Chapter 8

Basics of Hermitian Geometry

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

In this chapter, we attempt to generalize the basic results of Euclidean geometry presented in Chapter 5 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 $\overline{z} = x-iy$, and the absolute value (or length, or modulus) of z as $|z|$. Recall that $|z|^2 = z\overline{z} = x^2 + y^2$.

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

$$\begin{aligned} f(u+v) &= f(u) + f(v), \\ f(\lambda u) &= \overline{\lambda} f(u), \end{aligned}$$

for all $u, v \in E$ and all $\lambda \in \mathbb{C}$. The set of all semilinear maps $f\colon E \to \mathbb{C}$ is denoted as $\overline{E}^*$.

It is trivially verified that $\overline{E}^*$ is a vector space over $\mathbb{C}$. It is not quite the dual space E^* of E.

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.

If E has finite dimension n, it is easy to see that $\overline{E}^*$ has the same dimension n

We can now define sesquilinear forms and Hermitian forms.

Definition 8.1.2 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

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

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

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

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 lemma.

Lemma 8.1.3 *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.

Lemma 8.1.3 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 $\varphi\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined such that

$$\varphi((x_1, y_1), (x_2, y_2)) = x_1y_2 - x_2y_1$$

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 8.

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

Definition 8.1.4 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 [**?**], define a pre-Hilbert space as what we define to be a Hermitian space.

We prefer following the terminology used in Schwartz [**?**] and Bourbaki [**?**].

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 (see Chapter **??**).

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

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

Lemma 8.1.5 *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, see Lang [**?**, **?**], or Schwartz [**?**, **?**]).

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 5 rather quickly, leaving out most proofs, except when they need to be seriously amended. This will be the case for the Cartan-Dieudonné theorem.

8.2 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 5.2.1).

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.

Lemma 5.2.2 and 5.2.3 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 $\varphi^l_u \colon E \to \mathbb{C}$ be the map defined such that

$$\varphi^l_u(v) = u \cdot v,$$

for all $v \in E$.

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

$$\varphi^r_v(u) = u \cdot v,$$

for all $u \in E$.

Since the Hermitian product is linear in its first argument u, the map φ^r_v is a linear form in E^*, and since it is semilinear in its second argument v, the map φ^l_u is a semilinear form in $\overline{E}^*$.

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

$$\flat^l(u) = \varphi^l_u, \quad \text{and} \quad \flat^r(v) = \varphi^r_v.$$

Lemma 8.2.1 *let E be a Hermitian space E.*

(1) The map $\flat^l\colon E \to \overline{E}^*$ defined such that

$$\flat^l(u) = \varphi^l_u,$$

is linear and injective.

(2) The map $\flat^r\colon E \to E^*$ defined such that

$$\flat^r(v) = \varphi^r_v,$$

is semilinear and injective.

When E is also of finite dimension, the maps $\flat^l\colon E \to \overline{E}^*$ and $\flat^r\colon E \to E^*$ are canonical isomorphisms.

The inverse of the isomorphism $\flat^l \colon E \to \overline{E}^*$ is denoted as $\sharp^l \colon \overline{E}^* \to E$, and the inverse of the isomorphism $\flat^r \colon E \to E^*$ is denoted as $\sharp^r \colon E^* \to E$.

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

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

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" $\flat^r \colon E \to E^*$ is not surjective when E has infinite dimension.

This result will be salvaged in Section **??** 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 $\flat^l \colon E \to \overline{E}^*$ is crucial to the existence of adjoint maps.

Indeed, Lemma 8.2.1 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

$$v \mapsto u \cdot f(v)$$

is clearly a semilinear form in $\overline{E}^*$, and by lemma 8.2.1, there is a unique vector in E denoted as $f^*(u)$, such that

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

for every $v \in E$.

The following lemma shows that the map f^* is linear.

Lemma 8.2.2 *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 lemma 8.2.2 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, lemma 8.2.1 can be used to show that any Hermitian space of finite dimension has an orthonormal basis. The proof is unchanged.

Lemma 8.2.3 *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!

Lemma 8.2.4 *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 lemma 5.2.7 also apply here, except that in the QR-decomposition, Q is a unitary matrix.

