Chapter 25

Isometries of Hermitian Spaces

25.1 The Cartan-Dieudonné Theorem, Hermitian Case

The Cartan-Dieudonné theorem can be generalized (Theorem 25.1.3), 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.

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 25.1.1 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 25.1.2 *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 $\rho_{-\theta} \circ s(u) = v$.

Proof. (1) Consider the (usual) reflection about the hyperplane orthogonal to $w = v - e^{-i\theta}u$. We have

$$s(u) = u - 2\,\frac{(u \cdot (v - e^{-i\theta}u))}{\|v - e^{-i\theta}u\|^2}\,(v - e^{-i\theta}u).$$

We need to compute

$$-2u \cdot (v - e^{-i\theta}u) \quad \text{and} \quad (v - e^{-i\theta}u) \cdot (v - e^{-i\theta}u).$$

Since $u \cdot v = e^{i\theta}|u \cdot v|$, we have

$$e^{-i\theta}u \cdot v = |u \cdot v| \quad \text{and} \quad e^{i\theta}v \cdot u = |u \cdot v|.$$

Using the above and the fact that $\|u\| = \|v\|$, we get

$$\begin{aligned} -2u \cdot (v - e^{-i\theta}u) &= 2e^{i\theta}\,\|u\|^2 - 2u \cdot v, \\ &= 2e^{i\theta}(\|u\|^2 - |u \cdot v|), \end{aligned}$$

and

$$\begin{aligned} (v - e^{-i\theta}u) \cdot (v - e^{-i\theta}u) &= \|v\|^2 + \|u\|^2 - e^{-i\theta}u \cdot v - e^{i\theta}v \cdot u, \\ &= 2(\|u\|^2 - |u \cdot v|), \end{aligned}$$

and thus,

$$-2\,\frac{(u \cdot (v - e^{-i\theta}u))}{\|(v - e^{-i\theta}u)\|^2}\,(v - e^{-i\theta}u) = e^{i\theta}(v - e^{-i\theta}u).$$

But then,

$$s(u) = u + e^{i\theta}(v - e^{-i\theta}u) = u + e^{i\theta}v - u = e^{i\theta}v,$$

and $s(u) = e^{i\theta}v$, as claimed.

(2) This part is easier. Consider the Hermitian reflection

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

We have

$$\begin{aligned}\rho_\theta(v) &= v + (e^{i\theta} - 1)\,\frac{(v\cdot v)}{\|v\|^2}\,v,\\ &= v + (e^{i\theta} - 1)v,\\ &= e^{i\theta}v.\end{aligned}$$

Thus, $\rho_\theta(v) = e^{i\theta}v$. Since ρ_θ is linear, changing v to $e^{i\theta}v$, we get

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

and thus, $\rho_{-\theta} \circ s(u) = v$. ☐

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 [100] and Ciarlet [33], 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 25.1.2 shows that the Cartan-Dieudonné is salvaged, since we can send u to v by a sequence of two Hermitian reflections when $u \neq v$ and $\|u\| = \|v\|$, and since the inverse of a Hermitian reflection is a Hermitian reflection. Actually, because we are over the complex field, a linear map always have (complex) eigenvalues, and we can get a slightly improved result.

Theorem 25.1.3 *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, $1 \leq j \leq n$.

When $n = 1$, every isometry $f \in \mathbf{U}(E)$ is either the identity or a Hermitian reflection ρ_θ, since for any nonnull vector u, we have $f(u) = e^{i\theta}u$ for some θ. We let u_1 be any nonnull unit vector.

Let us now consider the case where $n \geq 2$. Since $\mathbb{C}$ is algebraically closed, the characteristic polynomial $\det(f - \lambda\mathrm{id})$ of f has n complex roots which must be the form $e^{i\theta}$, since they have absolute value 1. Pick any such eigenvalue $e^{i\theta_1}$, and pick any eigenvector $u_1 \neq 0$ of f for $e^{i\theta_1}$ of unit length. If $F = \mathbb{C}u_1$ is the subspace spanned by u_1, we have $f(F) = F$, since $f(u_1) = e^{i\theta_1}u_1$. By lemma 10.2.6, $f(F^\perp) \subseteq F^\perp$ and $E = F \oplus F^\perp$. Furthermore, it is obvious that the restriction of f to $F^\perp$ is unitary. Since $\dim(F^\perp) = n - 1$, we can apply the induction hypothesis to $F^\perp$, and we get an orthonormal basis of eigenvectors $(u_2, \ldots, u_n)$ for $F^\perp$ 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, $2 \leq j \leq n$ Since $E = F \oplus F^\perp$ and $F = \mathbb{C}u_1$, the claim is proved. But then, if ρ_j is the Hermitian reflection about the hyperplane H_j orthogonal to u_j and of angle θ_j, it is obvious that

