Corollary 3.3.30 (QR factorisation)**.** *Let $A \in \mathrm{GL}_n(\mathbb{R})$. Then, there exists an orthogonal matrix $Q \in O(n)$ and a upper-triangular matrix R such that*

$$A = QR.$$

Proof: This is a simple restatement of the Gram-Schmidt process. Suppose that

$$A = [a_1 \ \cdots \ a_n].$$

Then $\mathcal{B} = (a_1, \dots, a_n)$ is an ordered basis of $\mathbb{R}^n$. Apply the Gram-Schmidt process (with respect to the dot product) to obtain an orthonormal basis $\mathcal{B}' = (b_1, \dots, b_n)$ as above. Then, we have

$$\begin{aligned} b_1 &= \tfrac{1}{r_1} a_1 \\ b_2 &= \tfrac{1}{r_2}(a_2 - (a_2 \cdot b_1) b_1) \\ &\vdots \\ b_n &= \tfrac{1}{r_n}(a_n - (a_n \cdot b_1) b_1 - \dots - (a_n \cdot b_{n-1}) b_{n-1}) \end{aligned}$$

where $r_i \in \mathbb{R}_{>0}$ is the length of the c_i vectors from the Gram-Schmidt process. We have also slightly modified the Gram-Schmidt process (in what way?) but you can check that $(b_1, \dots, b_n)$ is an orthonormal basis.[74]

By moving all b_i terms to the left hand side of the above equations we obtain the table

$$\begin{aligned} r_1 b_1 &= a_1 \\ (a_2 \cdot b_1) b_1 + r_2 b_2 &= a_2 \\ &\vdots \\ (a_n \cdot b_1) b_1 + \dots + (a_n \cdot b_{n-1}) b_{n-1} + r_n b_n &= a_n \end{aligned}$$

and we can rewrite these equations using matrices: if

$$Q = [b_1 \ \cdots \ b_n] \in O(n), \quad R = \begin{bmatrix} r_1 & a_2 \cdot b_1 & a_3 \cdot b_1 & \cdots & a_n \cdot b_1 \\ 0 & r_2 & a_3 \cdot b_2 & \cdots & a_n \cdot b_2 \\ 0 & 0 & r_3 & \cdots & a_n \cdot b_3 \\ \vdots & & & \ddots & \vdots \\ 0 & & \cdots & & r_n \end{bmatrix},$$

then we see that the above equations correspond to

$$QR = A.$$

□

3.4 Hermitian spaces

In this section we will give a (very) brief introduction to the definition and fundamental properties of Hermitian forms and Hermitian spaces. A Hermitian form can be considered as a 'quasi-bilinear form' on complex vector spaces.

Definition 3.4.1. Let V be a $\mathbb{C}$-vector space. A function

$$H : V \times V \to \mathbb{C} \ ; \ (u, v) \mapsto H(u, v),$$

is called a *Hermitian form on V* if

(HF1) for any $u, v, w \in V$, $\lambda \in \mathbb{C}$,

$$H(u + \lambda v, w) = H(u, w) + \lambda H(v, w),$$

[74] Do this!

(HF2) for any $u, v \in V$,
$$H(u, v) = \overline{H(v, u)}, \qquad (\textit{Hermitian symmetric})$$
where, if $z = a + \sqrt{-1}b \in \mathbb{C}$, we define the *complex conjugate of z* to be the complex number
$$\overline{z} = a - \sqrt{-1}b \in \mathbb{C}.$$

We denote the set of all Hermitian forms on V by $\text{Herm}(V)$.

Remark 3.4.2. It is a direct consequence of the above definition that if H is a Hermitian form on V we have
$$H(u, v + \lambda w) = H(u, v) + \overline{\lambda}H(v, w),$$
for any $u, v, w \in V, \lambda \in \mathbb{C}$.

