Chapter 12

Hermitian Spaces

12.1 Sesquilinear Forms, Hermitian Forms, Hermitian Spaces, Pre-Hilbert Spaces

In this chapter, we generalize the basic results of Euclidean geometry presented in Chapter 10 to *vector spaces over the complex numbers*.

Some complications arise, due to *complex conjugation*.

Recall that for any complex number $z \in \mathbb{C}$, if $z = x + iy$ where $x, y \in \mathbb{R}$, we let $\Re z = x$, the real part of z, and $\Im z = y$, the imaginary part of z.

We also denote the *conjugate* of $z = x + iy$ as $\bar{z} = x - iy$, and the absolute value (or length, or modulus) of z as $|z|$. Recall that $|z|^2 = z\bar{z} = x^2 + y^2$.

There are many natural situations where a map $\varphi \colon E \times E \to \mathbb{C}$ is linear in its first argument and only semilinear in its second argument.

For example, the natural inner product to deal with functions $f \colon \mathbb{R} \to \mathbb{C}$, especially Fourier series, is

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)\overline{g(x)}dx,$$

which is semilinear (but not linear) in g.

Definition 12.1. Given two vector spaces E and F over the complex field $\mathbb{C}$, a function $f \colon E \to F$ is *semilinear* if

$$f(u + v) = f(u) + f(v),$$
$$f(\lambda u) = \overline{\lambda} f(u),$$

for all $u, v \in E$ and all $\lambda \in \mathbb{C}$.

Remark: Instead of defining semilinear maps, we could have defined the vector space $\overline{E}$ as the vector space with the same carrier set E, whose addition is the same as that of E, but whose multiplication by a complex number is given by

$$(\lambda, u) \mapsto \overline{\lambda}u.$$

Then, it is easy to check that a function $f\colon E \to \mathbb{C}$ is semilinear iff $f\colon \overline{E} \to \mathbb{C}$ is linear.

We can now define sesquilinear forms and Hermitian forms.

Definition 12.2. Given a complex vector space E, a function $\varphi \colon E \times E \to \mathbb{C}$ is a *sesquilinear form* iff it is linear in its first argument and semilinear in its second argument, which means that

$$\varphi(u_1 + u_2, v) = \varphi(u_1, v) + \varphi(u_2, v),$$
$$\varphi(u, v_1 + v_2) = \varphi(u, v_1) + \varphi(u, v_2),$$
$$\varphi(\lambda u, v) = \lambda \varphi(u, v),$$
$$\varphi(u, \mu v) = \overline{\mu} \varphi(u, v),$$

for all $u, v,\ u_1, u_2,\ v_1, v_2 \in E$, and all $\lambda, \mu \in \mathbb{C}$. A function $\varphi \colon E \times E \to \mathbb{C}$ is a *Hermitian form* iff it is sesquilinear and if

$$\varphi(v, u) = \overline{\varphi(u, v)}$$

for all all $u, v \in E$.

Obviously, $\varphi(0, v) = \varphi(u, 0) = 0$.

Also note that if $\varphi \colon E \times E \to \mathbb{C}$ is sesquilinear, we have

$$
\begin{aligned}
\varphi(\lambda u + \mu v, \lambda u + \mu v) = |\lambda|^2 \varphi(u, u) + \lambda \overline{\mu} \varphi(u, v) \\
+ \overline{\lambda} \mu \varphi(v, u) + |\mu|^2 \varphi(v, v),
\end{aligned}
$$

and if $\varphi \colon E \times E \to \mathbb{C}$ is Hermitian, we have

$$
\begin{aligned}
\varphi(\lambda u + \mu v, &\lambda u + \mu v) \\
&= |\lambda|^2 \varphi(u, u) + 2\Re(\lambda \overline{\mu} \varphi(u, v)) + |\mu|^2 \varphi(v, v).
\end{aligned}
$$

Note that restricted to real coefficients, a sesquilinear form is bilinear (we sometimes say $\mathbb{R}$-bilinear).

The function $\Phi \colon E \to \mathbb{C}$ defined such that $\Phi(u) = \varphi(u, u)$ for all $u \in E$ is called the *quadratic form* associated with φ.

The standard example of a Hermitian form on $\mathbb{C}^n$ is the map φ defined such that

$$\varphi((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = x_1 \overline{y_1} + x_2 \overline{y_2} + \cdots + x_n \overline{y_n}.$$

This map is also positive definite, but before dealing with these issues, we show the following useful proposition.

Proposition 12.1. *Given a complex vector space E, the following properties hold:*

(1) A sesquilinear form $\varphi \colon E \times E \to \mathbb{C}$ is a Hermitian form iff $\varphi(u, u) \in \mathbb{R}$ for all $u \in E$.

(2) If $\varphi \colon E \times E \to \mathbb{C}$ is a sesquilinear form, then

$$\begin{aligned} 4\varphi(u, v) = {} & \varphi(u + v, u + v) - \varphi(u - v, u - v) \\ & + i\varphi(u + iv, u + iv) - i\varphi(u - iv, u - iv), \end{aligned}$$

and

$$\begin{aligned} 2\varphi(u, v) = {} & (1 + i)(\varphi(u, u) + \varphi(v, v)) \\ & - \varphi(u - v, u - v) - i\varphi(u - iv, u - iv). \end{aligned}$$

These are called polarization identities.

