# III. Inner-Product Spaces, 89-116

DOI: 10.3792/euclid/9781429799980-3

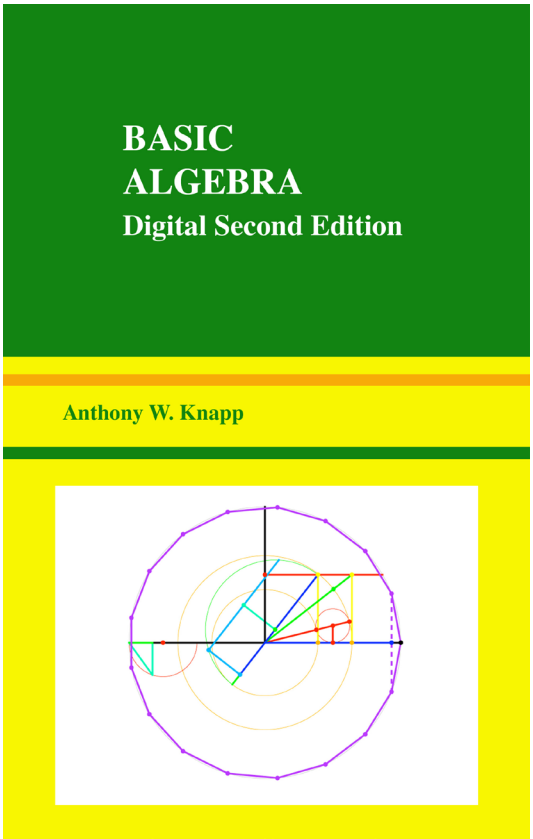
from

***Basic Algebra***

***Digital Second Edition***

Anthony W. Knapp

Full Book DOI: 10.3792/euclid/9781429799980
ISBN: 978-1-4297-9998-0



Anthony W. Knapp
81 Upper Sheep Pasture Road
East Setauket, N.Y. 11733–1729, U.S.A.
Email to: aknapp@math.stonybrook.edu
Homepage: www.math.stonybrook.edu/~aknapp

Title: Basic Algebra
Cover: Construction of a regular heptadecagon, the steps shown in color sequence; see page 505.

Mathematics Subject Classification (2010): 15–01, 20–01, 13–01, 12–01, 16–01, 08–01, 18A05, 68P30.

First Edition, ISBN-13 978-0-8176-3248-9

Published by Birkhäuser Boston

Digital Second Edition, not to be sold, no ISBN

Published by the Author

# CHAPTER III

# Inner-Product Spaces

**Abstract.** This chapter investigates the effects of adding the additional structure of an inner product to a finite-dimensional real or complex vector space.

Section 1 concerns the effect on the vector space itself, defining inner products and their corresponding norms and giving a number of examples and formulas for the computation of norms. Vector-space bases that are orthonormal play a special role.

Section 2 concerns the effect on linear maps. The inner product makes itself felt partly through the notion of the adjoint of a linear map. The section pays special attention to linear maps that are self-adjoint, i.e., are equal to their own adjoints, and to those that are unitary, i.e., preserve norms of vectors.

Section 3 proves the Spectral Theorem for self-adjoint linear maps on finite-dimensional inner-product spaces. The theorem says in part that any self-adjoint linear map has an orthonormal basis of eigenvectors. The Spectral Theorem has several important consequences, one of which is the existence of a unique positive semidefinite square root for any positive semidefinite linear map. The section concludes with the polar decomposition, showing that any linear map factors as the product of a unitary linear map and a positive semidefinite one.

## 1. Inner Products and Orthonormal Sets

In this chapter we examine the effect of adding further geometric structure to the structure of a real or complex vector space as defined in Chapter II. To be a little more specific in the cases of $\mathbb{R}^2$ and $\mathbb{R}^3$, the development of Chapter II amounted to working with points, lines, planes, coordinates, and parallelism, but nothing further. In the present chapter, by comparison, we shall take advantage of additional structure that captures the notions of distances and angles.

We take $\mathbb{F}$ to be $\mathbb{R}$ or $\mathbb{C}$, continuing to call its members the scalars. We do not allow $\mathbb{F}$ to be $\mathbb{Q}$ in this chapter; the main results will make essential use of additional facts about $\mathbb{R}$ and $\mathbb{C}$ beyond those of addition, subtraction, multiplication, and division. The relevant additional facts are summarized in Sections A3 and A4 of the appendix.[1]

[1] The theory of Chapter II will be observed in Chapter IV to extend to any "field" $\mathbb{F}$ in place of $\mathbb{Q}$ or $\mathbb{R}$ or $\mathbb{C}$, but the theory of the present chapter is limited to $\mathbb{R}$ and $\mathbb{C}$, as well as some other special fields that we shall not try to isolate.

Many of the results that we obtain will be limited to the finite-dimensional case. The theory of inner-product spaces that we develop has an infinite-dimensional generalization, but useful results for the generalization make use of a hypothesis of "completeness" for an inner-product space that we are not in a position to verify in examples.[2]

Let $V$ be a vector space over $\mathbb{F}$. An **inner product** on $V$ is a function from $V \times V$ into $\mathbb{F}$, which we here denote by $(\cdot\,,\,\cdot)$, with the following properties:

(i) the function $u \mapsto (u, v)$ of $V$ into $\mathbb{F}$ is linear,
(ii) the function $v \mapsto (u, v)$ of $V$ into $\mathbb{F}$ is **conjugate linear** in the sense that it satisfies $(u, v_1 + v_2) = (u, v_1) + (u, v_2)$ for $v_1$ and $v_2$ in $V$ and $(u, cv) = \bar{c}(u, v)$ for $v$ in $V$ and $c$ in $\mathbb{F}$,
(iii) $(u, v) = \overline{(v, u)}$ for $u$ and $v$ in $V$,
(iv) $(v, v) \geq 0$ for all $v$ in $V$,
(v) $(v, v) = 0$ only if $v = 0$ in $V$.