$$f = \rho_{\theta_n} \circ \cdots \circ \rho_{\theta_1}.$$

When $n \geq 2$, we have $\mathrm{id} = s \circ s$ for every reflection s. $\square$

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. Indeed, by the above theorem, with respect to the basis $(u_1, \ldots, u_n)$, $\det(f) = e^{i(\theta_1 + \cdots + \theta_n)}$, and letting $\theta = \theta_1 + \cdots + \theta_n$ and ρ_θ be the Hermitian reflection about the hyperplane orthogonal to u_1 and of angle θ, since $\rho_\theta \circ \rho_{-\theta} = \mathrm{id}$, we have

$$f = (\rho_\theta \circ \rho_{-\theta}) \circ f = \rho_\theta \circ (\rho_{-\theta} \circ f).$$

Letting $g = \rho_{-\theta} \circ f$, it is obvious that $\det(g) = 1$. 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 [128]) 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! This wonderful fact is proved in the next lemma.

Lemma 25.1.4 *Let E be a nontrivial Hermitian space. For any two distinct orthogonal vectors u, v such that $\|u\| = \|v\|$, we have*

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

Proof. Since u and v are orthogonal, each one is in the hyperplane orthogonal to the other, and thus,

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

Consequently, using linearity,

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

and since both $\rho_{v,\,-\theta} \circ \rho_{u,\,\theta}$ and $h_{v-\,u} \circ h_{v-\,e^{-i\theta}u}$ are the identity on the orthogonal complement of $\{u, v\}$, they are equal. Since we also have

$$\begin{aligned} h_{u+\,v}(u) &= -v, \\ h_{u+\,v}(v) &= -u, \\ h_{u+\,e^{i\theta}v}(u) &= -e^{i\theta}v, \\ h_{u+\,e^{i\theta}v}(v) &= -e^{-i\theta}u, \end{aligned}$$

it is immediately verified that

$$h_{v-\,u} \circ h_{v-\,e^{-i\theta}u} = h_{u+\,v} \circ h_{u+\,e^{i\theta}v}.$$

$\square$

We will use Lemma 25.1.4 as follows.

Lemma 25.1.5 *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

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

Proof. It is obvious from the definitions that

$$f = \rho_{u_n,\,\theta_n} \circ \cdots \circ \rho_{u_1,\,\theta_1},$$

and since the determinant of f is

$$D(f) = e^{i\theta_1} \cdots e^{i\theta_n} = e^{i(\theta_1+\cdots+\theta_n)}$$

and $\theta_1 + \cdots + \theta_n = 0$, we have $D(f) = e^0 = 1$, and f is a rotation. Letting

$$f_k = \rho_{u_k,\,-(\theta_1+\cdots+\theta_{k-1})} \circ \rho_{u_{k-1},\,\theta_1+\cdots+\theta_{k-1}} \circ \cdots \circ \rho_{u_3,\,-(\theta_1+\theta_2)} \circ \rho_{u_2,\,\theta_1+\theta_2} \circ \rho_{u_2,\,-\theta_1} \circ \rho_{u_1,\,\theta_1},$$

we prove by induction on k, $2 \leq k \leq n$, that

$$f_k(u_j) = \begin{cases} e^{i\theta_j}u_j & \text{if } 1 \leq j \leq k-1, \\ e^{-i(\theta_1+\cdots+\theta_{k-1})}u_k & \text{if } j = k, \text{ and} \\ u_j & \text{if } k+1 \leq j \leq n. \end{cases}$$

