \langle u | \phi(v) \rangle$$

for any $u, v \in \mathbb{C}^n$, so that $\phi^\dagger$ is an endomorphism of $H^n = (\mathbb{C}^n, \cdot)$.

Definition 12.1.4 Given a matrix $A = (a_{ij}) \in \mathbb{C}^{n,n}$, its adjoint $A^\dagger \in \mathbb{C}^{n,n}$ is the matrix whose entries are given by $(A^\dagger)_{ab} = \overline{a_{ba}}$.

Thus, adjoining a matrix is the composition of two compatible involutions, the transposition and the complex conjugation.

Exercise 12.1.5 Clearly

$$A = \begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix} \quad \Rightarrow \quad A^\dagger = \begin{pmatrix} 1 & 0 \\ \bar{\alpha} & \bar{\beta} \end{pmatrix}.$$

Exercise 12.1.6 By using the matrix calculus we described in the previous chapters, it comes as no surprise that the following relations hold.

$$\begin{aligned} &(A^\dagger)^\dagger = A, \\ &(AB)^\dagger = B^\dagger A^\dagger, \\ &(A + \alpha B)^\dagger = (A^\dagger + \bar{\alpha} B^\dagger) \end{aligned}$$

for any $A, B \in \mathbb{C}^{n,n}$ and $\alpha \in \mathbb{C}$. The second line indeed parallels the Remark 8.2.1. If we have two endomorphisms $\phi, \psi \in \mathrm{End}(H^n)$, one has

$$\langle (\phi\psi)^\dagger(u) | v \rangle = \langle u | \phi\,\psi(v) \rangle = \langle \phi^\dagger(u) | \psi(v) \rangle = \langle \psi^\dagger \phi^\dagger(u) | v \rangle,$$

for any $u, v \in H^n$. With $\alpha \in \mathbb{C}$, it is also

$$\langle(\phi + \alpha\,\psi)^{\dagger} u|v\rangle \;=\; \langle u|(\phi + \alpha\,\psi)v\rangle \;=\; \langle\phi(u)|v\rangle \;+\; \alpha\langle\psi(u)|v\rangle \;=\; \langle(\phi^{\dagger} + \bar{\alpha}\psi^{\dagger})(u)|v\rangle.$$

Again using the properties of the hermitian product together with the definition of adjoint, it is

$$\langle(\phi^{\dagger})^{\dagger} u|v\rangle \;=\; \langle u|\phi^{\dagger}(v)\rangle \;=\; \langle\phi(u)|v\rangle$$

The above lines establish the following identities

$$\begin{aligned}(\phi^{\dagger})^{\dagger} &= \phi,\\ (\phi\,\psi)^{\dagger} &= \psi^{\dagger}\phi^{\dagger},\\ (\phi + \alpha\,\psi)^{\dagger} &= \phi^{\dagger} + \bar{\alpha}\psi^{\dagger}\end{aligned}$$

which are the operator counterpart of the matrix identities described above.

Definition 12.1.7 An endomorphism ϕ on H^n is called

(a) *self-adjoint*, or *hermitian*, if
$$\phi = \phi^{\dagger},$$
that is if $\langle\phi(u)|v\rangle = \langle u|\phi(v)\rangle$ for any $u, v \in H^n$,

(b) *unitary*, if
$$\phi\phi^{\dagger} = \phi^{\dagger}\phi = I_n,$$
that is if $\langle\phi(u)|\phi(v)\rangle = \langle u|v\rangle$ for any $u, v \in H^n$,

(c) *normal*, if $\phi\phi^{\dagger} = \phi^{\dagger}\phi$.

In parallel to these, a matrix $A \in \mathbb{C}^{n,n}$ is called

(a) *self-adjoint*, or *hermitian*, if $A^{\dagger} = A$,
(b) *unitary*, if $AA^{\dagger} = A^{\dagger}A = I_n$,
(c) *normal*, if $AA^{\dagger} = A^{\dagger}A$.

Remark 12.1.8 Clearly the condition of unitarity for ϕ is equivalent to the condition $\phi^{\dagger} = \phi^{-1}$. Also, both unitary and self-adjoint endomorphisms are normal. From the Remark 12.1.6 it follows that for any endomorphism ψ, the compositions $\psi\psi^{\dagger}$ and $\psi^{\dagger}\psi$ are self-adjoint.

Remark 12.1.9 The notion of adjoint of an endomorphism can be introduced also on euclidean spaces E^n, where it is identified, at a matrix level, by the transposition. Then, it is clear that the notion of self-adjointness in H^n generalises that in E^n, since if $A = {}^tA$ in E^n, then $A = A^{\dagger}$ in H^n, while orthogonal matrices in E^n are unitary matrices in H^n with real entries.

The following theorem is the natural generalisation for hermitian spaces of a similar result for euclidean spaces. Its proof, that we omit, mimics indeed that of the Theorem 10.1.11.

Theorem 12.1.10 *Let $\mathcal{C}$ be an orthonormal basis for the hermitian vector space H^n and let $\mathcal{B}$ be any other basis. The matrix $M^{\mathcal{C},\mathcal{B}}$ of the change of basis from $\mathcal{C}$ to $\mathcal{B}$ is unitary if and only if $\mathcal{B}$ is orthonormal.*

The following proposition, gives an *ex-post* motivation for the definitions above.

