10

Basics of Hermitian Geometry

10.1 Sesquilinear and Hermitian Forms, Pre-Hilbert Spaces and Hermitian Spaces

In this chapter we generalize the basic results of Euclidean geometry presented in Chapter 6 to vector spaces over the complex numbers. Such a generalization is inevitable, and not simply a luxury. For example, linear maps may not have real eigenvalues, but they always have complex eigenvalues. Furthermore, some very important classes of linear maps can be diagonalized if they are extended to the complexification of a real vector space. This is the case for orthogonal matrices, and, more generally, normal matrices. Also, complex vector spaces are often the natural framework in physics or engineering, and they are more convenient for dealing with Fourier series. However, 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 $z = x+iy$ by $\overline{z} = x-iy$, and the absolute value (or length, or modulus) of z by $|z|$. 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, which means that $\varphi(u, \mu v) = \overline{\mu}\varphi(u, v)$, as opposed to $\varphi(u, \mu v) = \mu\varphi(u, v)$. 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. Thus, when generalizing a result from the real case of a Euclidean space to the complex case, we always have to check very carefully that our proofs do not rely on linearity in the second argument. Otherwise, we need to revise our proofs, and sometimes the result is simply wrong!

Before defining the natural generalization of an inner product, it is convenient to define semilinear maps.

Definition 10.1.1 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 by $\overline{E}^*$.

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 (if $(e_1, \ldots, e_n)$ is a basis for E, check that the semilinear maps $(\overline{e_1}, \ldots, \overline{e_n})$ defined such that

$$\overline{e_i}\left(\sum_{j=1}^{n} \lambda_j e_j\right) = \overline{\lambda_i},$$

form a basis of $\overline{E}^*$.)

We can now define sesquilinear forms and Hermitian forms.

Definition 10.1.2 Given a complex vector space E, a function $\varphi\colon E \times E \to \mathbb{C}$ is a *sesquilinear form* if 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 $\varphi: E \times E \to \mathbb{C}$ is a *Hermitian form* if 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 $\varphi: E \times E \to \mathbb{C}$ is sesquilinear, we have

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

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

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

Note that restricted to real coefficients, a sesquilinear form is bilinear (we sometimes say $\mathbb{R}$-bilinear). The function $\Phi: E \to \mathbb{C}$ 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 10.1.3 *Given a complex vector space E, the following properties hold:*

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

(2) If $\varphi: E \times E \to \mathbb{C}$ 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

$$2\varphi(u, v) = (1+i)(\varphi(u, u)+\varphi(v, v))-\varphi(u-v, u-v)-i\varphi(u-iv, u-iv).$$

These are called polarization identities.

Proof. (1) If φ is a Hermitian form, then

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

implies that

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

and thus $\varphi(u, u) \in \mathbb{R}$. If φ is sesquilinear and $\varphi(u, u) \in \mathbb{R}$ for all $u \in E$, then

$$\varphi(u + v, u + v) = \varphi(u, u) + \varphi(u, v) + \varphi(v, u) + \varphi(v, v),$$

which proves that

$$\varphi(u, v) + \varphi(v, u) = \alpha,$$

where α is real, and changing u to iu, we have

$$i(\varphi(u, v) - \varphi(v, u)) = \beta,$$

where β is real, and thus

$$\varphi(u, v) = \frac{\alpha - i\beta}{2} \quad \text{and} \quad \varphi(v, u) = \frac{\alpha + i\beta}{2},$$

proving that φ is Hermitian.

(2) These identities are verified by expanding the right-hand side, and we leave them as an exercise. $\square$

Lemma 10.1.3 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

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

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 10.1.4 Given a complex vector space E, a Hermitian form $\varphi\colon E \times E \to \mathbb{C}$ is *positive* if $\varphi(u, u) \geq 0$ for all $u \in E$, and *positive definite* if $\varphi(u, u) > 0$ 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 [109], define a pre-Hilbert space as what we define as a Hermitian space. We prefer following the terminology used in Schwartz [149] and Bourbaki [21]. 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 $\varphi(u, v)$ by

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

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

Example 10.1 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 10.2 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 $\varphi: l^2 \times l^2 \to \mathbb{C}$ 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 26).