Proposition 12.1 shows that a sesquilinear form is completely determined by the quadratic form $\Phi(u) = \varphi(u, u)$, even if φ is not Hermitian.

This is false for a real bilinear form, unless it is symmetric.

For example, the bilinear form $\varphi \colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ defined such that

$$\varphi((x_1, y_1), (x_2, y_2)) = x_1 y_2 - x_2 y_1$$

is not identically zero, and yet, it is null on the diagonal.

However, a real symmetric bilinear form is indeed determined by its values on the diagonal, as we saw in Chapter 10.

As in the Euclidean case, Hermitian forms for which $\varphi(u, u) \geq 0$ play an important role.

Definition 12.3. Given a complex vector space E, a Hermitian form $\varphi \colon E \times E \to \mathbb{C}$ is *positive* iff $\varphi(u, u) \geq 0$ for all $u \in E$, and *positive definite* iff $\varphi(u, u) > 0$ for all $u \neq 0$. A pair $\langle E, \varphi \rangle$ where E is a complex vector space and φ is a Hermitian form on E is called a *pre-Hilbert space* if φ is positive, and a *Hermitian (or unitary) space* if φ is positive definite.

We warn our readers that some authors, such as Lang [24], define a pre-Hilbert space as what we define to be a Hermitian space.

We prefer following the terminology used in Schwartz [28] and Bourbaki [7].

The quantity $\varphi(u, v)$ is usually called the *Hermitian product* of u and v. We will occasionally call it the inner product of u and v.

Given a pre-Hilbert space $\langle E, \varphi \rangle$, as in the case of a Euclidean space, we also denote $\varphi(u, v)$ as

$$u \cdot v, \quad \text{or} \quad \langle u, v \rangle, \quad \text{or} \quad (u|v),$$

and $\sqrt{\Phi(u)}$ as $\|u\|$.

Example 1. The complex vector space $\mathbb{C}^n$ under the Hermitian form

$$\varphi((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n}$$

is a Hermitian space.

Example 2. Let l^2 denote the set of all countably infinite sequences $x = (x_i)_{i \in \mathbb{N}}$ of complex numbers such that $\sum_{i=0}^{\infty} |x_i|^2$ is defined (i.e. the sequence $\sum_{i=0}^{n} |x_i|^2$ converges as $n \to \infty$).

It can be shown that the map $\varphi \colon l^2 \times l^2 \to \mathbb{C}$ defined such that

$$\varphi\left((x_i)_{i \in \mathbb{N}}, (y_i)_{i \in \mathbb{N}}\right) = \sum_{i=0}^{\infty} x_i\overline{y_i}$$

is well defined, and l^2 is a Hermitian space under φ. Actually, l^2 is even a Hilbert space.

Example 3. Consider the set $\mathcal{C}_{piece}[a, b]$ of piecewise bounded continuous functions $f \colon [a, b] \to \mathbb{C}$ under the Hermitian form

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

It is easy to check that this Hermitian form is positive, but it is not definite. Thus, under this Hermitian form, $\mathcal{C}_{piece}[a, b]$ is only a pre-Hilbert space.

Example 4. Let $\mathcal{C}[a, b]$ be the set of complex-valued continuous functions $f \colon [a, b] \to \mathbb{C}$ under the Hermitian form

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

It is easy to check that this Hermitian form is positive definite. Thus, $\mathcal{C}[a, b]$ is a Hermitian space.

Example 5. Let $E = \mathrm{M}_n(\mathbb{C})$ be the vector space of complex $n \times n$ matrices. We define the Hermitian product of two matrices $A, B \in \mathrm{M}_n(\mathbb{C})$ as

$$\langle A, B \rangle = \sum_{i,j=1}^{n} a_{ij} \bar{b}_{ij},$$

which can be conveniently written as

$$\langle A, B \rangle = \mathrm{tr}(A^\top \overline{B}) = \mathrm{tr}(B^* A).$$

Since this can be viewed as the standard Hermitian product on $\mathbb{C}^{n^2}$, it is a Hermitian product on $\mathrm{M}_n(\mathbb{C})$. The corresponding norm

$$\|A\|_F = \sqrt{\mathrm{tr}(A^* A)}$$

is the Frobenius norm (see Section 7.2).

If E is finite-dimensional and if $\varphi \colon E \times E \to \mathbb{R}$ is a sequilinear form on E, given any basis $(e_1, \ldots, e_n)$ of E, we can write $x = \sum_{i=1}^{n} x_i e_i$ and $y = \sum_{j=1}^{n} y_j e_j$, and we have

$$\varphi(x, y) = \sum_{i,j=1}^{n} x_i \overline{y}_j \varphi(e_i, e_j).$$

If we let $G = (g_{ij})$ be the matrix given by $g_{ij} = \varphi(e_j, e_i)$, and if x and y are the column vectors associated with $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$, then we can write

$$\varphi(x, y) = x^\top G^\top \overline{y} = y^* G x,$$

where $\overline{y}$ corresponds to $(\overline{y}_1, \ldots, \overline{y}_n)$.