As a consequence of lemma 5.2.6 (or lemma 8.2.4), 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 $u = u_1e_1 + \cdots + u_ne_n$ and $v = v_1e_1 + \cdots + v_ne_n$, the Hermitian product $u \cdot v$ is expressed as

$$u \cdot v = (u_1e_1 + \cdots + u_ne_n) \cdot (v_1e_1 + \cdots + v_ne_n) = \sum_{i=1}^{n} u_i\overline{v_i},$$

and the norm $\|u\|$ as

$$\|u\| = \|u_1e_1 + \cdots + u_ne_n\| = \sqrt{\sum_{i=1}^{n} |u_i|^2}.$$

Lemma 5.2.8 also holds unchanged.

Lemma 8.2.5 *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 $F^{\perp\perp} = F$.*

Affine Hermitian spaces are defined just as affine Euclidean spaces, except that we modify definition 5.2.9 to require that the complex vector space E be a Hermitian space.

We denote as $\mathbb{E}^m_{\mathbb{C}}$ the Hermitian affine space obtained from the affine space $\mathbb{A}^m_{\mathbb{C}}$ by defining on the vector space $\mathbb{C}^m$ the standard Hermitian product

$$(x_1, \ldots, x_m) \cdot (y_1, \ldots, y_m) = x_1\overline{y_1} + \cdots + x_m\overline{y_m}.$$

The corresponding Hermitian norm is

$$\|(x_1, \ldots, x_m)\| = \sqrt{|x_1|^2 + \cdots + |x_m|^2}.$$

Lemma 7.2.2 also holds for Hermitian spaces, and the proof is the same.

Lemma 8.2.6 *Let E be a Hermitian space of finite dimension n, and let $f\colon E \to E$ be an isometry. For any subspace F of E, if $f(F) = F$, then $f(F^\perp) \subseteq F^\perp$ and $E = F \oplus F^\perp$.*

8.3 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 5.3 extend to Hermitian spaces, *except that orthogonal transformations are called unitary transformation*, but Lemma 5.3.2 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 5.3.1.

Definition 8.3.1 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 $u \in E$.

Lemma 5.3.2 can be salvaged by strengthening condition (2).

Lemma 8.3.2 *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.

8.4 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.

The Cartan-Dieudonné theorem can be generalized (Theorem 8.4.8), but this requires allowing *new types of hyperplane reflections* that we call Hermitian reflections.

After doing so, every isometry in $\mathbf{U}(n)$ can always be written as a composition of at most n Hermitian reflections (for $n \geq 2$).

Better yet, every rotation in $\mathbf{SU}(n)$ can be expressed as the composition of at most $2n-2$ (standard) hyperplane reflections!

This implies that every unitary transformation in $\mathbf{U}(n)$ is the composition of at most $2n - 1$ isometries, with at most one Hermitian reflection, the other isometries being (standard) hyperplane reflections.

The crucial Lemma 7.1.3 is false as is, and needs to be amended.

The QR-decomposition of arbitrary complex matrices in terms of Householder matrices can also be generalized, using a trick.

Definition 8.4.1 Given a complex $m \times n$ matrix A, the *transpose $A^\top$ of A* is the $n \times m$ matrix $A^\top = (a^\top_{i,\,j})$ defined such that

$$a^\top_{i,\,j} = a_{j,\,i}$$

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

$$b_{i,\,j} = \overline{a}_{i,\,j}$$

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

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

Lemma 8.4.2 *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 E, iff the rows of A form an orthonormal basis of E.

Lemma 5.4.1 also motivates the following definition.

Definition 8.4.3 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 lemma 8.3.2 (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 unitary and A is its matrix with respect to any orthonormal basis, we have $|D(A)| = 1$.

Definition 8.4.4 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 $|D(f)| = 1$, where $D(f)$ denotes the determinant of f. The isometries such that $D(f) = 1$ 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 $D(f) \neq 1$ are called *improper isometries, or improper unitary transformations, or flip transformations*.

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

Lemma 8.4.5 *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!

In order to generalize the Cartan-Dieudonné theorem and the QR-decomposition in terms of Householder transformations, we need to introduce new kinds of hyperplane reflections.

This is not really surprising, since in the Hermitian case, there are improper isometries whose determinant can be any unit complex number.

Hyperplane reflections are generalized as follows.