Proposition 12.1.11 *If $\mathcal{E}$ is the canonical basis for H^n, with $\phi \in \mathrm{End}(H^n)$, it holds that*

$$M^{\mathcal{E},\mathcal{E}}_{\phi^\dagger} = (M^{\mathcal{E},\mathcal{E}}_{\phi})^\dagger$$

Proof Let $\mathcal{E} = (e_1, \dots, e_n)$ be the canonical basis for H^n. If $M^{\mathcal{E},\mathcal{E}}_{\phi} \in \mathbb{C}^{n,n}$ is the matrix that represents the action of ϕ on H^n with respect to the basis $\mathcal{E}$, its entries are given (see 8.7) by

$$(M^{\mathcal{E},\mathcal{E}}_{\phi})_{ab} = \langle e_a | \phi(e_b) \rangle.$$

By denoting $\phi_{ab} = (M^{\mathcal{E},\mathcal{E}}_{\phi})_{ab}$, the action of ϕ is given by $\phi(e_a) = \sum_{b=1}^{n} \phi_{ba} e_b$, so we can compute

$$(M^{\mathcal{E},\mathcal{E}}_{\phi^\dagger})_{ab} = \langle e_a | \phi^\dagger(e_b) \rangle = \langle \phi(e_a) | e_b \rangle = \sum_{c=1}^{n} \langle \phi_{ca} e_c | e_b \rangle = \overline{\phi_{ba}}$$

As an application of this proposition, the next proposition also generalises to hermitian spaces analogous results proven in Chap. 10 for euclidean spaces.

Proposition 12.1.12 *The endomorphism ϕ on H^n is self-adjoint (resp. unitary, resp. normal) if and only if there exists an orthonormal basis $\mathcal{B}$ for H^n with respect to which the matrix $M^{\mathcal{B},\mathcal{B}}_{\phi}$ is self-adjoint (resp. unitary, resp. normal).*

Exercise 12.1.13 Consider upper triangular matrices in $\mathbb{C}^{2,2}$,

$$M = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \quad \Rightarrow \quad M^\dagger = \begin{pmatrix} \bar{a} & 0 \\ \bar{b} & \bar{c} \end{pmatrix}.$$

One explicitly computes

$$MM^\dagger = \begin{pmatrix} a\bar{a} + b\bar{b} & b\bar{c} \\ c\bar{b} & b\bar{b} + c\bar{c} \end{pmatrix}, \qquad M^\dagger M = \begin{pmatrix} a\bar{a} & b\bar{a} \\ a\bar{b} & b\bar{b} + c\bar{c} \end{pmatrix},$$

and the matrix M is normal, $MM^\dagger = M^\dagger M$, if and only if $b\bar{b} = 0 \Leftrightarrow b = 0$. Thus an upper triangular matrix in 2-dimension is normal if and only if it is diagonal. In such a case, the matrix is self-adjoint if the diagonal entries are real, and unitary if the diagonal entries have norm 1.

Exercise 12.1.14 We consider the following family of matrices in $\mathbb{C}^{2,2}$,

$$M = \begin{pmatrix} a & b \\ c & 0 \end{pmatrix}, \qquad M^{\dagger} = \begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & 0 \end{pmatrix}.$$

It is

$$MM^{\dagger} = \begin{pmatrix} a\bar{a} + b\bar{b} & a\bar{c} \\ c\bar{a} & +c\bar{c} \end{pmatrix}, \qquad M^{\dagger}M = \begin{pmatrix} a\bar{a} + c\bar{c} & b\bar{a} \\ a\bar{b} & +b\bar{b} \end{pmatrix}.$$

The conditions for which M is normal are

$$b\bar{b} = c\bar{c}, \qquad a\bar{c} = b\bar{a}.$$

These are solved by $b = Re^{\mathrm{i}\beta},\ c = Re^{\mathrm{i}\gamma},\ A = |A|e^{\mathrm{i}\alpha}$ with $2\alpha = (\beta + \gamma) \bmod 2\pi$, where $R > 0$ and $|A| > 0$ are arbitrary moduli for complex numbers.

Exercise 12.1.15 With the Dirac's notation as in (8.6), an endomorphism ϕ and its adjoint are written as

$$\phi = \sum_{a,b=1}^{n} \phi_{ab}\ |e_a\rangle\langle e_b| \quad \text{ and } \quad \phi^{\dagger} = \sum_{a,b=1}^{n} \overline{\phi_{ba}}\ |e_a\rangle\langle e_b|$$

with $\phi_{ab} = \langle e_a|\phi(e_b)\rangle = (M_{\phi}^{\mathcal{E},\mathcal{E}})_{ab}$ with respect to the orthonormal basis $\mathcal{E} = (e_1, \ldots, e_n)$.