As in Section 10.1, we are committing the slight abuse of notation of letting x denote both the vector $x = \sum_{i=1}^{n} x_i e_i$ and the column vector associated with $(x_1, \ldots, x_n)$ (and similarly for y). The "correct" expression for $\varphi(x, y)$ is

$$\varphi(x, y) = \mathbf{y}^* G \mathbf{x} = \mathbf{x}^\top G^\top \overline{\mathbf{y}}.$$

Observe that in $\varphi(x, y) = y^* G x$, the matrix involved is the transpose of the matrix $(\varphi(e_i, e_j))$. The reason for this is that we want G to be positive definite when φ is positive definite, not $G^\top$.

Furthermore, observe that φ is Hermitian iff $G = G^*$, and φ is positive definite iff the matrix G is positive definite, that is,

$$(Gx)^\top \overline{x} = x^* G x > 0 \quad \text{for all } x \in \mathbb{C}^n, \ x \neq 0.$$

The matrix G associated with a Hermitian product is called the *Gram matrix* of the Hermitian product with respect to the basis $(e_1, \ldots, e_n)$.

Proposition 12.2. *Let E be a finite-dimensional vector space, and let $(e_1, \ldots, e_n)$ be a basis of E.*

1. *For any Hermitian inner product $\langle -, - \rangle$ on E, if $G = (g_{ij})$ with $g_{ij} = \langle e_j, e_i \rangle$ is the Gram matrix of the Hermitian product $\langle -, - \rangle$ w.r.t. the basis $(e_1, \ldots, e_n)$, then G is Hermitian positive definite.*

2. *For any change of basis matrix P, the Gram matrix of $\langle -, - \rangle$ with respect to the new basis is P^*GP.*

3. *If A is any $n \times n$ Hermitian positive definite matrix, then*
$$\langle x, y \rangle = y^* A x$$
is a Hermitian product on E.

The following result reminiscent of the first polarization identity of Proposition 12.1 can be used to prove that two linear maps are identical.

Proposition 12.3. *Given any Hermitian space E with Hermitian product $\langle -, - \rangle$, for any linear map $f \colon E \to E$, if $\langle f(x), x \rangle = 0$ for all $x \in E$, then $f = 0$.*

One should be careful not to apply Proposition 12.3 to a linear map on a real Euclidean space, because it is false! The reader should find a counterexample.

The Cauchy-Schwarz inequality and the Minkowski inequalities extend to pre-Hilbert spaces and to Hermitian spaces.

Proposition 12.4. *Let $\langle E, \varphi \rangle$ be a pre-Hilbert space with associated quadratic form Φ. For all $u, v \in E$, we have the* Cauchy-Schwarz inequality:

$$|\varphi(u, v)| \leq \sqrt{\Phi(u)}\sqrt{\Phi(v)}.$$

Furthermore, if $\langle E, \varphi \rangle$ is a Hermitian space, the equality holds iff u and v are linearly dependent.

We also have the Minkovski inequality:

$$\sqrt{\Phi(u + v)} \leq \sqrt{\Phi(u)} + \sqrt{\Phi(v)}.$$

Furthermore, if $\langle E, \varphi \rangle$ is a Hermitian space, the equality holds iff u and v are linearly dependent, where in addition, if $u \neq 0$ and $v \neq 0$, then $u = \lambda v$ for some real λ such that $\lambda > 0$.

As in the Euclidean case, if $\langle E, \varphi \rangle$ is a Hermitian space, the Minkovski inequality

$$\sqrt{\Phi(u + v)} \leq \sqrt{\Phi(u)} + \sqrt{\Phi(v)}$$

shows that the map $u \mapsto \sqrt{\Phi(u)}$ is a *norm* on E.

The norm induced by φ is called the *Hermitian norm induced by φ*.

We usually denote $\sqrt{\Phi(u)}$ as $\|u\|$, and the Cauchy-Schwarz inequality is written as

$$|u \cdot v| \leq \|u\| \, \|v\| \,.$$

Since a Hermitian space is a normed vector space, it is a topological space under the topology induced by the norm (a basis for this topology is given by the open balls $B_0(u, \rho)$ of center u and radius $\rho > 0$, where

$$B_0(u, \rho) = \{v \in E \mid \|v - u\| < \rho\}.$$

If E has finite dimension, every linear map is continuous.

Remark: As in the case of real vector spaces, a norm on a complex vector space is induced by some positive definite Hermitian product $\langle -, - \rangle$ iff it satisfies the *parallelogram law*:

$$\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2).$$

This time, the Hermitian product is recovered using the polarization identity from Proposition 12.1:

$$4\langle u, v \rangle = \|u + v\|^2 - \|u - v\|^2 + i \|u + iv\|^2 - i \|u - iv\|^2.$$

The Cauchy-Schwarz inequality

$$|u \cdot v| \leq \|u\| \, \|v\|$$

shows that $\varphi \colon E \times E \to \mathbb{C}$ is continuous, and thus, that $\|\ \|$ is continuous.

If $\langle E, \varphi \rangle$ is only pre-Hilbertian, $\|u\|$ is called a *semi-norm*.