We say that a Hermitian form is

'linear in the first argument, antilinear[75] in the second argument'

Definition 3.4.3. Let V be a $\mathbb{C}$-vector space, $\mathcal{B} = (b_1, \dots, b_n) \subset V$ an ordered basis and H a Hermitian form on V. Define *the matrix of H with respect to* $\mathcal{B}$, to be the matrix
$$[H]_{\mathcal{B}} = [a_{ij}], \quad a_{ij} = H(b_i, b_j).$$

The Hermitian symmetric property of a Hermitian form implies that
$$[H]_{\mathcal{B}} = \overline{[H]}_{\mathcal{B}}^t,$$
where, for any matrix $A = [a_{ij}] \in Mat_{m,n}(\mathbb{C})$, we define
$$\overline{A} = [b_{ij}], \quad b_{ij} = \overline{a_{ij}}.$$

A matrix $A \in Mat_n(\mathbb{C})$ is called a *Hermitian matrix* if
$$A = \overline{A}^t.$$

For any $A \in Mat_n(\mathbb{C})$, we will write
$$A^h \stackrel{def}{=} \overline{A}^t;$$
hence, a matrix $A \in Mat_n(\mathbb{C})$ is Hermitian if $A^h = A$.

Lemma 3.4.4. *For any* $A, B \in Mat_n(\mathbb{C}), \eta \in \mathbb{C}$ *we have*

- $(A + B)^h = A^h + B^h$,
- $(AB)^h = B^h A^h$,
- $(\eta A)^h = \overline{\eta} A^h$.

Lemma 3.4.5. *Let V be a $\mathbb{C}$-vector space, $\mathcal{B} \subset V$ an ordered basis of V and H a Hermitian form on V. Then, for any $u, v \in V$, we have*
$$H(u, v) = [u]_{\mathcal{B}}^t [H]_{\mathcal{B}} \overline{[v]}_{\mathcal{B}}.$$

Moreover, if $A \in Mat_n(\mathbb{C})$ is any matrix such that
$$H(u, v) = [u]_{\mathcal{B}}^t A \overline{[v]}_{\mathcal{B}},$$
for every $u, v \in V$, then $A = [H]_{\mathcal{B}}$.

Example 3.4.6. 1. Consider the function

$$H : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C} \ ; \ (\underline{z}, \underline{w}) \mapsto z_1\overline{w}_1 + \sqrt{-1}z_2\overline{w}_1 - \sqrt{-1}z_1\overline{w}_2.$$

H is a Hermitian form on $\mathbb{C}^2$.

2. The function

$$H : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C} \ ; \ (\underline{z}, \underline{w}) \mapsto z_1 w_1 + z_2 w_2,$$

is NOT a Hermitian form on $\mathbb{C}^2$: it is easy to see that

$$H\left(\begin{bmatrix} 1 \\ \sqrt{-1} \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = 1 + \sqrt{-1} \neq 1 - \sqrt{-1} = \overline{H\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ \sqrt{-1} \end{bmatrix}\right)}.$$

3. The function

$$H : \mathbb{C} \times \mathbb{C} \to \mathbb{C} \ ; \ (z, w) \mapsto z\overline{w},$$

is a Hermitian form on $\mathbb{C}$.

4. Let $A = a_{ij} \in Mat_n(\mathbb{C})$ be a Hermitian matrix. Then, we define

$$H_A : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C} \ ; \ (\underline{z}, \underline{w}) \mapsto \underline{z}^t A \underline{\overline{w}} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} z_i \overline{w}_j.$$

H_A is a Hermitian form on $\mathbb{C}^n$. Moreover, **any Hermitian form H on $\mathbb{C}^n$ is of the form $H = H_A$, for some Hermitian matrix $A \in Mat_n(\mathbb{C})$.**