With $u = (u_1, \cdots, u_n)$ and $v = (v_1, \ldots, v_n)$ vectors in H^n we have the endomorphism $P = |u\rangle\langle v|$. If we decompose the identity endomorphism (see the point (c) from the Proposition 10.3.7) as

$$\mathrm{id} = \sum_{s=1}^{n} |e_s\rangle\langle e_s|$$

we can write

$$P = |u\rangle\langle v| = \sum_{ab=1}^{n} |e_a\rangle\langle e_a|u\rangle\langle v|e_b\rangle\langle e_b| = \sum_{ab=1}^{n} P_{ab}\ |e_a\rangle\langle e_b|$$

with $P_{ab} = u_a\overline{v_b} = \langle e_a|P(e_b)\rangle$. Clearly then

$$P^{\dagger} = |v\rangle\langle u|.$$

Example 12.1.16 Let ϕ an endomorphism H^n with matrix $M_\phi^{\mathcal{E},\mathcal{E}}$ with respect to the canonical orthonormal basis, thus $(M_\phi^{\mathcal{E},\mathcal{E}})_{ab} = \langle e_a|\phi(e_b)\rangle$. If $\mathcal{B} = (b_1, \ldots, b_n)$ is a second orthonormal basis for H^n, we have two decompositions

$$\mathrm{id} \;=\; \sum_{k=1}^{n} |\,e_k\rangle\langle e_k| \;=\; \sum_{s=1}^{n} |\,e_s\rangle\langle e_s|.$$

Thus, by *inserting* these two expressions of the identity operators, we have

$$\langle e_a|\phi(e_b)\rangle \;=\; \sum_{k,s=1}^{n} \langle e_a|b_k\rangle\langle b_k|\phi(b_s)\rangle\langle b_s|e_b\rangle,$$

giving in components,

$$(M_\phi^{\mathcal{E},\mathcal{E}})_{ab} \;=\; \sum_{k,s=1}^{n} \langle e_a|b_k\rangle(M_\phi^{\mathcal{B},\mathcal{B}})_{ks}\langle b_s|e_b\rangle.$$

The matrix of the change of basis from $\mathcal{E}$ to $\mathcal{B}$ has entries $\langle e_a|b_k\rangle = (M^{\mathcal{E},\mathcal{B}})_{ak}$, with its inverse matrix entries given by $(M^{\mathcal{B},\mathcal{E}})_{sb} = \langle b_s|e_b\rangle$. From the previous examples we see that

$$(M^{\mathcal{B},\mathcal{E}\dagger})_{ak} = \overline{(M^{\mathcal{B},\mathcal{E}})_{ka}} \;=\; \overline{\langle b_k|e_a\rangle} \;=\; \langle e_a|b_k\rangle \;=\; (M^{\mathcal{E},\mathcal{B}})_{ak}$$

thus finding that the change of basis is given by a unitary matrix.

Proposition 12.1.17 *For any endomorphism ϕ in H^n, there is an orthogonal vector space decomposition*

$$H^n \;=\; \mathrm{Im}(\phi) \,\oplus\, \ker(\phi^\dagger)$$

Proof If u is any vector in H^n, the vector $\phi(u)$ cover over all of $\mathrm{Im}(\phi)$, so the condition $\langle\phi(u)|w\rangle = 0$ characterises the elements $w \in (\mathrm{Im}(\phi))^\perp$. It is now easy to compute

$$0 \;=\; \langle\phi(u)|w\rangle \;=\; \langle u|\phi^\dagger(w)\rangle.$$

Since u is arbitrary and the hermitian product is not degenerate, we have $\ker(\phi^\dagger) = (\mathrm{Im}(\phi))^\perp$. □

12.2 Spectral Theory for Normal Endomorphisms

We prove a few results for normal endomorphisms which will be useful for spectral theorems.

Proposition 12.2.1 *Let ϕ be a normal endomorphism of H^n.*

(a) With $u \in H^n$, we can write

$$\|\phi(u)\|^2 = \langle\phi(u)|\phi(u)\rangle = \langle u|\phi^\dagger\phi(u)\rangle = \langle u|\phi\phi^\dagger(u)\rangle = \langle\phi^\dagger(u)|\phi^\dagger(u)\rangle = \|\phi^\dagger(u)\|^2.$$

Since the order of these computations can be reversed, we have the following characterisation.

$$\phi\phi^\dagger = \phi^\dagger\phi \quad \Leftrightarrow \quad \|\phi(u)\| = \|\phi^\dagger(u)\| \quad \textit{for all} \quad u \in H^n.$$

(b) From this it also follows that $\ker(\phi) = \ker(\phi^\dagger)$. *So from the Proposition 12.1.17, we have the following orthogonal decomposition,*

$$H^n = \mathrm{Im}(\phi) \oplus \ker(\phi).$$

(c) Clearly $(\phi - \lambda I)$ is a normal endomorphism if ϕ is such. This gives $\ker(\phi - \lambda I) = \ker(\phi^\dagger - \bar{\lambda} I)$, *meaning that if λ is an eigenvalue of a normal endomorphism ϕ, then $\bar{\lambda}$ is an eigenvalue for $\phi^\dagger$, with the same eigenspaces.*

(d) Let λ, μ be two distinct eigenvalues for ϕ, with $\phi(v) = \lambda v$ and $\phi(w) = \mu w$. Then we have

$$(\lambda - \mu)\langle v|w\rangle = \langle\bar{\lambda}v|w\rangle - \langle v|\mu w\rangle = \langle\phi^\dagger(v)|w\rangle - \langle v|\phi(w)\rangle = 0.$$

We can conclude that the eigenspaces corresponding to distinct eigenvalues for a normal endomorphism are mutually orthogonal. □

We are ready to characterise a normal operator in terms of its spectral properties. The proof of the following result generalises to hermitian spaces the proof of the Theorem 10.4.5 on the diagonalization of symmetric endomorphisms on euclidean spaces.