Definition 8.4.6 Let E be a Hermitian space of finite dimension. For any hyperplane H, for any nonnull vector w orthogonal to H, so that $E = H \oplus G$, where $G = \mathbb{C}w$, a *Hermitian reflection about H of angle θ* is a linear map of the form $\rho_{H,\theta}\colon E \to E$, defined such that

$$\rho_{H,\theta}(u) = p_H(u) + e^{i\theta} p_G(u),$$

for any unit complex number $e^{i\theta} \neq 1$ (i.e. $\theta \neq k2\pi$).

Since $u = p_H(u) + p_G(u)$, the Hermitian reflection $\rho_{H,\theta}$ is also expressed as

$$\rho_{H,\theta}(u) = u + (e^{i\theta} - 1) p_G(u),$$

or as

$$\rho_{H,\theta}(u) = u + (e^{i\theta} - 1) \frac{(u \cdot w)}{\|w\|^2} w.$$

Note that the case of a standard hyperplane reflection is obtained when $e^{i\theta} = -1$, i.e., $\theta = \pi$.

We leave as an easy exercise to check that $\rho_{H,\theta}$ is indeed an isometry, and that the inverse of $\rho_{H,\theta}$ is $\rho_{H,-\theta}$.

If we pick an orthonormal basis $(e_1, \ldots, e_n)$ such that $(e_1, \ldots, e_{n-1})$ is an orthonormal basis of H, the matrix of $\rho_{H,\theta}$ is

$$\begin{pmatrix} I_{n-1} & 0 \\ 0 & e^{i\theta} \end{pmatrix}$$

We now come to the main surprise. Given any two distinct vectors u and v such that $\|u\| = \|v\|$, *there isn't always a hyperplane reflection mapping u to v*, but this can be done using two Hermitian reflections!

Lemma 8.4.7 *Let E be any nontrivial Hermitian space.*

(1) For any two vectors $u, v \in E$ such that $u \neq v$ and $\|u\| = \|v\|$, if $u \cdot v = e^{i\theta}|u \cdot v|$, then the (usual) reflection s about the hyperplane orthogonal to the vector $v - e^{-i\theta}u$ is such that $s(u) = e^{i\theta}v$.

(2) For any nonnull vector $v \in E$, for any unit complex number $e^{i\theta} \neq 1$, there is a Hermitian reflection ρ_θ such that

$$\rho_\theta(v) = e^{i\theta}v.$$

As a consequence, for u and v as in (1), we have $\rho_{-\theta} \circ s(u) = v$.

Remarks: (1) If we use the vector $v + e^{-i\theta}u$ instead of $v - e^{-i\theta}u$, we get $s(u) = -e^{i\theta}v$.

(2) Certain authors, such as Kincaid and Cheney [**?**] and Ciarlet [**?**], use the vector $u + e^{i\theta}v$ instead of our vector $v + e^{-i\theta}u$. The effect of this choice is that they also get $s(u) = -e^{i\theta}v$.

(3) If $v = \|u\|\, e_1$, where e_1 is a basis vector, $u \cdot e_1 = a_1$, where a_1 is just the coefficient of u over the basis vector e_1.

Then, since $u \cdot e_1 = e^{i\theta}|a_1|$, the choice of the plus sign in the vector $\|u\|\, e_1 + e^{-i\theta}u$ has the effect that the coefficient of this vector over e_1 is $\|u\| + |a_1|$, and no cancellations takes place, which is preferable for numerical stability (we need to divide by the square norm of this vector).

The last part of Lemma 8.4.7 shows that the Cartan-Dieudonné is salvaged.

Actually, because we are over the complex field, a linear map always have (complex) eigenvalues, and we can get a slightly improved result.

Theorem 8.4.8 *Let E be a Hermitian space of dimension $n \geq 1$. Every isometry $f \in \mathbf{U}(E)$ is the composition $f = \rho_n \circ \rho_{n-1} \circ \cdots \circ \rho_1$ of n isometries ρ_j, where each ρ_j is either the identity or a Hermitian reflection (possibly a standard hyperplane reflection). When $n \geq 2$, the identity is the composition of any hyperplane reflection with itself.*

Proof. We prove by induction on n that there is an orthonormal basis of eigenvectors $(u_1, \ldots, u_n)$ of f such that

$$f(u_j) = e^{i\theta_j} u_j,$$

where $e^{i\theta_j}$ is an eigenvalue associated with u_j, for all j, $1 \leq j \leq n$.