The base case was treated in Lemma 25.1.4. Now, the proof of Lemma 25.1.4 also showed that

$$\begin{aligned} \rho_{u_{k+1},\,-(\theta_1+\cdots+\theta_k)} \circ \rho_{u_k,\,\theta_1+\cdots+\theta_k}(u_k) &= e^{i(\theta_1+\cdots+\theta_k)}u_k, \\ \rho_{u_{k+1},\,-(\theta_1+\cdots+\theta_k)} \circ \rho_{u_k,\,\theta_1+\cdots+\theta_k}(u_{k+1}) &= e^{-i(\theta_1+\cdots+\theta_k)}u_{k+1}, \end{aligned}$$

and thus, using the induction hypothesis for k ($2 \leq k \leq n-1$), we have

$$\begin{aligned} f_{k+1}(u_j) &= \rho_{u_{k+1},\,-(\theta_1+\cdots+\theta_k)} \circ \rho_{u_k,\,\theta_1+\cdots+\theta_k} \circ f_k(u_j) = e^{i\theta_j}u_j, \quad 1 \leq j \leq k-1, \\ f_{k+1}(u_k) &= \rho_{u_{k+1},\,-(\theta_1+\cdots+\theta_k)} \circ \rho_{u_k,\,\theta_1+\cdots+\theta_k} \circ f_k(u_k) = e^{i(\theta_1+\cdots+\theta_k)}e^{-i(\theta_1+\cdots+\theta_{k-1})}u_k = e^{i\theta_k}u_k, \\ f_{k+1}(u_{k+1}) &= \rho_{u_{k+1},\,-(\theta_1+\cdots+\theta_k)} \circ \rho_{u_k,\,\theta_1+\cdots+\theta_k} \circ f_k(u_{k+1}) = e^{-i(\theta_1+\cdots+\theta_k)}u_{k+1}, \\ f_{k+1}(u_j) &= \rho_{u_{k+1},\,-(\theta_1+\cdots+\theta_k)} \circ \rho_{u_k,\,\theta_1+\cdots+\theta_k} \circ f_k(u_j) = u_j, \quad k+1 \leq j \leq n, \end{aligned}$$

which proves the induction step.

As a summary, we proved that

$$f_n(u_j) = \begin{cases} e^{i\theta_j} u_j & \text{if } 1 \leq j \leq n-1, \\ e^{-i(\theta_1 + \cdots + \theta_{n-1})} u_n & \text{when } j = n, \end{cases}$$

but since $\theta_1 + \cdots + \theta_n = 0$, we have $\theta_n = -(\theta_1 + \cdots + \theta_{n-1})$, and the last expression is in fact

$$f_n(u_n) = e^{i\theta_n} u_n.$$

Therefore, we proved that

$$f = \rho_{u_n, \theta_n} \circ \cdots \circ \rho_{u_1, \theta_1} = \rho_{u_n, -(\theta_1 + \cdots + \theta_{n-1})} \circ \rho_{u_{n-1}, \theta_1 + \cdots + \theta_{n-1}} \circ \cdots \circ \rho_{u_2, -\theta_1} \circ \rho_{u_1, \theta_1},$$

and using Lemma 25.1.4, we also have

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

which completes the proof. $\square$

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

Theorem 25.1.6 *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.*

Proof. By Theorem 25.1.3, $f \in \mathbf{SU}(n)$ can be written as a composition

$$\rho_{u_n, \theta_n} \circ \cdots \circ \rho_{u_1, \theta_1},$$

where $(u_1, \ldots, u_n)$ is an orthonormal basis of eigenvectors. Since f is a rotation, $\det(f) = 1$, and this implies that $\theta_1 + \cdots + \theta_n = 0$. By lemma 25.1.5,

$$f = h_{u_n - u_{n-1}} \circ h_{u_n - e^{-i(\theta_1 + \cdots + \theta_{n-1})} u_{n-1}} \circ \cdots \circ h_{u_2 - u_1} \circ h_{u_2 - e^{-i\theta_1} u_1},$$

a composition of $2n-2$ hyperplane reflections. In general, if $f \in \mathbf{U}(n)$, by the remark after Theorem 25.1.3, f can be written as $f = \rho_\theta \circ g$, where $g \in \mathbf{SU}(n)$ is a rotation, and ρ_θ is a Hermitian reflection. We conclude by applying what we just proved to g. $\square$

As a corollary of Theorem 25.1.6, 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 [46]. Evidently, it was settled since the publication of the third edition of the book [46].