Theorem 12.2.2 *An endomorphism ϕ of H^n is normal if and only there exists an orthonormal basis for H^n made of eigenvectors for ϕ.*

Proof If $\mathcal{B} = (b_1, \ldots, b_n)$ is an orthonormal basis of eigenvectors for ϕ, with corresponding eigenvalues $(\lambda_1, \ldots, \lambda_n)$, we can write

$$\phi = \sum_{a=1}^{n} \lambda_a \, |b_a\rangle\langle b_a| \qquad \text{and} \qquad \phi^\dagger = \sum_{a=1}^{n} \overline{\lambda_a} \, |b_a\rangle\langle b_a|$$

which directly yields (see the Exercise 12.1.15)

$$\phi\phi^{\dagger} \;=\; \sum_{a=1}^{n} |\lambda_a|^2 \; |b_a\rangle\langle b_a| \;=\; \phi^{\dagger}\phi.$$

The converse, the less trivial part of the statement, is proven once again by induction.

Consider first a normal operator ϕ on the two dimensional hermitian space H^2. With respect to any basis, the characteristic polynomial $p_{\phi}(T)$ has two complex roots, from the fundamental theorem of algebra. A normal endomorphism of H^2 with only the zero eigenvalue, would be the null endomorphism. So we can assume there is a root $\lambda \neq 0$, with v a (normalised) eigenvectors, that is $\phi(v) = \lambda v$ with $\|v\| = 1$. If $\mathcal{C} = (v, w)$ is an orthonormal basis for H^2 that completes v, we have, from point (c) above,

$$\langle \phi(w)|v\rangle \;=\; \langle w|\phi^{\dagger}(v)\rangle \;=\; \langle w|v\rangle\bar{\lambda} \;=\; 0.$$

Being $\lambda \neq 0$, this shows that $\phi(w)$ is orthogonal to $\mathcal{L}(v)$, so that there must exists a scalar μ, such that $\phi(w) = \mu w$. In turn this shows that if ϕ is a normal endomorphism of H^2, then H^2 has an orthonormal basis of eigenvectors for ϕ.

Inductively, let us assume that the statement is valid when the dimension of the hermitian space is $n-1$. The n-dimensional case is treated analogously to what done above. If ϕ is a normal endomorphism of H^n, its characteristic polynomial $p_{\phi}(T)$ has at least a non zero complex root, λ say, with v a corresponding normalised eigenvector: $\phi(v) = \lambda v$, with $\|v\| = 1$. (Again, a normal endomorphism of H^n with only the zero eigenvalue is the null endomorphism.) We have $H^n = V_{\lambda} \oplus V_{\lambda}^{\perp}$ and v can be completed to an orthonormal basis $\mathcal{C} = (v, w_1, \ldots, w_n)$ for H^n. If $w \in V_{\lambda}^{\perp}$ we compute as above

$$\langle \phi(w)|v\rangle \;=\; \langle w|\phi^{\dagger}(v)\rangle \;=\; \langle w|v\rangle\bar{\lambda} \;=\; 0.$$

This shows that ϕ maps $V_{\lambda}^{\perp}$ to itself, while also $\phi^{\dagger}$ maps $V_{\lambda}^{\perp}$ to itself since,

$$\langle \phi^{\dagger}(w)|v\rangle \;=\; \langle w|\phi(v)\rangle \;=\; \langle w|v\rangle\lambda \;=\; 0.$$

The restriction of ϕ to $V_{\lambda}^{\perp}$ is then a normal operator on a $(n-1)$ dimensional hermitian space, and by assumption there exists an orthonormal basis $(u_1, \ldots, u_{n-1})$ for $V_{\lambda}^{\perp}$ made of eigenvectors for ϕ. The basis $\mathcal{E} = (v, u_1, \ldots, u_{n-1})$ is an orthonormal basis for H^n of eigenvectors for ϕ. □

Remark 12.2.3 Since the field of real numbers is not algebraically closed (and the fundamental theorem of algebra is valid on $\mathbb{C}$), it is worth stressing that an analogue of this theorem for normal endomorphisms on euclidean spaces does not hold. A matrix $A \in \mathbb{R}^{n,n}$ such that $({}^{t}A)\,A = A\,({}^{t}A)$, needs not be diagonalisable. An example is given by an antisymmetric (skew-adjoint, see Sect. 11.1) matrix A, which clearly commutes with ${}^tA$, being nonetheless not diagonalisable.

We showed in the Remark 12.1.8 that self-adjoint and unitary endomorphisms are normal. Within the set of normal endomorphisms, they can be characterised in terms of their spectrum.

If λ is an eigenvalue of a self-adjoint endomorphism ϕ, with $\phi(v) = \lambda v$, then

$$\lambda v \;=\; \phi(v) \;=\; \phi^{\dagger}(v) = \bar{\lambda} v$$

and thus one has $\lambda = \bar{\lambda}$. If λ is an eigenvalue for a unitary operator ϕ, with $\phi(v) = \lambda v$, then

$$\|v\|^2 \;=\; \|\phi(v)\|^2 \;=\; |\lambda|^2 \|v\|^2,$$

which gives $|\lambda| = 1$. It is easy to show also the converse of these claims, so to have the following.

Theorem 12.2.4 *A normal operator on H^n is self-adjoint if and only if its eigenvalues are real. A normal operator on H^n is unitary if and only if its eigenvalues have modulus 1.*

As a corollary, by merging the previous two theorems, we have a characterisation of self-adjoint and unitary operators in terms of their spectral properties, as follows.