In this case, the condition

$$\|u\| = 0 \quad \text{implies} \quad u = 0$$

is not necessarily true.

However, the Cauchy-Schwarz inequality shows that if $\|u\| = 0$, then $u \cdot v = 0$ for all $v \in E$.

We will now basically mirror the presentation of Euclidean geometry given in Chapter 10 rather quickly, leaving out most proofs, except when they need to be seriously amended.

12.2   Orthogonality, Duality, Adjoint of A Linear Map

In this section, we assume that we are dealing with Hermitian spaces. We denote the Hermitian inner product as $u \cdot v$ or $\langle u, v \rangle$.

The concepts of orthogonality, orthogonal family of vectors, orthonormal family of vectors, and orthogonal complement of a set of vectors, are unchanged from the Euclidean case (Definition 10.2).

For example, the set $\mathcal{C}[-\pi, \pi]$ of continuous functions $f \colon [-\pi, \pi] \to \mathbb{C}$ is a Hermitian space under the product

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)\overline{g(x)}dx,$$

and the family $(e^{ikx})_{k \in \mathbb{Z}}$ is orthogonal.

Proposition 10.3 and 10.4 hold without any changes.

It is easy to show that

$$\left\|\sum_{i=1}^{n} u_i\right\|^2 = \sum_{i=1}^{n} \|u_i\|^2 + \sum_{1 \leq i < j \leq n} 2\Re(u_i \cdot u_j).$$

Analogously to the case of Euclidean spaces of finite dimension, the Hermitian product induces a *canonical bijection* (i.e., independent of the choice of bases) between the vector space E and the space E^*.

This is one of the places where conjugation shows up, but in this case, troubles are minor.

Given a Hermitian space E, for any vector $u \in E$, let $\varphi_u^l \colon E \to \mathbb{C}$ be the map defined such that

$$\varphi_u^l(v) = \overline{u \cdot v}, \quad \text{for all } v \in E.$$

Similarly, for any vector $v \in E$, let $\varphi_v^r \colon E \to \mathbb{C}$ be the map defined such that

$$\varphi_v^r(u) = u \cdot v, \quad \text{for all } u \in E.$$

Since the Hermitian product is linear in its first argument u, the map φ_v^r is a linear form in E^*, and since it is semilinear in its second argument v, the map φ_u^l is also a linear form in E^*.

Thus, we have two maps $\flat^l \colon E \to E^*$ and $\flat^r \colon E \to E^*$, defined such that

$$\flat^l(u) = \varphi_u^l, \quad \text{and} \quad \flat^r(v) = \varphi_v^r.$$

Actually, it is easy to show that $\varphi_u^l = \varphi_u^r$ and $\flat^l = \flat^r$.

Therefore, we use the notation φ_u for both φ_u^l and φ_u^r, and $\flat$ for both $\flat^l$ and $\flat^r$.

Theorem 12.5. *let E be a Hermitian space E. The map $\flat \colon E \to E^*$ defined such that*

$$\flat(u) = \varphi_u^l = \varphi_u^r \quad \text{for all } u \in E$$

is semilinear and injective. When E is also of finite dimension, the map $\flat \colon \overline{E} \to E^*$ is a canonical isomorphism.

The inverse of the isomorphism $\flat\colon \overline{E} \to E^*$ is denoted by $\sharp\colon E^* \to \overline{E}$.

As a corollary of the isomorphism $\flat\colon \overline{E} \to E^*$, if E is a Hermitian space of finite dimension, then every linear form $f \in E^*$ corresponds to a unique $v \in E$, such that

$$f(u) = u \cdot v, \quad \text{for every } u \in E.$$

In particular, if f is not the null form, the kernel of f, which is a hyperplane H, is precisely the set of vectors that are orthogonal to v.

Remark. The "musical map" $\flat\colon \overline{E} \to E^*$ is not surjective when E has infinite dimension.

This result will be salvaged by restricting our attention to continuous linear maps, and by assuming that the vector space E is a *Hilbert space*.

The existence of the isomorphism $\flat\colon \overline{E} \to E^*$ is crucial to the existence of adjoint maps.

Indeed, Theorem 12.5 allows us to define the adjoint of a linear map on a Hermitian space.

Let E be a Hermitian space of finite dimension n, and let $f\colon E \to E$ be a linear map.

For every $u \in E$, the map

$$v \mapsto \overline{u \cdot f(v)}$$

is clearly a linear form in E^*, and by Theorem 12.5, there is a unique vector in E denoted by $f^*(u)$, such that

$$\overline{f^*(u) \cdot v} = \overline{u \cdot f(v)},$$

that is,

$$f^*(u) \cdot v = u \cdot f(v), \quad \text{for every } v \in E.$$

Proposition 12.6. *Given a Hermitian space E of finite dimension, for every linear map $f \colon E \to E$, there is a unique linear map $f^* \colon E \to E$, such that*

$$f^*(u) \cdot v = u \cdot f(v),$$

for all $u, v \in E$. The map f^* is called the adjoint of f (w.r.t. to the Hermitian product).

The fact that

$$v \cdot u = \overline{u \cdot v}$$

implies that the adjoint f^* of f is also characterized by

$$f(u) \cdot v = u \cdot f^*(v),$$

for all $u, v \in E$. It is also obvious that $f^{**} = f$.