Lemma 3.4.7. *Let $H \in \mathrm{Herm}(V)$, $\mathcal{B}, \mathcal{C} \subset V$ ordered bases on V. Then, if $P = P_{\mathcal{C} \leftarrow \mathcal{B}}$ is the change of coordinate matrix from $\mathcal{B}$ to $\mathcal{C}$, then*

$$P^h [H]_{\mathcal{C}} P = [H]_{\mathcal{B}}.$$

Definition 3.4.8. Let $H \in \mathrm{Herm}(V)$. We say that H is *nondegenerate* if $[H]_{\mathcal{B}}$ is invertible, for any basis $\mathcal{B} \subset V$. The previous lemma ensures that this notion of nondegeneracy is well-defined (ie, does not depend on the choice of basis $\mathcal{B}$).[76]

Theorem 3.4.9 (Classification of Hermitian forms)**.** *Let V be a $\mathbb{C}$-vector space, $n = \dim V$ and $H \in \mathrm{Herm}(V)$ be nondegenerate. Then, there is an ordered basis $\mathcal{B} \subset V$ such that*

$$[H]_{\mathcal{B}} = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix}, \quad d_i \in \{1, -1\}.$$

Hence, if $u, v \in V$ with

$$[u]_{\mathcal{B}} = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}, \ [v]_{\mathcal{B}} = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix},$$

then we have

$$H(u, v) = \sum_{i=1}^n d_i \xi_i \overline{\eta}_i.$$

Proof: The proof is similar to the proof of Theorem 3.2.6 and uses the following facts: for any Hermitian form $H \in \mathrm{Herm}(V)$, there exists $v \in V$ such that $H(v, v) \neq 0$; if $H \in \mathrm{Herm}(V)$ is nondegenerate then for any subspace $U \subset V$ we have $V = U \oplus U^\perp$. The first fact follows from an analagous 'polarisation identity' for Hermitian forms. □

[76] Note that the determinant of A^h is equal to $\overline{\det A}$: indeed, we have

$$\det(A^h) = \det(\overline{A}^t) = \det \overline{A} = \overline{\det A}.$$

Definition 3.4.10. A *Hermitian (or unitary) space* is a pair (V, H), where V is a $\mathbb{C}$-vector space and H is a Hermitian form on V such that $[H]_{\mathcal{B}} = I_n$, for some basis $\mathcal{B}$. This condition implies that H is nondegenerate.

If (V, H) is a Hermitian space and $E \subset V$ is a nonempty subset then we define the *orthogonal complement of E (with respect to H)* to be the subspace

$$E^{\perp} = \{v \in V \mid H(v, u) = 0, \text{ for every } u \in E\}.$$

We say that $z, w \in V$ are *orthogonal (with respect to H)* if $H(z, w) = 0$. We say that $E \subset V$ is *orthogonal* if $H(s, t) = 0$, for every $s \neq t \in E$.

A basis $\mathcal{B} \subset V$ is an *orthogonal basis* if $\mathcal{B}$ is an orthogonal set. A basis $\mathcal{B} \subset V$ is an *orthonormal basis* if it is an orthogonal basis and $H(b, b) = 1$, for every $b \in \mathcal{B}$.

We define $\mathbb{H}^n = (\mathbb{C}^n, H_{I_n})$, where

$$H_{I_n}(\underline{z}, \underline{w}) = z_1\overline{w}_1 + ... + z_n\overline{w}_n.$$

As in the Euclidean case we obtain the notion of a 'Hermitian morphism': a *Hermitian morphism* $f : (V, H_V) \to (W, H_W)$ is a linear morphism such that

$$H_W(f(u), f(v)) = H_V(u, v), \text{ for any } u, v \in V.$$

In particular, if (V, H) is a Hermitian space then we denote the set of all Hermitian isomorphisms of (V, H) by $U(V, H)$, or simply $U(V)$ when there is no confusion. A Hermitian isomorphism is also called a *unitary transformation of V*. Thus,

$$U(V) = \{f : V \to V \mid H(u, v) = H(f(u), f(v)), \text{ for any } u, v \in V\}.$$

We denote $U(n) = U(\mathbb{H}^n)$ and it is straightforward to verify[77] that

$$U(n) = \{T_A \in \text{End}_{\mathbb{C}}(\mathbb{C}^n) \mid A \in Mat_n(\mathbb{C}) \text{ and } A^hA = I_n\}.$$

We say that $A \in Mat_n(\mathbb{C})$ is a *unitary matrix* if

$$A^hA = I_n.$$

Thus, we can identify the set of unitary transformations of $\mathbb{H}^n$ with the set of unitary matrices. Moreover, this association is an **isomorphism of groups**.