Corollary 12.2.5 *An endomorphism ϕ on H^n is self-adjoint if and only if its spectrum is real and there exists an orthonormal basis for H^n of eigenvectors for ϕ. An endomorphism ϕ on H^n is unitary if and only if its spectrum is a subset of the unit circle in $\mathbb{C}$, and there exists an orthonormal basis for H^n of eigenvectors for ϕ.*

Exercise 12.2.6 Consider the hermitian space H^2, with $\mathcal{E} = (e_1, e_2)$ its canonical orthonormal basis, and the endomorphism ϕ represented with respect to $\mathcal{E}$ by

$$M_{\phi}^{\mathcal{E},\mathcal{E}} \;=\; \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} \qquad \text{with } a \in \mathbb{R}.$$

This endomorphism is not diagonalisable over $\mathbb{R}$, since it is antisymmetric (see Sect. 11.1) and the Remark 12.2.3. Being normal with respect to the hermitian structure in H^2, there exists an orthonormal basis for H^2 of eigenvectors for ϕ. The eigenvalue equation is $p_{\phi}(T) = T^2 + a^2 = 0$, so the eigenvalues are $\lambda_{\pm} = \pm \mathrm{i}a$, with normalised eigenvectors $u_{\pm}$ given by

$$\lambda_{\pm} \;=\; \pm \mathrm{i}\, a \qquad u_{\pm} \;=\; \tfrac{1}{\sqrt{2}}(1, \pm \mathrm{i})_{\mathcal{E}},$$

while the unitary conjugation that diagonalises the matrix $M_{\phi}^{\mathcal{E},\mathcal{E}}$ is given by

$$\tfrac{1}{2}\begin{pmatrix} 1 & -\mathrm{i} \\ 1 & \mathrm{i} \end{pmatrix} \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{pmatrix} = \begin{pmatrix} \mathrm{i}\, a & 0 \\ 0 & -\mathrm{i}\, a \end{pmatrix}.$$

The comparison of this with the content of the Example 12.1.16 follows by writing the matrix giving the change of basis from $\mathcal{E}$ to $\mathcal{B} = (u_+, u_-)$ as

$$M^{\mathcal{B},\mathcal{E}} = \tfrac{1}{2}\begin{pmatrix} 1 & -\mathrm{i} \\ 1 & \mathrm{i} \end{pmatrix} = \begin{pmatrix} \langle u_+|e_1\rangle & \langle u_+|e_2\rangle \\ \langle u_-|e_1\rangle & \langle u_-|e_2\rangle \end{pmatrix}.$$

We next study a family of normal endomorphisms, which will be useful when considering the properties of unitary matrices. The following definition comes naturally from the Definition 12.1.7.

Definition 12.2.7 An endomorphism ϕ in H^n is named *skew-adjoint* if $\langle u|\phi(v)\rangle + \langle \phi^\dagger(u)|v\rangle = 0$ for any $u, v \in H^n$. A matrix $A \in \mathbb{C}^{n,n}$ is named *skew-adjoint* if $A^\dagger = -A$.

We list some important results on skew-adjoint endomorphisms and matrices.

(a) It is clear that an endomorphism ϕ on H^n is skew-adjoint if and only if there exists an orthonormal basis $\mathcal{E}$ for H^n with respect to which the matrix $M_\phi^{\mathcal{E},\mathcal{E}}$ is skew-adjoint.
(b) Skew-adjoint endomorphisms are normal. We know from the Proposition 12.2.1 point (c), that if λ is an eigenvalue for the endomorphism ϕ, then $\bar{\lambda}$ is an eigenvalue for $\phi^\dagger$. This means that if λ is an eigenvalue For a skew-adjoint endomorphism ϕ, then $\bar{\lambda} = -\lambda$, so any eigenvalue for a skew-adjoint endomorphism is either purely imaginary or zero.
(c) There exists an orthonormal basis $\mathcal{E} = (e_1, \dots, e_n)$ of eigenvectors for ϕ such that

$$\phi = \sum_{a=1}^{n} \mathrm{i}\,\lambda_a\,|e_a\rangle\langle e_a| \qquad \text{with} \quad \lambda_a \in \mathbb{R}.$$

(d) The real vector space of skew-adjoint matrices $A = -A^\dagger \in \mathbb{C}^{n,n}$ is a matrix Lie algebra (see the Definition 11.1.6), that is the commutator of skew-adjoint matrices is a skew-adjoint matrix; it is denoted $\mathfrak{u}(n)$ and it has dimension n.

Remark 12.2.8 In parallel with the Remark 11.1.9, self-adjoint matrices do not make up a Lie algebra since the commutator of two self-adjoint matrices is a skew-adjoint matrix.

Exercise 12.2.9 On the hermitian space H^3 we consider the endomorphism ϕ whose representing matrix is, with respect to the canonical basis $\mathcal{E}$, given by

$$M_\phi^{\mathcal{E},\mathcal{E}} = \begin{pmatrix} 0 & \mathrm{i} & a \\ \mathrm{i} & 0 & 0 \\ -a & 0 & 0 \end{pmatrix},$$

with a a real parameter. Since $(M_{\phi}^{\mathcal{E},\mathcal{E}})^{\dagger} = -M_{\phi}^{\mathcal{E},\mathcal{E}}$, then ϕ is skew-adjoint (and thus normal). Its characteristic equation

$$p_{\phi}(T) = -T(1 + a^2 + T^2) = 0$$

has solutions $\lambda = 0$ and $\lambda_{\pm} = \pm i\sqrt{1 + a^2}$. Explicit calculations show that the eigenspaces are given by $\ker(\phi) = V_{\lambda=0} = \mathcal{L}(u_0)$ and $V_{\lambda_{\pm}} = \mathcal{L}(u_{\pm})$ with

$$\begin{aligned} u_0 &= \tfrac{1}{\sqrt{1+a^2}}\,(0, ia, 1), \\ u_{\pm} &= \tfrac{1}{\sqrt{2(1+a^2)}}\,(\sqrt{1 + a^2}, \pm 1, \pm ia). \end{aligned}$$

It is immediate to see that the set $\mathcal{B} = (u_0, u_{\pm})$ gives an orthonormal basis for H^3.