Given two Hermitian spaces E and F, where the Hermitian product on E is denoted as $\langle -, - \rangle_1$ and the Hermitian product on F is denoted as $\langle -, - \rangle_2$, given any linear map $f \colon E \to F$, it is immediately verified that the proof of Proposition 12.6 can be adapted to show that there is a unique linear map $f^* \colon F \to E$ such that

$$\langle f(u), v \rangle_2 = \langle u, f^*(v) \rangle_1$$

for all $u \in E$ and all $v \in F$. The linear map f^* is also called the *adjoint* of f.

As in the Euclidean case, a linear map $f \colon E \to E$ (where E is a finite-dimensional Hermitian space) is *self-adjoint* if $f = f^*$. The map f is *positive semidefinite* iff

$$\langle f(x), x \rangle \geq 0 \quad \text{all } x \in E;$$

positive definite iff

$$\langle f(x), x \rangle > 0 \quad \text{all } x \in E, \ x \neq 0.$$

An interesting corollary of Proposition 12.3 is that a positive semidefinite linear map must be self-adjoint. In fact, we can prove a slightly more general result.

Proposition 12.7. *Given any finite-dimensional Hermitian space E with Hermitian product $\langle -, - \rangle$, for any linear map $f \colon E \to E$, if $\langle f(x), x \rangle \in \mathbb{R}$ for all $x \in E$, then f is self-adjoint. In particular, any positive semidefinite linear map $f \colon E \to E$ is self-adjoint.*

As in the Euclidean case, Theorem 12.5 can be used to show that any Hermitian space of finite dimension has an orthonormal basis. The proof is unchanged.

Proposition 12.8. *Given any nontrivial Hermitian space E of finite dimension $n \geq 1$, there is an orthonormal basis $(u_1, \ldots, u_n)$ for E.*

The *Gram–Schmidt orthonormalization procedure* also applies to Hermitian spaces of finite dimension, without any changes from the Euclidean case!

Proposition 12.9. *Given any nontrivial Hermitian space E of finite dimension $n \geq 1$, from any basis $(e_1, \ldots, e_n)$ for E, we can construct an orthonormal basis $(u_1, \ldots, u_n)$ for E, with the property that for every k, $1 \leq k \leq n$, the families $(e_1, \ldots, e_k)$ and $(u_1, \ldots, u_k)$ generate the same subspace.*

Remarks: The remarks made after Proposition 10.8 also apply here, except that in the QR-decomposition, Q is a unitary matrix.

As a consequence of Proposition 10.7 (or Proposition 12.9), given any Hermitian space of finite dimension n, if $(e_1, \ldots, e_n)$ is an orthonormal basis for E, then for any two vectors $u = u_1 e_1 + \cdots + u_n e_n$ and $v = v_1 e_1 + \cdots + v_n e_n$, the Hermitian product $u \cdot v$ is expressed as

$$u \cdot v = (u_1 e_1 + \cdots + u_n e_n) \cdot (v_1 e_1 + \cdots + v_n e_n) = \sum_{i=1}^{n} u_i \overline{v_i},$$

and the norm $\|u\|$ as

$$\|u\| = \|u_1 e_1 + \cdots + u_n e_n\| = \sqrt{\sum_{i=1}^{n} |u_i|^2}.$$

Proposition 10.9 also holds unchanged.

Proposition 12.10. *Given any nontrivial Hermitian space E of finite dimension $n \geq 1$, for any subspace F of dimension k, the orthogonal complement $F^{\perp}$ of F has dimension $n - k$, and $E = F \oplus F^{\perp}$. Furthermore, we have $F^{\perp\perp} = F$.*

12.3 Linear Isometries (also called Unitary Transformations)

In this section, we consider linear maps between Hermitian spaces that preserve the Hermitian norm.

All definitions given for Euclidean spaces in Section 10.5 extend to Hermitian spaces, *except that orthogonal transformations are called unitary transformation*, but Proposition 10.10 only extends with a modified condition (2).

Indeed, the old proof that (2) implies (3) does not work, and the implication is in fact false! It can be repaired by strengthening condition (2). For the sake of completeness, we state the Hermitian version of Definition 10.3.

Definition 12.4. Given any two nontrivial Hermitian spaces E and F of the same finite dimension n, a function $f \colon E \to F$ is *a unitary transformation, or a linear isometry* iff it is linear and

$$\|f(u)\| = \|u\|,$$

for all $u \in E$.

Proposition 10.10 can be salvaged by strengthening condition (2).

Proposition 12.11. *Given any two nontrivial Hermitian space E and F of the same finite dimension n, for every function $f \colon E \to F$, the following properties are equivalent:*

(1) f is a linear map and $\|f(u)\| = \|u\|$, for all $u \in E$;

(2) $\|f(v) - f(u)\| = \|v - u\|$ and $f(iu) = if(u)$, for all $u, v \in E$;

(3) $f(u) \cdot f(v) = u \cdot v$, for all $u, v \in E$.

Furthermore, such a map is bijective.

Observe that from $f(iu) = if(u)$, for $u = 0$, we get $f(0) = if(0)$, which implies that $f(0) = 0$.