As a consequence of Theorem 3.4.9 we can show that there is essentially only one Hermitian space of any given dimension.

Theorem 3.4.11. *Let (V, H) be a Hermitian space, $n = \dim V$. Then, there is a Hermitian isomorphism*

$$f : (V, H) \to \mathbb{H}^n.$$

Remark 3.4.12. There are generalisations to Hermitian spaces of most of the results that apply to Euclidean spaces (section 3.3). In particular, we obtain notions of length and Cauchy-Schwarz/triangle inequalities. For details see [1], section 9.2.

[77] Every linear endomorphism f of $\mathbb{C}^n$ is of the form $f = T_A$, for some $A \in Mat_n(\mathbb{C})$. Then, for f to be a Hermitian morphism we must have

$$\underline{z}^t\overline{\underline{w}} = (A\underline{z})^t\overline{A\underline{w}} = \underline{z}^tA^t\overline{A}\overline{\underline{w}}, \text{ for every } \underline{z}, \underline{w} \in \mathbb{C}^n.$$

This implies that $A^t\overline{A} = I_n$, which is equivalent to the condition $A^hA = I_n$.

3.5 The spectral theorem

In this section we will discuss the diagonalisabliity properties of morphisms in Euclidean/Hermitian spaces. The culmination of this discussion is the **spectral theorem**: this states that self-adjoint morphisms are orthogonally/unitarily diagonalisable and have real eigenvalues. This means that such morphisms are diagonalisable and, moreover, there exists an orthonormal basis of eigenvectors.

Throughout section 3.5 we will only be considering Euclidean (resp. Hermitian) spaces $(V, \langle , \rangle)$ (resp. (V, H)) and, as such, will denote such a space by V, the inner product (resp. Hermitian form) being implicitly assumed given.

First we will consider f-invariant subspaces $U \subset V$ and their orthogonal complements, for an orthogonal/unitary transformation $f : V \rightarrow V$.

Proposition 3.5.1. *Let $f : V \rightarrow V$ be an orthogonal (resp. unitary) transformation of the Euclidean (resp. Hermitian) space V and $U \subset V$ be an f-invariant subspace. Then, $U^\perp$ is f^+-invariant, where $f^+ : V \rightarrow V$ is the adjoint of f (with respect to the corresponding inner product/Hermitian form).*[78]

Proof: To say that U is $f-$-invariant means that, for every $u \in U$, $f(u) \in U$. Consider the orthogonal complement of U in V, $U^\perp$ and let $w \in U^\perp$. Then, we want to show that $f^+(w) \in U^\perp$. Now, for each $u \in U$, we have

$$H(u, f^+(w)) = H(f(u), w) = 0,$$

as $f(u) \in U$. Hence, $f^+(w) \in U^\perp$ and $U^\perp$ is f^+-invariant. □

Lemma 3.5.2. *Let (V, H) be a Hermitian space and $U \subset V$ be a subspace. Then, the restriction of H to U is nondegenerate.*

Proof: Suppose that $v \in U$ is such that $H(u, v) = 0$, for every $u \in U$. Then, $V = U \oplus U^\perp$ (as H is nondegenerate). Hence, if $w \in V$ then $w = u + z$, with $u \in U, z \in U^\perp$ and

$$H(w, v) = H(u + z, v) = H(u, v) + H(z, v) = 0 + 0 = 0.$$

Hence, using nondegeneracy of H on V we have $v = 0_V$ and the restriction of H to U is nondegenerate. □