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. As a consequence, a version of Theorem 7.2.5 holds for rotations in a Hermitian space of dimension at least 3.

Theorem 25.1.7 *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}.*

Proof. It is identical to that of Theorem 7.2.5, except that it uses Theorem 25.1.6 instead of Theorem 7.2.1. The second part of the Lemma also holds, because if $u \neq 0$ is an eigenvector of f for 1, then u is one of the vectors in the orthonormal basis of eigenvectors used in 25.1.3. The details are left as an exercise. $\square$

We now show that the QR-decomposition in terms of (complex) Householder matrices holds for complex matrices. We need the version of Lemma 25.1.2 and a trick at the end of the argument, but the proof is basically unchanged.

Lemma 25.1.8 *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.

Proof. The proof is very similar to the proof of Lemma 7.3.1, but it needs to be modified a little bit since Lemma 25.1.2 is weaker than Lemma 7.1.3. We explain how to modify the induction step, leaving the base case and the rest of the proof as an exercise.

As in the proof of Lemma 7.3.1, the vectors $(e_1, \ldots, e_k)$ form a basis for the subspace denoted as U_k', the vectors $(e_{k+1}, \ldots, e_n)$ form a basis for the subspace denoted as U_k'', the subspaces U_k' and U_k'' are orthogonal, and $E = U_k' \oplus U_k''$. Let

$$u_{k+1} = h_k \circ \cdots \circ h_2 \circ h_1(v_{k+1}).$$

We can write

$$u_{k+1} = u_{k+1}' + u_{k+1}'',$$

where $u_{k+1}' \in U_k'$ and $u_{k+1}'' \in U_k''$. Let

$$r_{k+1,k+1} = \left\| u_{k+1}'' \right\|, \quad \text{and} \quad e^{i\theta_{k+1}} |u_{k+1}'' \cdot e_{k+1}| = u_{k+1}'' \cdot e_{k+1}.$$

If $u_{k+1}'' = e^{i\theta_{k+1}} r_{k+1,k+1}\, e_{k+1}$, we let $h_{k+1} = \mathrm{id}$. Otherwise, by Lemma 25.1.2, there is a unique hyperplane reflection h_{k+1} such that

$$h_{k+1}(u_{k+1}'') = e^{i\theta_{k+1}} r_{k+1,k+1}\, e_{k+1},$$

where h_{k+1} is the reflection about the hyperplane H_{k+1} orthogonal to the vector

$$w_{k+1} = r_{k+1,k+1}\, e_{k+1} - e^{-i\theta_{k+1}} u_{k+1}''.$$

At the end of the induction, we have a triangular matrix R, but the diagonal entries $e^{i\theta_j} r_{j,\,j}$ of R may be complex. Letting

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

we observe that the diagonal entries of the matrix of vectors

$$r'_j = h_{n+1} \circ h_n \circ \cdots \circ h_2 \circ h_1(v_j)$$

is triangular with nonnegative entries. $\square$

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 25.1.8 immediately implies the QR-decomposition for arbitrary complex $n \times n$-matrices, where Q is now unitary (see Kincaid and Cheney [100], Golub and Van Loan [75], Trefethen and Bau [170], or Ciarlet [33]).

Lemma 25.1.9 *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.

Proof. It is essentially identical to the proof of Lemma 7.3.2, and we leave the details as an exercise. For the last statement, observe that $h_{n+1} \circ \cdots \circ h_1$ is also an isometry. $\square$

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.

25.2 Affine Isometries (Rigid Motions)

In this section, we study very briefly the affine isometries of a Hermitian space. Most results holding for Euclidean affine spaces generalize without any problems to Hermitian spaces.

The characterization of the set of fixed points of an affine map is unchanged. Similarly, every affine isometry f (of a Hermitian space) can be written uniquely as

$$f = t \circ g, \quad \text{with} \quad t \circ g = g \circ t,$$

where g is an isometry having a fixed point, and t is a translation by a vector τ such that $\overrightarrow{f}(\tau) = \tau$, and with some additional nice properties (see lemma 25.2.6). A generalization of the Cartan-Dieudonné theorem can easily be shown: every affine isometry in $\mathbf{Is}(n, \mathbb{C})$ can be written as the composition of at most $2n - 1$ isometries if it has a fixed point, or else as the composition of at most $2n + 1$ isometries, where all these isometries are hyperplane reflections except for possibly one Hermitian reflection. We also prove that every rigid motion in $\mathbf{SE}(n, \mathbb{C})$ is the composition of at most $2n - 2$ flips (for $n \geq 3$).