Remarks: (i) In the Euclidean case, we proved that the assumption

(2') $\|f(v) - f(u)\| = \|v - u\|$, for all $u, v \in E$, and $f(0) = 0$;

implies (3). For this, we used the polarization identity

$$2u \cdot v = \|u\|^2 + \|v\|^2 - \|u - v\|^2.$$

In the Hermitian case, the polarization identity involves the complex number i.

In fact, the implication (2') implies (3) is *false* in the Hermitian case! Conjugation $z \mapsto \overline{z}$ satifies (2') since

$$|\overline{z_2} - \overline{z_1}| = |\overline{z_2 - z_1}| = |z_2 - z_1|,$$

and yet, it is not linear!

(ii) If we modify (2) by changing the second condition by now requiring that there is some $\tau \in E$ such that

$$f(\tau + iu) = f(\tau) + i(f(\tau + u) - f(\tau))$$

for all $u \in E$, then the function $g \colon E \to E$ defined such that

$$g(u) = f(\tau + u) - f(\tau)$$

satisfies the old conditions of (2), and the implications $(2) \to (3)$ and $(3) \to (1)$ prove that g is linear, and thus that f is affine.

In view of the first remark, some condition involving i is needed on f, in addition to the fact that f is distance-preserving.

We are now going to take a closer look at the isometries $f \colon E \to E$ of a Hermitian space of finite dimension.

12.4 The Unitary Group, Unitary Matrices

In this section, as a mirror image of our treatment of the isometries of a Euclidean space, we explore some of the fundamental properties of the unitary group and of unitary matrices.

Definition 12.5. Given a complex $m \times n$ matrix A, the *transpose $A^\top$ of A* is the $n \times m$ matrix $A^\top = (a_{ij}^\top)$ defined such that

$$a_{ij}^\top = a_{ji}$$

and the *conjugate $\overline{A}$ of A* is the $m \times n$ matrix $\overline{A} = (b_{ij})$ defined such that

$$b_{ij} = \overline{a_{ij}}$$

for all i, j, $1 \leq i \leq m$, $1 \leq j \leq n$. The *adjoint A^* of A* is the matrix defined such that

$$A^* = \overline{(A^\top)} = (\overline{A})^\top.$$

Proposition 12.12. *Let E be any Hermitian space of finite dimension n, and let $f \colon E \to E$ be any linear map. The following properties hold:*

(1) The linear map $f \colon E \to E$ is an isometry iff

$$f \circ f^* = f^* \circ f = \mathrm{id}.$$

(2) For every orthonormal basis $(e_1, \ldots, e_n)$ of E, if the matrix of f is A, then the matrix of f^* is the adjoint A^* of A, and f is an isometry iff A satisfies the idendities

$$A\,A^* = A^* A = I_n,$$

where I_n denotes the identity matrix of order n, iff the columns of A form an orthonormal basis of $\mathbb{C}^n$, iff the rows of A form an orthonormal basis of $\mathbb{C}^n$.

Proposition 10.12 also motivates the following definition.

Definition 12.6. A complex $n \times n$ matrix is a *unitary matrix* iff

$$A \, A^* = A^* A = I_n.$$

Remarks: The conditions $A \, A^* = I_n$, $A^* A = I_n$, and $A^{-1} = A^*$, are equivalent.

Given any two orthonormal bases $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$, if P is the change of basis matrix from $(u_1, \ldots, u_n)$ to $(v_1, \ldots, v_n)$, it is easy to show that the matrix P is unitary.

The proof of Proposition 12.11 (3) also shows that if f is an isometry, then the image of an orthonormal basis $(u_1, \ldots, u_n)$ is an orthonormal basis.

If f is a unitary transformation and A is its matrix with respect to any orthonormal basis, we have $|\det(A)| = 1$.

Definition 12.7. Given a Hermitian space E of dimension n, the set of isometries $f \colon E \to E$ forms a subgroup of $\mathbf{GL}(E, \mathbb{C})$ denoted as $\mathbf{U}(E)$, or $\mathbf{U}(n)$ when $E = \mathbb{C}^n$, called the *unitary group (of E)*. For every isometry, f, we have $|\det(f)| = 1$, where $\det(f)$ denotes the determinant of f. The isometries such that $\det(f) = 1$ are called *rotations, or proper isometries, or proper unitary transformations*, and they form a subgroup of the special linear group $\mathbf{SL}(E, \mathbb{C})$ (and of $\mathbf{U}(E)$), denoted as $\mathbf{SU}(E)$, or $\mathbf{SU}(n)$ when $E = \mathbb{C}^n$, called the *special unitary group (of E)*. The isometries such that $\det(f) \neq 1$ are called *improper isometries, or improper unitary transformations, or flip transformations*.

The Gram–Schmidt orthonormalization procedure immediately yields the QR-decomposition for matrices.

Proposition 12.13. *Given any $n \times n$ complex matrix A, if A is invertible then there is a unitary matrix Q and an upper triangular matrix R with positive diagonal entries such that $A = QR$.*

The proof is absolutely the same as in the real case!

We have the following version of the Hadamard inequality for complex matrices.