Remarks. (1) Any isometry $f \in \mathbf{U}(n)$ can be express as $f = \rho_\theta \circ g$, where $g \in \mathbf{SU}(n)$ is a rotation, and ρ_θ is a Hermitian reflection.

As a consequence, there is a bijection between $S^1 \times \mathbf{SU}(n)$ and $\mathbf{U}(n)$, where S^1 is the unit circle (which corresponds to the group of complex numbers $e^{i\theta}$ of unit length). In fact, it is a homeomorphism.

(2) We abandoned the style of proof used in theorem 7.2.1, because in the Hermitian case, eigenvalues and eigenvectors always exist, and the proof is simpler that way (in the real case, an isometry may not have any real eigenvalues!).

The sacrifice is that the theorem yields no information on the number of hyperplane reflections. We shall rectify this situation shortly.

We will now reveal the beautiful trick (found in Mneimné and Testard [**?**]) that allows us to prove that every rotation in $\mathbf{SU}(n)$ is the composition of at most $2n-2$ (standard) hyperplane reflections.

For what follows, it is more convenient to denote the Hermitian reflection $\rho_{H,\theta}$ about a hyperplane H as $\rho_{u,\,\theta}$, where u is any vector orthogonal to H, and to denote a standard reflection about the hyperplane H as h_u (it is trivial that these do not depend on the choice of u in $H^{\perp}$).

Then, given any two distinct orthogonal vectors u, v such that $\|u\| = \|v\|$, consider the composition
$\rho_{v,\,-\theta} \circ \rho_{u,\,\theta}$.

The trick is that this composition can be expressed as two standard hyperplane reflections!

Lemma 8.4.9 *Let E be a nontrivial Hermitian space. For any two distinct orthogonal vectors u, v such that $\|u\| = \|v\|$, we have*

$$\begin{aligned}\rho_{v,\,-\theta} \circ \rho_{u,\,\theta} &= h_{v-\,u} \circ h_{v-\,e^{-i\theta}u} \\ &= h_{u+v} \circ h_{u+e^{i\theta}v}.\end{aligned}$$

Lemma 8.4.10 *Let E be a nontrivial Hermitian space, and let $(u_1, \ldots, u_n)$ be some orthonormal basis for E. For any $\theta_1, \ldots, \theta_n$ such that $\theta_1 + \cdots + \theta_n = 0$, if $f \in \mathbf{U}(n)$ is the isometry defined such that*

$$f(u_j) = e^{i\theta_j} u_j,$$

for all j, $1 \leq j \leq n$, then f is a rotation ($f \in \mathbf{SU}(n)$), and

$$\begin{aligned} f &= \rho_{u_n,\,\theta_n} \circ \cdots \circ \rho_{u_1,\,\theta_1} \\ &= \rho_{u_n,\,-(\theta_1+\cdots+\theta_{n-1})} \circ \rho_{u_{n-1},\,\theta_1+\cdots+\theta_{n-1}} \circ \cdots \circ \rho_{u_2,\,-\theta_1} \circ \rho_{u_1,\,\theta_1} \\ &= h_{u_n-\,u_{n-1}} \circ h_{u_n-\,e^{-i(\theta_1+\cdots+\theta_{n-1})}u_{n-1}} \circ \cdots \circ h_{u_2-\,u_1} \circ h_{u_2-\,e^{-i\theta_1}u_1} \\ &= h_{u_{n-1}+u_n} \circ h_{u_{n-1}+\,e^{i(\theta_1+\cdots+\theta_{n-1})}u_n} \circ \cdots \circ h_{u_1+u_2} \circ h_{u_1+\,e^{i\theta_1}u_2}.\end{aligned}$$

We finally get our *improved version* of the Cartan-Dieudonné theorem.

Theorem 8.4.11 *Let E be a Hermitian space of dimension $n \geq 1$. Every rotation $f \in \mathbf{SU}(E)$ different from the identity is the composition of at most $2n-2$ hyperplane reflections. Every isometry $f \in \mathbf{U}(E)$ different from the identity is the composition of at most $2n-1$ isometries, all hyperplane reflections, except for possibly one Hermitian reflection. When $n \geq 2$, the identity is the composition of any reflection with itself.*

As a corollary of Theorem 8.4.11, the following interesting result can be shown (this is not hard, do it!).