Exercise 12.2.10 We close this section by studying an endomorphism which is *not* normal, and indeed diagonalisable with an eigenvector basis which is not orthonormal. In H^2 with respect to $\mathcal{E} = (e_1, e_2)$, consider the endomorphism whose representing matrix is

$$M = \begin{pmatrix} 0 & 1 \\ a & 0 \end{pmatrix}$$

with $a \in \mathbb{R}$. Than M is normal if and only if $a = 1$. The characteristic equation is

$$p_M(T) = T^2 - a = 0$$

so its spectral decomposition is given by

$$\lambda_{\pm} = \pm\sqrt{a}, \qquad V_{\lambda_{\pm}} = \mathcal{L}(u_{\pm}) \quad \text{with} \quad u_{\pm} = (1, \pm\sqrt{a})_{\mathcal{E}}.$$

Being $\langle u_+ | u_- \rangle = 1 - a$, the eigenvectors are orthogonal if and only if M is normal.

12.3 The Unitary Group

If $A, B \in \mathbb{C}^{n,n}$ are two unitary matrices, $A^{\dagger}A = I_n$ and $B^{\dagger}B = I_n$ (see the Definition 12.1.7), one has $(AB)^{\dagger}AB = B^{\dagger}A^{\dagger}AB = I_n$. Furthermore, $\det(A^{\dagger}) = \overline{\det(A)}$, so from $\det(AA^{\dagger}) = 1$ we have $|\det(A)| = 1$. Clearly, the identity matrix I_n is unitary and these leads to the following definition.

Definition 12.3.1 The collection of $n \times n$ unitary matrices is a group, called the *unitary group* of order n and denoted $\mathrm{U}(n)$. The subset $\mathrm{SU}(n) = \{A \in \mathrm{U}(n) : \det(A) = 1\}$ is a subgroup of $\mathrm{U}(n)$, called the *special unitary group* of order n.

Remark 12.3.2 With the the natural inclusion of real matrices as complex matrices whose entries are invariant under complex conjugation, it is clear that $\mathrm{O}(n)$ is a subgroup of $\mathrm{U}(n)$ and $\mathrm{SO}(n)$ is a subgroup of $\mathrm{SU}(n)$.

Now, the exponential of a matrix as in the Definition 11.2.1 can be extended to complex matrices. Thus, for a matrix $A \in \mathbb{C}^{n,n}$, its exponential is defined by by the expansion,

$$e^A = \sum_{k=0}^{\infty} \tfrac{1}{k!} A^k .$$

Then, all properties in the Proposition 11.2.2 have a counterpart for complex matrices, with point (e) there now reading $e^{A^\dagger} = (e^A)^\dagger$.

Theorem 12.3.3 *Let $M, U \in \mathbb{C}^{n,n}$. One has the following results.*

(a) If $M^\dagger = -M$, then $e^M \in \mathrm{U}(n)$. If $M^\dagger = -M$ and $\mathrm{tr}(M) = 0$, then $e^M \in \mathrm{SU}(n)$.
(b) Conversely, if $UU^\dagger = I_n$, there exists a skew-adjoint matrix $M = -M^\dagger$ such that $U = e^M$. If U is a special unitary matrix, there exists a skew-adjoint traceless matrix, $M = -M^\dagger$ with $\mathrm{tr}(M) = 0$, such that $U = e^M$.

Proof Let M be a skew-adjoint matrix. From the previous section we know that there exists a unitary matrix V such that $M = V\Delta_M V^\dagger$, with $\Delta_M = \mathrm{diag}(\mathrm{i}\rho_1, \dots, \mathrm{i}\rho_n)$ for $\rho_a \in \mathbb{R}$. We can then write

$$e^M = e^{V\Delta_M V^\dagger} = V\, e^{\Delta_M}\, V^\dagger$$

with $e^{\Delta_M} = \mathrm{diag}(e^{\mathrm{i}\rho_1}, \dots, e^{\mathrm{i}\rho_n})$. This means that e^{Δ_M} is a unitary matrix, and we can conclude that the starting matrix e^M is unitary. If $\mathrm{tr}(M) = 0$, then e^M is a special unitary matrix.

Alternatively, the result can be shown as follows. If $M = -M^\dagger$, then

$$(e^M)^\dagger = e^{M^\dagger} = e^{-M} = (e^M)^{-1}.$$

This concludes the proof of point (a).