Proposition 12.14. *(Hadamard) For any complex $n \times n$ matrix $A = (a_{ij})$, we have*

$$| \det(A)| \leq \prod_{i=1}^{n} \left(\sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}$$

$$and \quad | \det(A)| \leq \prod_{j=1}^{n} \left(\sum_{i=1}^{n} |a_{ij}|^2 \right)^{1/2}.$$

Moreover, equality holds iff either A has orthogonal rows in the left inequality or orthogonal columns in the right inequality.

We also have the following version of Proposition 10.16 for Hermitian matrices.

Proposition 12.15. *(Hadamard) For any complex $n \times n$ matrix $A = (a_{ij})$, if A is Hermitian positive semidefinite, then we have*

$$\det(A) \leq \prod_{i=1}^{n} a_{ii}.$$

Moreover, if A is positive definite, then equality holds iff A is a diagonal matrix.

12.5 Orthogonal Projections and Involutions

In this section, we assume that the field K is not a field of characteristic 2.

Recall that a linear map $f \colon E \to E$ is an *involution* iff $f^2 = \mathrm{id}$, and is *idempotent* iff $f^2 = f$. We know from Proposition 4.5 that if f is idempotent, then

$$E = \mathrm{Im}(f) \oplus \mathrm{Ker}\,(f),$$

and that the restriction of f to its image is the identity.

For this reason, a linear involution is called a *projection*.

Proposition 12.16. *For any linear map $f \colon E \to E$, we have $f^2 = \mathrm{id}$ iff $\frac{1}{2}(\mathrm{id} - f)$ is a projection iff $\frac{1}{2}(\mathrm{id} + f)$ is a projection; in this case, f is equal to the difference of the two projections $\frac{1}{2}(\mathrm{id} + f)$ and $\frac{1}{2}(\mathrm{id} - f)$.*

Let $U^+ = \text{Ker}\left(\frac{1}{2}(\text{id} - f)\right)$ and let $U^- = \text{Im}(\frac{1}{2}(\text{id} - f))$.

If $f^2 = \text{id}$, then

$$(\text{id} - f) \circ (\text{id} + f) = \text{id} - f^2 = \text{id} - \text{id} = 0,$$

which implies that

$$\text{Im}\left(\frac{1}{2}(\text{id} + f)\right) \subseteq \text{Ker}\left(\frac{1}{2}(\text{id} - f)\right).$$

Conversely, if $u \in \text{Ker}\left(\frac{1}{2}(\text{id} - f)\right)$, then $f(u) = u$, so

$$\frac{1}{2}(\text{id} + f)(u) = \frac{1}{2}(u + u) = u,$$

and thus

$$\text{Ker}\left(\frac{1}{2}(\text{id} - f)\right) \subseteq \text{Im}\left(\frac{1}{2}(\text{id} + f)\right).$$

Therefore,

$$U^+ = \text{Ker}\left(\frac{1}{2}(\text{id} - f)\right) = \text{Im}\left(\frac{1}{2}(\text{id} + f)\right),$$

and so, $f(u) = u$ on U^+ and $f(u) = -u$ on U^-.

The involutions of E that are unitary transformations are characterized as follows.

Proposition 12.17. *Let $f \in \mathbf{GL}(E)$ be an involution. The following properties are equivalent:*

(a) The map f is unitary; that is, $f \in \mathbf{U}(E)$.

(b) The subspaces $U^- = \mathrm{Im}(\frac{1}{2}(\mathrm{id} - f))$ and $U^+ = \mathrm{Im}(\frac{1}{2}(\mathrm{id} + f))$ are orthogonal.

Furthermore, if E is finite-dimensional, then (a) and (b) are equivalent to

(c) The map is self-adjoint; that is, $f = f^*$.

A unitary involution is the identity on $U^+ = \mathrm{Im}(\frac{1}{2}(\mathrm{id} + f))$, and $f(v) = -v$ for all $v \in U^- = \mathrm{Im}(\frac{1}{2}(\mathrm{id} - f))$.

Furthermore, E is an orthogonal direct sum $E = U^+ \oplus U^-$.

We say that f is an *orthogonal reflection* about U^+.

In the special case where U^+ is a hyperplane, we say that f is a *hyperplane reflection*. We already studied hyperplane reflections in the Euclidean case; see Chapter 11.

If $f\colon E \to E$ is a projection $(f^2 = f)$, then

$$(\mathrm{id} - 2f)^2 = \mathrm{id} - 4f + 4f^2 = \mathrm{id} - 4f + 4f = \mathrm{id},$$

so $\mathrm{id} - 2f$ is an involution. As a consequence, we get the following result.

Proposition 12.18. *If $f\colon E \to E$ is a projection $(f^2 = f)$, then $\mathrm{Ker}\,(f)$ and $\mathrm{Im}(f)$ are orthogonal iff $f^* = f$.*

A projection such that $f = f^*$ is called an *orthogonal projection*.

If $(a_1 \ldots, a_k)$ are k linearly independent vectors in $\mathbb{R}^n$, let us determine the matrix P of the orthogonal projection onto the subspace of $\mathbb{R}^n$ spanned by $(a_1, \ldots, a_k)$.

Let A be the $n \times k$ matrix whose jth column consists of the coordinates of the vector a_j over the canonical basis $(e_1, \ldots, e_n)$.