Example 10.3 Let $\mathcal{C}_{\text{piece}}[a,b]$ be the set of piecewise bounded continuous functions $f: [a,b] \to \mathbb{C}$ 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, $\mathcal{C}_{\text{piece}}[a,b]$ is only a pre-Hilbert space.

Example 10.4 Let $\mathcal{C}[-\pi,\pi]$ be the set of complex-valued continuous functions $f: [-\pi,\pi] \to \mathbb{C}$ 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 10.1.5 *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 Minkowski 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 $\lambda > 0$.

Proof. For all $u, v \in E$ and all $\mu \in \mathbb{C}$, we have observed that

$$\varphi(u+\mu v, u+\mu v) = \varphi(u,u) + 2\Re(\overline{\mu}\varphi(u,v)) + |\mu|^2\varphi(v,v).$$

Let $\varphi(u, v) = \rho e^{i\theta}$, where $|\varphi(u, v)| = \rho$ ($\rho \geq 0$). Let $F: \mathbb{R} \to \mathbb{R}$ be the function defined such that

$$F(t) = \Phi(u + te^{i\theta} v),$$

for all $t \in \mathbb{R}$. The above shows that

$$F(t) = \varphi(u, u) + 2t|\varphi(u, v)| + t^2 \varphi(v, v) = \Phi(u) + 2t|\varphi(u, v)| + t^2 \Phi(v).$$

Since φ is assumed to be positive, we have $F(t) \geq 0$ for all $t \in \mathbb{R}$. If $\Phi(v) = 0$, we must have $\varphi(u, v) = 0$, since otherwise, $F(t)$ could be made negative by choosing t negative and small enough. If $\Phi(v) > 0$, in order for $F(t)$ to be nonnegative, the equation

$$\Phi(u) + 2t|\varphi(u, v)| + t^2 \Phi(v) = 0$$

must not have distinct real roots, which is equivalent to

$$|\varphi(u, v)|^2 \leq \Phi(u)\Phi(v).$$

Taking the square root on both sides yields the Cauchy–Schwarz inequality.

For the second part of the claim, if φ is positive definite, we argue as follows. If u and v are linearly dependent, it is immediately verified that we get an equality. Conversely, if

$$|\varphi(u, v)|^2 = \Phi(u)\Phi(v),$$

then the equation

$$\Phi(u) + 2t|\varphi(u, v)| + t^2 \Phi(v) = 0$$

has a double root t_0, and thus

$$\Phi(u + t_0 e^{i\theta} v) = 0.$$

Since φ is positive definite, we must have

$$u + t_0 e^{i\theta} v = 0,$$

which shows that u and v are linearly dependent.

If we square the Minkowski inequality, we get

$$\Phi(u + v) \leq \Phi(u) + \Phi(v) + 2\sqrt{\Phi(u)}\sqrt{\Phi(v)}.$$

However, we observed earlier that

$$\Phi(u + v) = \Phi(u) + \Phi(v) + 2\Re(\varphi(u, v)).$$

Thus, it is enough to prove that

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

but this follows from the Cauchy–Schwarz inequality

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

and the fact that $\Re z \leq |z|$.

If φ is positive definite and u and v are linearly dependent, it is immediately verified that we get an equality. Conversely, if equality holds in the Minkowski inequality, we must have

$$\Re(\varphi(u,v)) = \sqrt{\Phi(u)}\sqrt{\Phi(v)},$$

which implies that

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

since otherwise, by the Cauchy–Schwarz inequality, we would have

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

Thus, equality holds in the Cauchy–Schwarz inequality, and

$$\Re(\varphi(u,v)) = |\varphi(u,v)|.$$

But then, we proved in the Cauchy–Schwarz case that u and v are linearly dependent. Since we also just proved that $\varphi(u,v)$ is real and nonnegative, the coefficient of proportionality between u and v is indeed nonnegative. $\square$

As in the Euclidean case, if $\langle E, \varphi\rangle$ is a Hermitian space, the Minkowski 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 $\sqrt{\Phi(u)}$ by $\|u\|$, 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 [109, 110], Dixmier [50], or Schwartz [149, 150]. 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 *seminorm*. 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 6 rather quickly, leaving out most proofs, except when they need to be seriously amended. This will be the case for the Cartan–Dieudonné theorem.