3.5.1 Normal morphisms

Throughout this section we will assume that V is a Hermitian space, equipped with the Hermitian form H. The results all hold for Euclidean spaces with appropriate modifications to statements of results and to proofs.[79]

Definition 3.5.3 (Normal morphism)**.** Let V be a Hermitian space. We say that $f : V \rightarrow V$ is a *normal morphism* if we have

$$f \circ f^+ = f^+ \circ f.$$

[78]Given a linear morphism $f : V \rightarrow V$, where (V, H) is a Hermitian space, we define the *adjoint of f* to be the morphism

$$f^+ = \sigma_H^{-1} \circ f^* \circ \sigma_H : V \rightarrow V,$$

where

$$\sigma_H : V \rightarrow V^* \ ; \ v \mapsto \sigma_H(v), \text{ so that } (\sigma_H(v))(u) = H(u, v).$$

It is important to note that σ_H **is NOT** $\mathbb{C}$**-linear**: we have $\sigma_H(\lambda v) = \overline{\lambda}\sigma_H(v)$, for any $\lambda \in \mathbb{C}$. However, the composition $\sigma_H^{-1} \circ f^* \circ \sigma_H$ **IS linear** (check this). The definition of f^+ implies that, for every $u, v \in V$, we have

$$H(f(u), v) = H(u, f^+(v));$$

moreover, f^+ is the unique morphism such that this property holds.

As a result of the nonlinearity of σ_H we **DO NOT** have a nice formula for the matrix of f^+ in general. However, if $V = \mathbb{H}^n$ and $f = T_A \in \text{End}_{\mathbb{C}}(V)$, where $A \in Mat_n(\mathbb{C})$, then $f^+ = T_{A^h}$: indeed, for any $\underline{z}, \underline{w} \in \mathbb{C}^n$ we have

$$H(A\underline{z}, \underline{w}) = (A\underline{z})^t\overline{\underline{w}} = \underline{z}^t A^t \overline{\underline{w}} = \underline{z}^t \overline{A^h \underline{w}} = H(\underline{z}, A^h \underline{w}).$$

[79]We could consider a Euclidean space as being a **real Hermitian space**, since $x = \overline{x}$, for every $x \in \mathbb{R}$.

Example 3.5.4. Let V be a Hermitian (resp. Euclidean) space. Then, unitary (resp. orthogonal) transformations of V are normal.

However, not all normal morphisms are unitary/orthogonal transformations: for example, the morphism $T_A \in \mathrm{End}_{\mathbb{C}}(\mathbb{C}^3)$ defined by the matrix

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix},$$

is normal but does not define a unitary transformation of $\mathbb{H}^3$ (as $A^h A \neq I_3$).

Normal morphisms possess useful orthogonality properties of their eigenvectors.

Lemma 3.5.5. *Let $f : V \to V$ be a normal morphism of the Hermitian space (V, H), $f^+ : V \to V$ the adjoint of f (with respect to H). If $v \in V$ is an eigenvector of f with associated eigenvalue $\lambda \in \mathbb{C}$ then v is an eigenvector of f^+ with associated eigenvalue $\overline{\lambda} \in \mathbb{C}$.*

Proof: First, we claim that E_λ (the λ-eigenspace of f) is f^+-invariant: indeed, for any $u \in E_\lambda$ we want to show that $f^+(u) \in E_\lambda$. Then,

$$f(f^+(u)) = f^+(f(u)) = f^+(\lambda u) = \lambda f^+(u),$$

so that $f^+(u) \in E_\lambda$. Hence, f^+ defines an endomorphism of E_λ. Now, let $v \in E_\lambda$ be nonzero (so that $v \in V$ is an eigenvector of f with associated eigenvalue λ). Then, for any $u \in E_\lambda$ we have

$$H(u, f^+(v)) = H(f(u), v) = H(\lambda u, v) = H(u, \overline{\lambda} v) \implies H(u, f^+(v) - \overline{\lambda} v) = 0, \text{ for every } u \in E_\lambda.$$