The matrix P of the projection onto the subspace spanned by $(a_1 \ldots, a_k)$ is given by

$$P = A(A^\top A)^{-1} A^\top.$$

The reader should check that $P^2 = P$ and $P^\top = P$.

12.6   Dual Norms

If $(E, \| \ \|)$ and $(F, \| \ \|)$ are two normed vector spaces and if we let $\mathcal{L}(E; F)$ denote the set of all continuous (equivalently, bounded) linear maps from E to F, then, we can define the *operator norm* (or *subordinate norm*) $\| \ \|$ on $\mathcal{L}(E; F)$ as follows: for every $f \in \mathcal{L}(E; F)$,

$$\|f\| = \sup_{\substack{x \in E \\ x \neq 0}} \frac{\|f(x)\|}{\|x\|} = \sup_{\substack{x \in E \\ \|x\|=1}} \|f(x)\|.$$

In particular, if $F = \mathbb{C}$, then $\mathcal{L}(E; F) = E'$ is the *dual space* of E, and we get the operator norm denoted by $\| \ \|_*$ given by

$$\|f\|_* = \sup_{\substack{x \in E \\ \|x\|=1}} |f(x)|.$$

The norm $\| \ \|_*$ is called the *dual norm* of $\| \ \|$ on E'.

Let us now assume that E is a finite-dimensional Hermitian space, in which case $E' = E^*$.

Theorem 12.5 implies that for every linear form $f \in E^*$, there is a unique vector $y \in E$ so that

$$f(x) = \langle x, y \rangle, \quad \text{for all } x \in E,$$

and so we can write

$$\|f\|_* = \sup_{\substack{x \in E \\ \|x\|=1}} |\langle x, y \rangle|.$$

The above suggests defining a norm $\| \ \|^D$ on E.

Definition 12.8. If E is a finite-dimensional Hermitian space and $\|\ \|$ is any norm on E, for any $y \in E$ we let

$$\|y\|^D = \sup_{\substack{x \in E \\ \|x\|=1}} |\langle x, y \rangle|,$$

be the *dual norm* of $\|\ \|$ (on E). If E is a real Euclidean space, then the dual norm is defined by

$$\|y\|^D = \sup_{\substack{x \in E \\ \|x\|=1}} \langle x, y \rangle$$

for all $y \in E$.

Beware that $\|\ \|$ is generally *not* the Hermitian norm associated with the Hermitian inner product.

The dual norm shows up in convex programming; see Boyd and Vandenberghe [8], Chapters 2, 3, 6, 9.

The fact that $\| \ \|^D$ is a norm follows from the fact that $\| \ \|_*$ is a norm and can also be checked directly.

It is worth noting that the triangle inequality for $\| \ \|^D$ comes "for free," in the sense that it holds for any function $p \colon E \to \mathbb{R}$.

If $p \colon E \to \mathbb{R}$ is a function such that

(1) $p(x) \geq 0$ for all $x \in E$, and $p(x) = 0$ iff $x = 0$;

(2) $p(\lambda x) = |\lambda| p(x)$, for all $x \in E$ and all $\lambda \in \mathbb{C}$;

(3) p is continuous, in the sense that for some basis $(e_1, \ldots, e_n)$ of E, the function

$$(x_1, \ldots, x_n) \mapsto p(x_1 e_1 + \cdots + x_n e_n)$$

from $\mathbb{C}^n$ to $\mathbb{R}$ is continuous;

then we say that p is a *pre-norm*.

Obviously, every norm is a pre-norm, but a pre-norm may not satisfy the triangle inequality.

However, we just showed that the dual norm of any pre-norm is actually a norm.

It is not hard to show that

$$\|y\|^D = \sup_{\substack{x \in E \\ \|x\|=1}} |\langle x, y \rangle| = \sup_{\substack{x \in E \\ \|x\|=1}} \Re\langle x, y \rangle.$$

Proposition 12.19. *For all $x, y \in E$, we have*

$$|\langle x, y \rangle| \leq \|x\| \, \|y\|^D$$
$$|\langle x, y \rangle| \leq \|x\|^D \, \|y\| .$$

We can show that

$$\|y\|_1^D = \|y\|_\infty$$
$$\|y\|_\infty^D = \|y\|_1$$
$$\|y\|_2^D = \|y\|_2 .$$

Thus, the Euclidean norm is autodual. More generally, if $p, q \geq 1$ and $1/p + 1/q = 1$, we have

$$\|y\|_p^D = \|y\|_q .$$

It can also be shown that the dual of the spectral norm is the trace norm (or nuclear norm) from Section 16.5.

Theorem 12.20. *If E is a finite-dimensional Hermitian space, then for any norm $\| \ \|$ on E, we have*

$$\|y\|^{DD} = \|y\|$$

for all $y \in E$.

The proof makes use of the fact that a nonempty, closed, convex set has a supporting hyperplane through each of its boundary points, a result known as *Minkowski's lemma*.

This result is a consequence of the *Hahn–Banach theorem*; see Gallier [14].

More details on dual norms and unitarily invariant norms can be found in Horn and Johnson [19] (Chapters 5 and 7).