10.2 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 by $u \cdot v$ 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 6.2.1).

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 6.2.2 and 6.2.3 hold without any changes. It is easy to show that

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

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

$$\varphi_u^l(v) = u\cdot v,$$

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

$$\varphi_v^r(u) = u\cdot v,$$

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 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

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

Lemma 10.2.1 *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.*

Proof. (1) That $\flat^l\colon E \to \overline{E}^*$ is a linear map follows immediately from the fact that the Hermitian product is linear in its first argument. If $\varphi^l_u = \varphi^l_v$, then $\varphi^l_u(w) = \varphi^l_v(w)$ for all $w \in E$, which by definition of φ^l_u means that

$$u \cdot w = v \cdot w$$

for all $w \in E$, which by linearity on the left is equivalent to

$$(v - u) \cdot w = 0$$

for all $w \in E$, which implies that $u = v$, since the Hermitian product is positive definite. Thus, $\flat^l\colon E \to \overline{E}^*$ is injective. Finally, when E is of finite dimension n, $\overline{E}^*$ is also of dimension n, and then $\flat^l\colon E \to \overline{E}^*$ is bijective.

The proof of (2) is essentially the same as the proof of (1), except that the Hermitian product is semilinear in its second argument. $\square$

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

As a corollary of the isomorphism $\flat^r\colon E \to E^*$, if E is a Hermitian space of finite dimension, then 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 can 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 $\flat^l\colon E \to \overline{E}^*$ is crucial to the existence of adjoint maps. Indeed, Lemma 10.2.1 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 10.2.1, there is a unique vector in E denoted by $f^*(u)$ 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 10.2.2 *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).

Proof. Careful inspection of the proof of lemma 6.2.5 reveals that it applies unchanged. The only potential problem is in proving that $f^*(\lambda u) = \lambda f^*(u)$, but everything takes place in the first argument of the Hermitian product, and there, we have linearity. $\square$

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 by $\langle -, - \rangle_1$ and the Hermitian product on F is denoted by $\langle -, - \rangle_2$, given any linear map $f\colon E \to F$, it is immediately verified that the proof of Lemma 10.2.2 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 10.2.1 can be used to show that any Hermitian space of finite dimension has an orthonormal basis. The proof is unchanged.

Lemma 10.2.3 *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 10.2.4 *Given a 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.*

Remark: The remarks made after Lemma 6.2.7 also apply here, except that in the QR-decomposition, Q is a unitary matrix.

As a consequence of Lemma 6.2.6 (or Lemma 10.2.4), 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 6.2.8 also holds unchanged.

Lemma 10.2.5 *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 6.2.9 to require that the complex vector space $\overrightarrow{E}$ be a Hermitian space. We denote by $\mathbb{E}^m_{\mathbb{C}}$ 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 10.2.6 *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 $E = F \oplus F^{\perp}$.*

10.3 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 6.3 extend to Hermitian spaces, except that orthogonal transformations are called unitary transformation, but Lemma 6.3.2 extends only 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 6.3.1.

Definition 10.3.1 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*, if it is linear and

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

for all $u \in E$.

Lemma 6.3.2 can be salvaged by strengthening condition (2).

Lemma 10.3.2 *Given any two nontrivial Hermitian spaces 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.

Proof. The proof that (2) implies (3) given in Lemma 6.3.2 needs to be revised as follows. We use the polarization identity

$$2\varphi(u, v) = (1 + i)(\|u\|^2 + \|v\|^2) - \|u - v\|^2 - i\|u - iv\|^2.$$

Since $f(iv) = if(v)$, we get $f(0) = 0$ by setting $v = 0$, so the function f preserves distance and norm, and we get

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

which shows that f preserves the Hermitian inner product, as desired. The rest of the proof is unchanged. $\square$

Remarks:

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

$$\|f(v) - f(u)\| = \|v - u\| \quad \text{for all } u, v \in E \text{ and } f(0) = 0 \qquad (2')$$

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 $z \mapsto \overline{z}$ satisfies $(2')$ 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 be 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.