Then, by Lemma 3.5.2 we see that

$$f^+(v) - \overline{\lambda} v = 0_V \implies f^+(v) = \overline{\lambda} v,$$

and the result follows. □

Lemma 3.5.6. *Let $f : V \to V$ be a normal morphism of the Hermitian space V. Then, if $v_1, \dots, v_k \in V$ are eigenvectors of f corresponding to* distinct *eigenvectors $\xi_1, \dots, \xi_k$ (so that $\xi_i \neq \xi_j$, $i \neq j$), then $\{v_1, \dots, v_k\}$ is orthogonal.*

Proof: Consider v_i, v_j with $i \neq j$. Then, we have $f(v_i) = \xi_i v_i$ and $f(v_j) = \xi_j v_j$ as v_i, v_j are eigenvectors. Then,

$$\xi_i H(v_i, v_j) = H(\xi_i v_i, v_j) = H(f(v_i), v_j) = H(v_i, f^+(v_j)) = H(v_i, \overline{\xi}_j v_j) = \xi_j H(v_i, v_j),$$

so that

$$(\xi_i - \xi_j) H(v_i, v_j) = 0 \implies H(v_i, v_j) = 0, \text{ since } \xi_i \neq \xi_j.$$

□

Theorem 3.5.7 (Normal morphisms are orthogonally diagonalisable)**.** *Let (V, H) be a Hermitian space, $f : V \to V$ a normal morphism. Then, there exists an orthonormal basis of V consisting of eigenvectors of f.*

Proof: Since V is a $\mathbb{C}$-vector space we can find an eigenvector $v \in V$ of f with associated eigenvalue $\lambda \in \mathbb{C}$ (as there is always a root of the characteristic polynomial χ_f). Let $E_\lambda \subset V$ be the corresponding λ-eigenspace (so that $E_\lambda \neq \{0_V\}$). Consider the orthogonal complement $E_\lambda^\perp$ of E_λ (with respect to H). Then, since H is nondegenerate we have

$$V = E_\lambda \oplus E_\lambda^\perp.[80]$$

We are going to show that $E_\lambda^\perp$ is f-invariant: let $w \in E_\lambda^\perp$, so that for every $v \in E_\lambda$ we have

$$H(u, v) = 0.$$

[80] You can check that $E_\lambda \cap E_\lambda^\perp = \{0_V\}$.

We want to show that $f(w) \in E_\lambda^\perp$. Let $u \in E_\lambda$. Then, using Lemma 3.5.5, we obtain

$$H(f(w), u) = H(w, f^+(u)) = H(w, \overline{\lambda} u) = \lambda H(w, u) = 0.$$

Hence, $f(w) \in E_\lambda^\perp$ and $E_\lambda^\perp$ is f-invariant.

So, we have that $E_\lambda^\perp$ is both f-invariant and f^+-invariant (Proposition 3.5.1) and so f and f^+ define endomorphisms of $E_\lambda^\perp$. Moreover, we see that the restriction of f to $E_\lambda^\perp$ is normal. Hence, we can use an induction argument on $\dim V$ and assume that there exists an orthonormal basis of $E_\lambda^\perp$ consisting of eigenvectors of f, $\mathcal{B}_1$ say. Using the Gram-Schmidt process we can obtain an orthonormal basis of E_λ, $\mathcal{B}_2$ say. Then, $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ is an orthonormal basis (Lemma 3.5.6) and consists of eigenvectors of f. □

Corollary 3.5.8. *1. Let $A \in Mat_n(\mathbb{C})$ be such that*

$$AA^h = A^h A.$$

Then, there exists a unitary matrix $P \in U(n)$ (ie, $P^{-1} = P^h$) such that

$$P^h AP = D,$$

where D is a diagonal matrix.