Consider then a unitary matrix U. Since U is normal, there exists a unitary matrix V such that $U = V\Delta_U V^\dagger$ with $\Delta_U = \mathrm{diag}(e^{\mathrm{i}\varphi_1}, \dots, e^{\mathrm{i}\varphi_n})$, where $e^{\mathrm{i}\varphi_k}$ are the modulus 1 eigenvalues of U. Clearly, the matrix Δ_U can be written as

$$\Delta_U = e^{\delta_U}$$

with $\delta_U = \mathrm{diag}(\mathrm{i}\varphi_1, \dots, \mathrm{i}\varphi_n) = -(\delta_U)^\dagger$. This means that

$$U = V\, e^{\delta_U}\, V^\dagger = e^{V\delta_U V^\dagger}$$

with $(V\delta_U V^\dagger)^\dagger = -(V\delta_U V^\dagger)$. If $U \in \mathrm{SU}(n)$, then one has $\mathrm{tr}(V\delta_U V^\dagger) = 0$. This establishes point (b) and concludes the proof. □

Exercise 12.3.4 Consider the matrix, with $a, b \in \mathbb{R}$,

$$A = \begin{pmatrix} b & a \\ a & b \end{pmatrix} = A^{\dagger}.$$

Its eigenvalues λ are given by the solutions of the characteristic equation

$$p_A(T) = (b-T)^2 - a^2 = (b - T - a)(b - T + a) = 0.$$

Its spectral decomposition turns out to be

$$\lambda_{\pm} = b \pm a, \qquad V_{\lambda_{\pm}} = \mathcal{L}((1, \pm 1)).$$

To exponentiate the skew-adjoint matrix iA we can follow two ways.

- By normalising the eigenvectors, we have the conjugation with its diagonal form $A = V \Delta_A V^{\dagger}$,

$$\begin{pmatrix} b & a \\ a & b \end{pmatrix} = \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} b-a & 0 \\ 0 & b+a \end{pmatrix} \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

so we have

$$\begin{aligned} e^{\mathrm{i}A} &= e^{\mathrm{i}V\Delta_A V^{\dagger}} = V e^{\mathrm{i}\Delta_A} V^{\dagger} = \tfrac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} e^{\mathrm{i}(b-a)} & 0 \\ 0 & e^{\mathrm{i}(b+a)} \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\ &= \tfrac{1}{2} \begin{pmatrix} e^{\mathrm{i}(b-a)} + e^{\mathrm{i}(a+b)} & -e^{\mathrm{i}(b-a)} + e^{\mathrm{i}(a+b)} \\ -e^{\mathrm{i}(b-a)} + e^{\mathrm{i}(a+b)} & e^{\mathrm{i}(b-a)} + e^{\mathrm{i}(a+b)} \end{pmatrix} = \begin{pmatrix} e^{\mathrm{i}b} \cos a & \mathrm{i}e^{ib} \sin a \\ \mathrm{i}e^{\mathrm{i}b} \sin a & e^{\mathrm{i}b} \cos a \end{pmatrix}. \end{aligned}$$

  Notice that $\det(e^{\mathrm{i}A}) = e^{2\mathrm{i}b} = e^{\mathrm{i}\,\mathrm{tr}(A)}$.

- By setting

$$A = \tilde{A} + \tilde{B} = \begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix} + \begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix}$$

  we see that A is the sum of two commuting matrices, since $\tilde{B} = bI_2$. So we can write

$$e^{\mathrm{i}A} = e^{\mathrm{i}(\tilde{A}+\tilde{B})} = e^{\mathrm{i}\tilde{A}} e^{\mathrm{i}\tilde{B}}.$$

  Since $\tilde{B}$ is diagonal, $e^{\mathrm{i}\tilde{B}} = \mathrm{diag}(e^{\mathrm{i}b}, e^{ib})$. Computing as in the Exercise 11.2.4 we have

$$\tilde{A}^{2k} = \begin{pmatrix} (-1)^k a^{2k} & 0 \\ 0 & (-1)^k a^{2k} \end{pmatrix}, \qquad \tilde{A}^{2k+1} = \begin{pmatrix} 0 & (-1)^k \mathrm{i} a^{2k+1} \\ (-1)^k \mathrm{i} a^{2k+1} & 0 \end{pmatrix}$$

  so that

$$e^{\mathrm{i}\tilde{A}} = \begin{pmatrix} \cos a & \mathrm{i} \sin a \\ \mathrm{i} \sin a & \cos a \end{pmatrix}$$

and

$$e^{\mathrm{i}A} = \begin{pmatrix} \cos a & \mathrm{i}\sin a \\ \mathrm{i}\sin a & \cos a \end{pmatrix} \begin{pmatrix} e^{\mathrm{i}b} & 0 \\ 0 & e^{\mathrm{i}b} \end{pmatrix} = \begin{pmatrix} e^{\mathrm{i}b}\cos a & \mathrm{i}e^{\mathrm{i}b}\sin a \\ \mathrm{i}e^{\mathrm{i}b}\sin a & e^{\mathrm{i}b}\cos a \end{pmatrix}.$$

Exercise 12.3.5 In this exercise we describe how to reverse the construction of the previous one. That is, given the unitary matrix

$$U = \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix},$$

we determine the self-adjoint matrix $A = A^{\dagger}$ such that $U = e^{\mathrm{i}A}$. Via the usual techniques it is easy to show that the spectral decomposition of U is given by

$$\lambda_{\pm} = \frac{a \pm \mathrm{i}}{\sqrt{1+a^2}}, \quad \text{with} \quad V_{\lambda_{\pm}} = \mathcal{L}((1, \pm\mathrm{i})).$$