10.4 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. As an immediate corollary of the Gram–Schmidt orthonormalization procedure, we obtain the QR-decomposition for invertible matrices. In the Hermitian framework, the matrix of the adjoint of a linear map is not given by the transpose of the original matrix, but by its conjugate.

Definition 10.4.1 Given a complex $m \times n$ matrix A, the *transpose* $A^\top$ *of* A is the $n \times m$ matrix $A^\top = \left(a^\top_{i,\,j}\right)$ 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

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

Lemma 10.4.2 *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 identities

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

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.

Proof. (1) The proof is identical to that of Lemma 6.4.1 (1).

(2) If $(e_1, \ldots, e_n)$ is an orthonormal basis for E, let $A = (a_{i,j})$ be the matrix of f, and let $B = (b_{i,j})$ be the matrix of f^*. Since f^* is characterized by

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

for all $u, v \in E$, using the fact that if $w = w_1 e_1 + \cdots + w_n e_n$, we have $w_k = w \cdot e_k$, for all k, $1 \leq k \leq n$; letting $u = e_i$ and $v = e_j$, we get

$$b_{j,i} = f^*(e_i) \cdot e_j = e_i \cdot f(e_j) = \overline{f(e_j) \cdot e_i} = \overline{a_{i,j}},$$

for all i, j, $1 \leq i, j \leq n$. Thus, $B = A^*$. Now, if X and Y are arbitrary matrices over the basis $(e_1, \ldots, e_n)$, denoting as usual the jth column of X by X_j, and similarly for Y, a simple calculation shows that

$$Y^* X = (X_j \cdot Y_i)_{1 \leq i, j \leq n}.$$

Then it is immediately verified that if $X = Y = A$,

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

iff the column vectors $(A_1, \ldots, A_n)$ form an orthonormal basis. Thus, from (1), we see that (2) is clear. $\square$

Lemma 6.4.1 shows that the inverse of an isometry f is its adjoint f^*. Lemma 6.4.1 also motivates the following definition.

Definition 10.4.3 A complex $n \times n$ matrix is a *unitary matrix* if

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

Remarks:

(1) 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 10.3.2 (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.

(2) If f is unitary and A is its matrix with respect to any orthonormal basis, the characteristic polynomial $D(A - \lambda I)$ of A is a polynomial with complex coefficients, and thus it has n (complex) roots (counting multiplicities). If u is an eigenvector of f for λ, then from $f(u) = \lambda u$ and the fact that f is an isometry we get

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

which shows that $|\lambda| = 1$. Since the determinant $D(A)$ of f is the product of the eigenvalues of f, we have $|D(A)| = 1$. It is clear that the isometries of a Hermitian space of dimension n form a group, and that the isometries of determinant $+1$ form a subgroup.

This leads to the following definition.

Definition 10.4.4 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 by $\mathbf{U}(E)$, 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 by $\mathbf{SU}(E)$, 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*.

A very important example of unitary matrices is provided by Fourier matrices (up to a factor of $\sqrt{n}$), matrices that arise in the various versions of the discrete Fourier transform. For more on this topic, see the problems, and Strang [165, 168].

Now that we have the definition of a unitary matrix, we can explain how the Gram–Schmidt orthonormalization procedure immediately yields the QR-decomposition for matrices.

Lemma 10.4.5 *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!

Due to space limitations, we will not study the isometries of a Hermitian space in this chapter. However, the reader will find such a study in the supplements on the web site (see web page, Chapter 25).

10.5 Problems

Problem 10.1 Given a complex vector space E of finite dimension n, prove that $\overline{E}^*$ also has dimension n.
Hint. If $(e_1, \ldots, e_n)$ is a basis for E, check that the semilinear maps $\overline{e_i}$ defined such that

$$\overline{e_i}\left(\sum_{j=1}^{n} \lambda_j e_j\right) = \overline{\lambda_i}$$

form a basis of $\overline{E}^*$.