Definition 25.2.1 Given any two nontrivial Hermitian affine spaces E and F of the same finite dimension n, a function $f\colon E \to F$ is *an affine isometry (or rigid map)* iff it is an affine map and

$$\|\mathbf{f}(\mathbf{a})\mathbf{f}(\mathbf{b})\| = \|\mathbf{ab}\|,$$

for all $a, b \in E$. When $E = F$, an affine isometry $f\colon E \to E$ is also called a *rigid motion*.

Thus, an affine isometry is an affine map that preserves the distance. This is a rather strong requirement, but unlike the Euclidean case, not strong enough to force f to be an affine map.

The following simple lemma is left as an exercise.

Lemma 25.2.2 *Given any two nontrivial Hermitian affine spaces E and F of the same finite dimension n, an affine map $f\colon E \to F$ is an affine isometry iff its associated linear map $\overrightarrow{f}\colon \overrightarrow{E} \to \overrightarrow{F}$ is an isometry. An affine isometry is a bijection.*

As in the Euclidean case, given an affine isometry $f\colon E \to E$, if $\overrightarrow{f}$ is a rotation, we call f a *proper (or direct) affine isometry*, and if $\overrightarrow{f}$ is a an improper linear isometry, we call f a *an improper (or skew) affine isometry*. It is easily shown that the set of affine isometries $f\colon E \to E$ forms a group, and those for which $\overrightarrow{f}$ is a rotation is a subgroup. The group of affine isometries, or rigid motions, is a subgroup of the affine group $\mathbf{GA}(E, \mathbb{C})$ denoted as $\mathbf{Is}(E, \mathbb{C})$ (or $\mathbf{Is}(n, \mathbb{C})$ when $E = \mathbb{C}^n$). The subgroup of $\mathbf{Is}(E, \mathbb{C})$ consisting of the direct rigid motions is also a subgroup of $\mathbf{SA}(E, \mathbb{C})$, and it is denoted as $\mathbf{SE}(E, \mathbb{C})$ (or $\mathbf{SE}(n, \mathbb{C})$, when $E = \mathbb{C}^n$). The translations are the affine isometries f for which $\overrightarrow{f} = \mathrm{id}$, the identity map on $\overrightarrow{E}$. The following lemma is the counterpart of lemma 10.3.2 for isometries between Hermitian vector spaces.

Lemma 25.2.3 *Given any two nontrivial Hermitian affine 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 an affine map and $\|\mathbf{f(a)f(b)}\| = \|\mathbf{ab}\|$, for all $a, b \in E$.

(2) $\|\mathbf{f(a)f(b)}\| = \|\mathbf{ab}\|$, and there is some $\Omega \in E$ such that

$$f(\Omega + i\mathbf{ab}) = f(\Omega) + i(\mathbf{f(\Omega)f(\Omega + ab)}),$$

for all $a, b \in E$.

Proof. Obviously, (1) implies (2). The proof that that (2) implies (1) is similar to the proof of Lemma 7.4.3, but uses Lemma 10.3.2 instead of Lemma 6.3.2. The details are left as an exercise. □

Inspection of the proof shows immediately that Lemma 7.5.1 holds for Hermitian spaces. For the sake of completeness, we restate the Lemma in the complex case.

Lemma 25.2.4 *Let E be any complex affine space of finite dimension For every affine map $f\colon E \to E$, let $Fix(f) = \{a \in E \mid f(a) = a\}$ be the set of fixed points of f. The following properties hold:*

(1) If f has some fixed point a, so that $Fix(f) \neq \emptyset$, then $Fix(f)$ is an affine subspace of E such that

$$Fix(f) = a + E(1, \overrightarrow{f}) = a + \mathrm{Ker}\,(\overrightarrow{f} - \mathrm{id}),$$

where $E(1, \overrightarrow{f})$ is the eigenspace of the linear map $\overrightarrow{f}$ for the eigenvalue 1.