Notice that $|\lambda_{\pm}| = 1$ so we can write $\lambda_{\pm} = e^{\mathrm{i}\varphi_{\pm}}$ and, by normalising the eigenvectors for U,

$$U = V\Delta_U V^{\dagger} = \tfrac{1}{2} \begin{pmatrix} 1 & 1 \\ -\mathrm{i} & \mathrm{i} \end{pmatrix} \begin{pmatrix} e^{\mathrm{i}\varphi_-} & 0 \\ 0 & e^{\mathrm{i}\varphi_+} \end{pmatrix} \begin{pmatrix} 1 & \mathrm{i} \\ 1 & -\mathrm{i} \end{pmatrix},$$

with $V^{\dagger}V = I_2$. Since $\Delta_U = e^{\mathrm{i}\delta_U}$ with $\delta_U = \delta_U^{\dagger} = \mathrm{diag}(\varphi_-, \varphi_+)$, we write

$$U = Ve^{\mathrm{i}\delta_U}V^{\dagger} = e^{\mathrm{i}V\delta_U V^{\dagger}} = e^{\mathrm{i}A}$$

where $A = A^{\dagger}$ with

$$A = V\delta_U V^{\dagger} = \tfrac{1}{2} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} e^{i\varphi_-} & 0 \\ 0 & e^{i\varphi_+} \end{pmatrix} \begin{pmatrix} 1 & \mathrm{i} \\ 1 & -\mathrm{i} \end{pmatrix} = \tfrac{1}{2} \begin{pmatrix} \varphi_- + \varphi_+ & \mathrm{i}(\varphi_- - \varphi_+) \\ \mathrm{i}(\varphi_- - \varphi_+) & \varphi_- + \varphi_+ \end{pmatrix}.$$

Notice that the matrix A is not uniquely determined by U, since the angular variables $\varphi_{\pm}$ are defined up to 2π periodicity by

$$\cos\varphi_{\pm} = \frac{a}{\sqrt{1+a^2}}, \qquad \sin\varphi_{\pm} = \pm\frac{1}{\sqrt{1+a^2}}.$$

We close this section by considering one parameter groups of unitary matrices. We start with a self-adjoint matrix $A = A^{\dagger} \in \mathbb{C}^{n,n}$, and define the matrix

$$U_s = e^{\mathrm{i}sA}, \quad \text{for} \quad s \in \mathbb{R}.$$

From the properties of the exponential of a matrix, it is easy to show that, for any real s, s', the following identities hold.

(i) $(U_s)^\dagger U_s = I_n$,
that is U_s is unitary,
(ii) $U_0 = I_n$,
(iii) $(U_s)^\dagger \;=\; U_{-s}$,
(iv) $U_{s+s'} \;=\; U_s U_{s'} \;=\; U_{s'} U_s$,
thus in particular, these unitary matrices commute for different values of the parameter.

The map $\mathbb{R} \to \mathrm{U}(n)$ given by $s \mapsto U_s$ is, according to the definition in the Appendix A.4, a group homomorphism between $(\mathbb{R}, +)$ and $\mathrm{U}(n)$ (with group multiplication), that is between the abelian group $\mathbb{R}$ with respect to the sum and the non abelian group $\mathrm{U}(n)$ with respect to the matrix product. This leads to the following definition.

Definition 12.3.6 If U_s is a family (labelled by a real parameter s) of elements in $\mathrm{U}(n)$ such that, for any value of $s \in \mathbb{R}$, the above identities ii) – iv) are fulfilled, then U_s is called a *one parameter group of unitary matrices* of order n.

For any self-adjoint matrix A, we have a one parameter group of unitary matrices given by $U_s = e^{\mathrm{i}sA}$. The matrix A is usually called the *infinitesimal generator* of the one parameter group.

Proposition 12.3.7 *For any $A = A^\dagger \in \mathbb{C}^{n,n}$, the elements $U_s = e^{\mathrm{i}sA}$ give a one parameter group of unitary matrices in H^n. Conversely, if U_s is a one parameter group of unitary matrices in H^n, there exists a self-adjoint matrix $A = A^\dagger$ such that $U_s = e^{\mathrm{i}sA}$.*

Proof Let $U_s \in \mathrm{U}(n)$ be a one parameter group of unitary matrices. For each value $s \in \mathbb{R}$ the matrix U_s can be diagonalised, and since U_s commutes with any $U_{s'}$, it follows that there exists an orthonormal basis $\mathcal{B}$ for H^n of common eigenvectors for any U_s. So there is a unitary matrix V (providing the change of basis from $\mathcal{B}$ to the canonical base $\mathcal{E}$) such that

$$U_s \;=\; V\{\mathrm{diag}(e^{\mathrm{i}\varphi_1(s),\dots,\mathrm{i}\varphi_n(s)})\}V^\dagger$$

where $e^{\mathrm{i}\varphi_k(s)}$ are the eigenvalues of U_s. From the condition $U_s U_{s'} = U_{s+s'}$ it follows that the dependence of the eigenvalues on the parameter s is linear, and from $U_0 = I_n$ we know that $\varphi_k(s = 0) = 0$. We can eventually write

$$U_s \;=\; V\{\mathrm{diag}(e^{\mathrm{i}s\varphi_1,\dots,\mathrm{i}s\varphi_n})\}V^\dagger \;=\; V e^{\mathrm{i}s\,\delta} V^\dagger \;=\; e^{\mathrm{i}s\,V\delta V^\dagger}$$

where $\delta = \mathrm{diag}(\varphi_1, \dots, \varphi_n)$ is a self-adjoint matrix. We then set $A = V\delta V^\dagger = A^\dagger$ to be the infinitesimal generator of the given one parameter group of unitary matrices. □