Problem 10.2 Prove the polarization identities in Lemma 10.1.3 (2).

Problem 10.3 Given a Hermitian space E, for any orthonormal basis $(e_1, \ldots, e_n)$, if X and Y are arbitrary matrices over the basis $(e_1, \ldots, e_n)$, denoting as usual the jth column of X by X_j, and similarly for Y, prove that

$$Y^* X = (X_j \cdot Y_i)_{1 \leq i,j \leq n}.$$

Then prove that

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

iff the column vectors $(A_1, \ldots, A_n)$ form an orthonormal basis.

Problem 10.4 Given a Hermitian space E, prove that if f is an isometry, then f maps any orthonormal basis of E to an orthonormal basis.

Problem 10.5 Given p vectors $(u_1, \ldots, u_p)$ in a Hermitian space E of dimension $n \geq p$, the *Gram determinant (or Gramian)* of the vectors $(u_1, \ldots, u_p)$ is the determinant

$$\mathrm{Gram}(u_1, \ldots, u_p) = \begin{vmatrix} \|u_1\|^2 & \langle u_1, u_2 \rangle & \ldots & \langle u_1, u_p \rangle \\ \langle u_2, u_1 \rangle & \|u_2\|^2 & \ldots & \langle u_2, u_p \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle u_p, u_1 \rangle & \langle u_p, u_2 \rangle & \ldots & \|u_p\|^2 \end{vmatrix}.$$

(1) Prove that

$$\mathrm{Gram}(u_1, \ldots, u_n) = \lambda_E(u_1, \ldots, u_n)^2.$$

Hint. By Problem 10.3, if $(e_1, \ldots, e_n)$ is an orthonormal basis of E and A is the matrix of the vectors $(u_1, \ldots, u_n)$ over this basis, then

$$\det(A)^2 = \det(A^* A) = \det(A_i \cdot A_j),$$

where A_i denotes the ith column of the matrix A, and $(A_i \cdot A_j)$ denotes the $n \times n$ matrix with entries $A_i \cdot A_j$.

Problem 10.6 Let F_n be the symmetric $n \times n$ matrix (with complex coefficients)

$$F_n = \left(e^{i2\pi kl/n}\right)_{\substack{0 \le k \le n-1 \\ 0 \le l \le n-1}},$$

assuming that we index the entries in F_n over $[0, 1, \ldots, n-1] \times [0, 1, \ldots, n-1]$, the standard kth row now being indexed by $k-1$ and the standard lth column now being indexed by $l-1$. The matrix F_n is called a *Fourier matrix*.

(1) Letting $\overline{F_n} = \left(e^{-i2\pi kl/n}\right)_{\substack{0 \le k \le n-1 \\ 0 \le l \le n-1}}$ be the conjugate of F_n, prove that

$$F_n \overline{F_n} = \overline{F_n} F_n = n\, I_n.$$

The above shows that $F_n/\sqrt{n}$ is unitary.

(2) Define the *discrete Fourier transform* $\widehat{f}$ of a sequence $f = (f_0, \ldots, f_{n-1}) \in \mathbb{C}^n$ as

$$\widehat{f} = \overline{F_n} f.$$

Define the *inverse discrete Fourier transform* (taking c back to f) as

$$\overline{\widehat{c}} = F_n\, c,$$

where $c = (c_0, \ldots, c_{n-1}) \in \mathbb{C}^n$. Define the *circular shift matrix* S_n *(of order* n*)* as the matrix

$$S_n = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

consisting of cyclic permutations of its first column. For any sequence $f = (f_0, \ldots, f_{n-1}) \in \mathbb{C}^n$, we define the *circulant matrix* $H(f)$ as

$$H(f) = \sum_{j=0}^{n-1} f_j S_n^j,$$

where $S_n^0 = I_n$, as usual.