First, recall that a linear map $f \colon E \to E$ is *self-adjoint* (or *Hermitian*) iff $f = f^*$.

Then, the subgroup of $\mathbf{U}(n)$ generated by the Hermitian isometries is equal to the group

$$\mathbf{SU}(n)^{\pm} = \{f \in \mathbf{U}(n) \mid \det(f) = \pm 1\}.$$

Equivalently, $\mathbf{SU}(n)^{\pm}$ is equal to the subgroup of $\mathbf{U}(n)$ generated by the hyperplane reflections.

This problem had been left open by Dieudonné in [**?**]. Evidently, it was settled since the publication of the third edition of the book [**?**].

Inspection of the proof of Lemma 7.2.4 reveals that this lemma also holds for Hermitian spaces.

Thus, when $n \geq 3$, the composition of any two hyperplane reflections is equal to the composition of two flips.

Theorem 8.4.12 *Let E be a Hermitan space of dimension $n \geq 3$. Every rotation $f \in \mathbf{SU}(E)$ is the composition of an even number of flips $f = f_{2k} \circ \cdots \circ f_1$, where $k \leq n-1$. Furthermore, if $u \neq 0$ is invariant under f (i.e. $u \in \mathrm{Ker}\,(f - \mathrm{id})$), we can pick the last flip f_{2k} such that $u \in F_{2k}^{\perp}$, where F_{2k} is the subspace of dimension $n-2$ determining f_{2k}.*

We now show that the QR-decomposition in terms of (complex) Householder matrices holds for complex matrices.

We need the version of Lemma 8.4.7 and a trick at the end of the argument, but the proof is basically unchanged.

Lemma 8.4.13 *Let E be a nontrivial Hermitian space of dimension n. Given any orthonormal basis $(e_1, \ldots, e_n)$, for any n-tuple of vectors $(v_1, \ldots, v_n)$, there is a sequence of n isometries $h_1, \ldots, h_n$, such that h_i is a hyperplane reflection or the identity, and if $(r_1, \ldots, r_n)$ are the vectors given by*

$$r_j = h_n \circ \cdots \circ h_2 \circ h_1(v_j),$$

then every r_j is a linear combination of the vectors $(e_1, \ldots, e_j)$, $(1 \leq j \leq n)$. Equivalently, the matrix R whose columns are the components of the r_j over the basis $(e_1, \ldots, e_n)$ is an upper triangular matrix. Furthermore, if we allow one more isometry h_{n+1} of the form

$$h_{n+1} = \rho_{e_n, \varphi_n} \circ \cdots \circ \rho_{e_1, \varphi_1}$$

after $h_1, \ldots, h_n$, we can ensure that the diagonal entries of R are nonnegative.

Remark: For numerical stability, it may be preferable to use $w_{k+1} = r_{k+1,k+1}\, e_{k+1} + e^{-i\theta_{k+1}} u''_{k+1}$ instead of $w_{k+1} = r_{k+1,k+1}\, e_{k+1} - e^{-i\theta_{k+1}} u''_{k+1}$. The effect of that choice is that the diagonal entries in R will be of the form $-e^{i\theta_j} r_{j,\,j} = e^{i(\theta_j+\pi)} r_{j,\,j}$.

Of course, we can make these entries nonegative by applying

$$h_{n+1} = \rho_{e_n,\,\pi-\theta_n} \circ \cdots \circ \rho_{e_1,\pi-\theta_1}$$

after h_n.

As in the Euclidean case, Lemma 8.4.13 immediately implies the QR-decomposition for arbitrary complex $n \times n$-matrices, where Q is now unitary (see Kincaid and Cheney [**?**], or Ciarlet [**?**]).

Lemma 8.4.14 *For every complex $n \times n$-matrix A, there is a sequence $H_1, \ldots, H_n$ of matrices, where each H_i is either a Householder matrix or the identity, and an upper triangular matrix R, such that*

$$R = H_n \cdots H_2 H_1 A.$$

As a corollary, there is a pair of matrices Q, R, where Q is unitary and R is upper triangular, such that $A = QR$ (a QR-decomposition of A). Furthermore, R can be chosen so that its diagonal entries are nonnegative.

As in the Euclidean case, the QR-decomposition has applications to least squares problems. It is also possible to convert any complex matrix to bidiagonal form.