Remark 3.5.9. Suppose that $A \in Mat_n(\mathbb{R})$. Then, we have

$$A^h = A^t,$$

so that the condition

$$A^h A = AA^h \implies A^t A = AA^t.$$

Thus, if $A^t A = AA^t$ then Corollary 3.5.8 implies that A is diagonalisable. However, it is not necessarily true that there exists $P \in \mathrm{GL}_n(\mathbb{R})$ such that

$$P^{-1}AP = D,$$

with $D \in Mat_n(\mathbb{R})$. For example, consider the matrix

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in Mat_2(\mathbb{R}).$$

Then,

$$A^t A = I_2 = AA^t,$$

so that A is normal. Then, Corollary 3.5.8 implies that we can diagonalise A. However, the eigenvalues of A are $\pm\sqrt{-1}$ so that we must have

$$P^{-1}AP = \pm \begin{bmatrix} \sqrt{-1} & 0 \\ 0 & -\sqrt{-1} \end{bmatrix},$$

so that it is not possible that $P \in \mathrm{GL}_2(\mathbb{R})$.[81]

3.5.2 Self-adjoint operators and the spectral theorem

Definition 3.5.10. Let V be a Hermitian space. We say that a morphism $f \in \mathrm{End}_{\mathbb{C}}(V)$ is *self-adjoint* if $f = f^+$. **Self-adjoint morphisms are normal morphisms**.

Example 3.5.11. Let V be a Hermitian (resp. Euclidean) space. Then, $T_A \in \mathrm{End}(V)$ is self-adjoint if and only if A is Hermitian (resp. symmetric).

[81] Why?

Lemma 3.5.12. *Let V be a Hermitian space, $f \in \mathrm{End}_{\mathbb{C}}(V)$ a self-adjoint morphism. Then, all eigenvalues of f are real numbers.*

Proof: As f is self-adjoint then f is normal. Using Lemma 3.5.5 we know that if $v \in V$ is an eigenvector of f with associated eigenvalue $\lambda \in \mathbb{C}$, then $v \in V$ is an eigenvector of f^+ with associated eigenvalue $\overline{\lambda} \in \mathbb{C}$. As $f = f^+$ we must have that $\lambda = \overline{\lambda}$, which implies that $\lambda \in \mathbb{R}$. □

Since a self-adjoint morphism f is normal (indeed, we have $f \circ f^+ = f \circ f = f^+ \circ f$), then Theorem 3.5.7 implies that V admits an orthonormal basis consisting of eigenvectors of f. This result is commonly referred to as **The Spectral Theorem**.

Theorem 3.5.13 (Spectral theorem)**.** *Let V be a Hermitian space, $f \in \mathrm{End}_{\mathbb{C}}(V)$ a self-adjoint morphism. Then, there exists an orthonormal basis $\mathcal{B}$ of V consisting of eigenvectors of f and such that*

$$[f]_{\mathcal{B}} = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix} \in Mat_n(\mathbb{R}).$$

Corollary 3.5.14. *1. Let $A \in Mat_n(\mathbb{C})$ be Hermitian ($A^h = A$). Then, there exists a unitary matrix $P \in U(n)$ such that*

$$P^h A P = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix}, \text{ where } d_1, \dots, d_n \in \mathbb{R}.$$

2. Let $A \in Mat_n(\mathbb{R})$ be symmetric ($A^t = A$). Then, there exists an orthogonal matrix $P \in O(n)$ such that

$$P^t A P = D,$$

where D is diagonal.

Example 3.5.15. 1. Consider the matrix

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

Then, $A^t = A$ so that there exists $P \in O(3)$ such that $P^t A P$ is diagonal (Theorem 3.5.13).