Prove that

$$H(f) F_n = F_n \widehat{f}.$$

The above shows that the columns of the Fourier matrix F_n are the eigenvectors of the circulant matrix $H(f)$, and that the eigenvalue associated

with the lth eigenvector is $(\widehat{f})_l$, the lth component of the Fourier transform $\widehat{f}$ of f (counting from 0).
Hint. Prove that

$$S_n F_n = F_n \operatorname{diag}(v^1)$$

where $\operatorname{diag}(v^1)$ is the diagonal matrix with the following entries on the diagonal:

$$v^1 = \left(1, e^{-i2\pi/n}, \ldots, e^{-ik2\pi/n}, \ldots, e^{-i(n-1)2\pi/n}\right).$$

(3) If the sequence $f = (f_0, \ldots, f_{n-1})$ is even, which means that $f_{-j} = f_j$ for all $j \in \mathbb{Z}$ (viewed as a periodic sequence), or equivalently that $f_{n-j} = f_j$ for all j, $0 \leq j \leq n-1$, prove that the Fourier transform $\widehat{f}$ is expressed as

$$\widehat{f}(k) = \sum_{j=0}^{n-1} f_j \cos\left(2\pi jk/n\right),$$

and that the inverse Fourier transform (taking c back to f) is expressed as

$$\overline{\overline{c}}(k) = \sum_{j=0}^{n-1} c_j \cos\left(2\pi jk/n\right),$$

for every k, $0 \leq k \leq n-1$.

(4) Define the *convolution* $f \star g$ of two sequences $f = (f_0, \ldots, f_{n-1})$ and $g = (g_0, \ldots, g_{n-1})$ as

$$f \star g = H(f)\, g,$$

viewing f and g as column vectors.

Prove the *(circular) convolution rule*

$$\widehat{f \star g} = \widehat{f}\,\widehat{g},$$

where the multiplication on the right-hand side is just the inner product of the vectors $\widehat{f}$ and $\widehat{g}$.

Problem 10.7 Let $\varphi\colon E \times E \to \mathbb{C}$ be a sesquilinear form on a complex vector space E of finite dimension n. Given any basis $(e_1, \ldots, e_n)$ of E, let $A = (\alpha_{i\,j})$ be the matrix defined such that

$$\alpha_{i\,j} = \varphi(e_i, e_j),$$

$1 \leq i, j \leq n$. We call A *the matrix of* φ *w.r.t. the basis* $(e_1, \ldots, e_n)$.

(a) For any two vectors x and y, if X and Y denote the column vectors of coordinates of x and y w.r.t. the basis $(e_1, \ldots, e_n)$, prove that

$$\varphi(x, y) = X^\top A\overline{Y}.$$

(b) Recall that A is a *Hermitian* matrix if $A = A^* = \overline{A^\top}$. Prove that φ is Hermitian iff A is a Hermitian matrix. When is it true that

$$\varphi(x, y) = Y^* AX?$$

(c) If $(f_1, \ldots, f_n)$ is another basis of E and P is the change of basis matrix from $(e_1, \ldots, e_n)$ to $(f_1, \ldots, f_n)$, prove that the matrix of φ w.r.t. the basis $(f_1, \ldots, f_n)$ is

$$P^\top A\overline{P}.$$

The common rank of all matrices representing φ is called the *rank* of φ.

Problem 10.8 Let $\varphi\colon E \times E \to \mathbb{C}$ be a Hermitian form on a complex vector space E of finite dimension n. Two vectors x and y are said to be *conjugate w.r.t.* φ if $\varphi(x, y) = 0$. The main purpose of this problem is to prove that there is a basis of vectors that are pairwise conjugate w.r.t. φ.

(a) Prove that if $\varphi(x, x) = 0$ for all $x \in E$, then φ is identically null on E. For this, compute $\varphi(ix + y, ix + y)$ and $i\varphi(x + y, x + y)$, and conclude that $\varphi(x, y) = 0$.

Otherwise, we can assume that there is some vector $x \in E$ such that $\varphi(x, x) \neq 0$. Use induction to prove that there is a basis of vectors that are pairwise conjugate w.r.t. φ.