(2) The affine map f has a unique fixed point iff $E(1, \overrightarrow{f}) = \mathrm{Ker}\,(\overrightarrow{f} - \mathrm{id}) = \{0\}$.

Affine orthogonal symmetries are defined just as in the Euclidean case, and Lemma 7.6.1 also applies to complex affine spaces.

Lemma 25.2.5 *Given any affine complex space E, if $f\colon E \to E$ and $g\colon E \to E$ are orthogonal symmetries about parallel affine subspaces F_1 and F_2, then $g \circ f$ is a translation defined by the vector $2\mathbf{ab}$, where $\mathbf{ab}$ is any vector perpendicular to the common direction $\overrightarrow{F}$ of F_1 and F_2 such that $\|\mathbf{ab}\|$ is the distance between F_1 and F_2, with $a \in F_1$ and $b \in F_2$. Conversely, every translation by a vector τ is obtained as the composition of two orthogonal symmetries about parallel affine subspaces F_1 and F_2 whose common direction is orthogonal to $\tau = \mathbf{ab}$, for some $a \in F_1$ and some $b \in F_2$ such that the distance betwen F_1 and F_2 is $\|\mathbf{ab}\|/2$.*

It is easy to check that the proof of Lemma 7.6.2 also holds in the Hermitian case.

Lemma 25.2.6 *Let E be a Hermitian affine space of finite dimension n. For every affine isometry $f\colon E \to E$, there is a unique isometry $g\colon E \to E$ and a unique translation $t = t_\tau$, with $\overrightarrow{f}(\tau) = \tau$ (i.e., $\tau \in \mathrm{Ker}\,(\overrightarrow{f} - \mathrm{id})$), such that the set $Fix(g) = \{a \in E, \mid g(a) = a\}$ of fixed points of g is a nonempty affine subspace of E of direction*

$$\overrightarrow{G} = \mathrm{Ker}\,(\overrightarrow{f} - \mathrm{id}) = E(1, \overrightarrow{f}),$$

and such that

$$f = t \circ g \quad \textit{and} \quad t \circ g = g \circ t.$$

Furthermore, we have the following additional properties:

(a) $f = g$ and $\tau = 0$ iff f has some fixed point, i.e., iff $Fix(f) \neq \emptyset$.

(b) If f has no fixed points, i.e., $Fix(f) = \emptyset$, then $\dim(\mathrm{Ker}\,(\overrightarrow{f} - \mathrm{id})) \geq 1$.

The remarks made in the Euclidean case also apply to the Hermitian case. In particular, the fact that E has finite dimension is only used to prove (b).

A version of the Cartan-Dieudonné also holds for affine isometries, but it may not be possible to get rid of Hermitian reflections entirely.

Theorem 25.2.7 *Let E be an affine Hermitian space of dimension $n \geq 1$. Every affine isometry in $\mathbf{Is}(n, \mathbb{C})$ can be written as the composition of at most $2n-1$ isometries if it has a fixed point, or else as the composition of at most $2n+1$ isometries, where all these isometries are hyperplane reflections except for possibly one Hermitian reflection. When $n \geq 2$, the identity is the composition of any reflection with itself.*

Proof. The proof is very similar to the proof of Theorem 7.7.1, except that it uses Theorem 25.1.6 instead of Theorem 7.2.1. The details are left as an exercise. □

When $n \geq 3$, as in the Euclidean case, we can characterize the affine isometries in $\mathbf{SE}(n, \mathbb{C})$ in terms of flips, and we can even bound the number of flips by $2n - 2$.

Theorem 25.2.8 *Let E be a Hermitian affine space of dimension $n \geq 3$. Every rigid motion $f \in \mathbf{SE}(E, \mathbb{C})$ is the composition of an even number of flips $f = f_{2k} \circ \cdots \circ f_1$, where $k \leq n - 1$.*

Proof. It is very similar to the proof of theorem 7.7.2, but it uses Lemma 25.1.7 instead of Lemma 7.2.5. The details are left as an exercise. □

A more detailed study of the rigid motions of Hermitian spaces of dimension 2 and 3 would seem worthwhile, but we are not aware of any reference on this subject.