How do we determine P? We know that A is diagonalisable so we proceed as usual: we find that

$$\chi_A(\lambda) = (1-\lambda)(\lambda - \sqrt{3})(\lambda + \sqrt{3}).$$

Then, if we choose eigenvectors $v_1 \in E_1, v_2 \in E_{-\sqrt{3}}, v_3 \in E_{\sqrt{3}}$ such that $||v_i|| = 1$, then we have

$$P = [v_1\ v_2\ v_3] \in O(3).$$

For example, we can take

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6-2\sqrt{3}}} & \frac{1}{\sqrt{6+2\sqrt{3}}} \\ 0 & \frac{1-\sqrt{3}}{\sqrt{6-2\sqrt{3}}} & \frac{1+\sqrt{3}}{\sqrt{6+2\sqrt{3}}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6-2\sqrt{3}}} & \frac{-1}{\sqrt{6+2\sqrt{3}}} \end{bmatrix} \in O(3)$$

2. Consider the matrix

$$A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & -1-\sqrt{-1} \\ 0 & -1+\sqrt{-1} & 1 \end{bmatrix}.$$

Then, $A = A^h$ so that A is Hermitian. Hence, there exists $P \in U(3)$ such that

$$P^hAP = \begin{bmatrix} d_1 & & \\ & d_2 & \\ & & d_3 \end{bmatrix}, \ d_1, d_2, d_3 \in \mathbb{R}.$$

We first determine

$$\chi_A(\lambda) = -(1+\lambda)^2(\lambda - 2),$$

so that the eigenvalues are $\lambda_1 = -1, \lambda_2 = 2$. Then,

$$E_{-1} = \text{span}_{\mathbb{C}} \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1-\sqrt{-1} \\ -2 \end{bmatrix} \right\}.$$

Since

$$H_{I_3}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1-\sqrt{-1} \\ -2 \end{bmatrix} \right) = 1.0 + 0.(-1+\sqrt{-1}) + 0.(-2) = 0,$$

we have that

$$\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1-\sqrt{-1} \\ -2 \end{bmatrix} \right) = (v_1, v_2)$$

is an orthogonal basis of E_{-1}. In order to obtain an orthonormal basis we must scale v_1, v_2 by $H_{I_3}(v_i, v_i)$. Hence, as

$$H_{I_3}(v_1, v_1) = 1, \ H_{I_3}(v_2, v_2) = 0.0 + (-1-\sqrt{-1})(-1+\sqrt{-1}) + (-2).(-2) = 2 + 4 = 6,$$

we have that

$$\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{6}} \begin{bmatrix} 0 \\ -1-\sqrt{-1} \\ -2 \end{bmatrix} \right)$$

is an orthonormal basis of E_{-1}.

Now, we need only determine a vector $v_3 \in E_2$ for which $H_{I_3}(v_3, v_3) = 1$: such an example is

$$v_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 0 \\ -1-\sqrt{-1} \\ -1 \end{bmatrix}.$$

Hence, if we set

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{-1}{\sqrt{6}} - \sqrt{\frac{-1}{6}} & \frac{-1}{\sqrt{3}} - \sqrt{\frac{-1}{3}} \\ 0 & \frac{-2}{\sqrt{6}} & \frac{-1}{\sqrt{3}} \end{bmatrix},$$

then

$$P^hAP = \begin{bmatrix} -1 & & \\ & -1 & \\ & & 2 \end{bmatrix}.$$

3. Consider the matrix

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}.$$

As $A = A^t$ we can find $P \in O(3)$ such that

$$P^t AP = D,$$

where D is diagonal. We have that

$$\chi_A(\lambda) = -(1-\lambda)^2(\lambda - 4),$$

so that the eigenvalues of A are $\lambda_1 = 1, \lambda_2 = 4$.

We have that

$$E_1 = \text{span}_{\mathbb{R}} \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \right\},$$

where

$$\left(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \right),$$

is a basis of E_1. Using the Gram-Schmidt process we can obtain an orthonormal basis

$$\left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right) \subset E_1.$$

Now, we need to find $v_3 \in E_4$ such that $||v_3|| = 1$: we can take

$$v_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Then, if we let

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix},$$

then $P \in O(3)$ and

$$P^t AP = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 4 \end{bmatrix}.$$

References

[1] Shilov, Georgi E., *Linear Algebra*, Dover Publications 1977.