For the induction step, proceed as follows. Let $(e_1, e_2, \ldots, e_n)$ be a basis of E, with $\varphi(e_1, e_1) \neq 0$. Prove that there are scalars $\lambda_2, \ldots, \lambda_n$ such that each of the vectors

$$v_i = e_i + \lambda_i e_1,$$

is conjugate to e_1 w.r.t. φ, where $2 \leq i \leq n$, and that $(e_1, v_2, \ldots, v_n)$ is a basis.

(b) Let $(e_1, \ldots, e_n)$ be a basis of vectors that are pairwise conjugate w.r.t. φ, and assume that they are ordered such that

$$\varphi(e_i, e_i) = \begin{cases} \theta_i \neq 0 & \text{if } 1 \leq i \leq r, \\ 0 & \text{if } r + 1 \leq i \leq n, \end{cases}$$

where r is the rank of φ. Show that the matrix of φ w.r.t. $(e_1, \ldots, e_n)$ is a diagonal matrix, and that

$$\varphi(x, y) = \sum_{i=1}^{r} \theta_i x_i \overline{y_i},$$

where $x = \sum_{i=1}^n x_i e_i$ and $y = \sum_{i=1}^n y_i e_i$.

Prove that for every Hermitian matrix A there is an invertible matrix P such that

$$P^\top A\overline{P} = D,$$

where D is a diagonal matrix.

(c) Prove that there is an integer p, $0 \leq p \leq r$ (where r is the rank of φ), such that $\varphi(u_i, u_i) > 0$ for exactly p vectors of every basis $(u_1, \ldots, u_n)$ of vectors that are pairwise conjugate w.r.t. φ (*Sylvester's inertia theorem*).

Proceed as follows. Assume that in the basis $(u_1, \ldots, u_n)$, for any $x \in E$, we have

$$\varphi(x, x) = \alpha_1 |x_1|^2 + \cdots + \alpha_p |x_p|^2 - \alpha_{p+1} |x_{p+1}|^2 - \cdots - \alpha_r |x_r|^2,$$

where $x = \sum_{i=1}^n x_i u_i$, and that in the basis $(v_1, \ldots, v_n)$, for any $x \in E$, we have

$$\varphi(x, x) = \beta_1 |y_1|^2 + \cdots + \beta_q |y_q|^2 - \beta_{q+1} |y_{q+1}|^2 - \cdots - \beta_r |y_r|^2,$$

where $x = \sum_{i=1}^n y_i v_i$, with $\alpha_i > 0$, $\beta_i > 0$, $1 \leq i \leq r$.

Assume that $p > q$ and derive a contradiction. First, consider x in the subspace F spanned by

$$(u_1, \ldots, u_p, u_{r+1}, \ldots, u_n),$$

and observe that $\varphi(x, x) \geq 0$ if $x \neq 0$. Next, consider x in the subspace G spanned by

$$(v_{q+1}, \ldots, v_r),$$

and observe that $\varphi(x, x) < 0$ if $x \neq 0$. Prove that $F \cap G$ is nontrivial (i.e., contains some nonnull vector), and derive a contradiction. This implies that $p \leq q$. Finish the proof.

The pair $(p, r - p)$ is called the *signature* of φ.

(d) A Hermitian form φ is *definite* if for every $x \in E$, if $\varphi(x, x) = 0$, then $x = 0$.

Prove that a Hermitian form is definite iff its signature is either $(n, 0)$ or $(0, n)$. In other words, a Hermitian definite form has rank n and is either positive or negative.

(e) The *kernel* of a Hermitian form φ is the subspace consisting of the vectors that are conjugate to all vectors in E. We say that a Hermitian form φ is *nondegenerate* if its kernel is trivial (i.e., reduced to $\{0\}$).

Prove that a Hermitian form φ is nondegenerate iff its rank is n, the dimension of E. Is a definite Hermitian form φ nondegenerate? What about the converse?

Prove that if φ is nondegenerate, then there is a basis of vectors that are pairwise conjugate w.r.t. φ and such that φ is represented by the matrix

$$\begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix},$$

where (p, q) is the signature of φ.

(f) Given a nondegenerate Hermitian form φ on E, prove that for every linear map $f \colon E \to E$, there is a unique linear map $f^* \colon E \to E$ such that

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

for all $u, v \in E$. The map f^* is called the *adjoint of f (w.r.t. to φ)*. Given any basis $(u_1, \ldots, u_n)$, if Ω is the matrix representing φ and A is the matrix representing f, prove that f^* is represented by $(\Omega^\top)^{-1} A^* \Omega^\top$.

Prove that Lemma 10.2.1 also holds, i.e., the maps $\flat^l\colon E \to \overline{E}^*$ and $\flat^r\colon E \to E^*$ are canonical isomorphisms.

A linear map $f\colon E \to E$ is an *isometry w.r.t.* φ if

$$\varphi(f(x),\, f(y)) = \varphi(x,\, y)$$

for all $x, y \in E$. Prove that a linear map f is an isometry w.r.t. φ iff

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

Prove that the set of isometries w.r.t. φ is a group. This group is denoted by $\mathbf{U}(\varphi)$, and its subgroup consisting of isometries having determinant $+1$ by $\mathbf{SU}(\varphi)$. Given any basis of E, if Ω is the matrix representing φ and A is the matrix representing f, prove that $f \in \mathbf{U}(\varphi)$ iff

$$A^*\Omega^\top A = \Omega^\top.$$

Given another nondegenerate Hermitian form ψ on E, we say that φ and ψ are *equivalent* if there is a bijective linear map $h\colon E \to E$ such that

$$\psi(x,\, y) = \varphi(h(x),\, h(y)),$$

for all $x, y \in E$. Prove that the groups of isometries $\mathbf{U}(\varphi)$ and $\mathbf{U}(\psi)$ are isomomorphic (use the map $f \mapsto h \circ f \circ h^{-1}$ from $\mathbf{U}(\psi)$ to $\mathbf{U}(\varphi)$).

If φ is a nondegenerate Hermitian form of signature (p, q), prove that the group $\mathbf{U}(\varphi)$ is isomorphic to the group of $n \times n$ matrices A such that

$$A^\top \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix} \overline{A} = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}.$$

Remark: In view of question (f), the groups $\mathbf{U}(\varphi)$ and $\mathbf{SU}(\varphi)$ are also denoted by $\mathbf{U}(p, q)$ and $\mathbf{SU}(p, q)$ when φ has signature (p, q). They are Lie groups.

Problem 10.9 (a) If A is a real symmetric $n \times n$ matrix and B is a real skew symmetric $n \times n$ matrix, then $A + iB$ is Hermitian. Conversely, every Hermitian matrix can be written as $A + iB$, where A is real symmetric and B is real skew symmetric.

(b) Every complex $n \times n$ matrix can be written as $A + iB$, for some Hermitian matrices A, B.

Problem 10.10 (a) Given a complex $n \times n$ matrix A, prove that

$$\sum_{i,j=1}^{n} |a_{i,\,j}|^2 = \mathrm{tr}(A^*A) = \mathrm{tr}(AA^*).$$

(b) Prove that $\|A\| = \sqrt{\mathrm{tr}(A^*A)}$ defines a norm on matrices. Prove that

$$\|AB\| \le \|A\|\|B\|.$$

(c) When A is Hermitian, prove that

$$\|A\|^2 = \sum_{i=1}^{n} \lambda_i^2,$$

where the λ_i are the (real) eigenvalues of A.

Problem 10.11 Given a Hermitian matrix A, prove that $I_n + iA$ and $I_n - iA$ are invertible. Prove that $(I_n + iA)(I_n - iA)^{-1}$ is a unitary matrix.

Problem 10.12 Let E be a Hermitian space of dimension n. For any basis $(e_1, \ldots, e_n)$ of E, orthonormal or not, let G be the Gram matrix associated with $(e_1, \ldots, e_n)$, i.e., the matrix

$$G = (e_i \cdot e_j).$$

Given any linear map $f \colon E \to E$, if A is the matrix of f w.r.t. $(e_1, \ldots, e_n)$, prove that f is self-adjoint ($f^* = f$) iff

$$G^\top A = A^* G^\top.$$