The overbars in (ii) and (iii) indicate complex conjugation. Property (ii) reduces when $\mathbb{F} = \mathbb{R}$ to the fact that $v \mapsto (u, v)$ is linear. Properties (i) and (ii) together are summarized by saying that $(\cdot\,,\,\cdot)$ is **bilinear** if $\mathbb{F} = \mathbb{R}$ or **sesquilinear** if $\mathbb{F} = \mathbb{C}$. Property (iii) is summarized when $\mathbb{F} = \mathbb{R}$ by saying that $(\cdot\,,\,\cdot)$ is **symmetric**, or when $\mathbb{F} = \mathbb{C}$ by saying that $(\cdot\,,\,\cdot)$ is **Hermitian symmetric**.

An **inner-product space**, for purposes of this book, is a vector space over $\mathbb{R}$ or $\mathbb{C}$ with an inner product in the above sense.[3,4]

EXAMPLES.

(1) $V = \mathbb{R}^n$ with $(\cdot\,,\,\cdot)$ as the **dot product**, i.e., with $(x, y) = y^t x = x_1 y_1 + \cdots + x_n y_n$ if $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$. The traditional notation for the dot product is $x \cdot y$.

(2) $V = \mathbb{C}^n$ with $(\cdot\,,\,\cdot)$ defined by $(x, y) = \bar{y}^t x = x_1 \bar{y}_1 + \cdots + x_n \bar{y}_n$ if $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$. Here $\bar{y}$ denotes the entry-by-entry complex conjugate of $y$. The sesquilinear expression $(\cdot\,,\,\cdot)$ is different from the complex *bilinear* **dot product** $x \cdot y = x_1 y_1 + \cdots + x_n y_n$.

[2] A careful study in the infinite-dimensional case is normally made only after the development of a considerable number of topics in real analysis.

[3] When the scalars are complex, many books emphasize the presence of complex scalars by referring to the inner product as a "Hermitian inner product." This book does not need to distinguish the complex case very often and therefore will not use the modifier "Hermitian" with the term "inner product."

[4] Some authors, particularly in connection with mathematical physics, reverse the roles of the two variables, defining inner products to be conjugate linear in the first variable and linear in the second variable.

(3) $V$ equal to the vector space of all complex-valued polynomials with $(f, g) = \int_0^1 f(x)\overline{g(x)}\,dx$.

Let $V$ be an inner-product space. If $v$ is in $V$, define $\|v\| = \sqrt{(v, v)}$, calling $\|\cdot\|$ the **norm** associated with the inner product. The norm of $v$ is understood to be the nonnegative square root of the nonnegative real number $(v, v)$ and is well defined as a consequence of (iv). In the case of $\mathbb{R}^n$, $\|x\|$ is the Euclidean distance $\sqrt{x_1^2 + \cdots + x_n^2}$ from the origin to the column vector $x = (x_1, \ldots, x_n)$. In this interpretation the dot product of two nonzero vectors in $\mathbb{R}^n$ is shown in analytic geometry to be given by $x \cdot y = \|x\|\|y\| \cos\theta$, where $\theta$ is the angle between the vectors $x$ and $y$.

Direct expansion of norms squared of sums of vectors using bilinearity or sesquilinearity leads to certain formulas of particular interest. The formula that we shall use most frequently is

$$\|u + v\|^2 = \|u\|^2 + 2\,\mathrm{Re}(u, v) + \|v\|^2,$$

which generalizes from $\mathbb{R}^2$ a version of the **law of cosines** in trigonometry relating the lengths of the three sides of a triangle when one of the angles is known. With the additional hypothesis that $(u, v) = 0$, this formula generalizes from $\mathbb{R}^2$ the **Pythagorean Theorem**

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

Another such formula is the **parallelogram law**

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2 \qquad \text{for all } u \text{ and } v \text{ in } V,$$

which is proved by computing $\|u + v\|^2$ and $\|u - v\|^2$ by the law of cosines and adding the results. The name "parallelogram law" is explained by the geometric interpretation in the case of the dot product for $\mathbb{R}^2$ and is illustrated in Figure 3.1. That figure uses the familiar interpretation of vectors in $\mathbb{R}^2$ as arrows, two arrows being identified if they are translates of one another; thus the arrow from $v$ to $u$ represents the vector $u - v$.

The parallelogram law is closely related to a formula for recovering the inner product from the norm, namely

$$(u, v) = \frac{1}{4} \sum_k i^k \|u + i^k v\|^2,$$

where the sum extends for $k \in \{0, 2\}$ if the scalars are real and extends for $k \in \{0, 1, 2, 3\}$ if the scalars are complex. This formula goes under the name

**polarization**. To prove it, we expand $\|u + i^k v\|^2 = \|u\|^2 + 2\,\mathrm{Re}(u, i^k v) + \|v\|^2 = \|u\|^2 + 2\,\mathrm{Re}\big((-i)^k(u, v)\big) + \|v\|^2$. Multiplying by $i^k$ and summing on $k$ shows that $\sum_k i^k \|u + i^k v\|^2 = 2\sum_k i^k\,\mathrm{Re}\big((-i)^k(u, v)\big)$. If $k$ is even, then $i^k\,\mathrm{Re}((-i)^k z) = \mathrm{Re}\, z$ for any complex $z$, while if $k$ is odd, then $i^k\,\mathrm{Re}((-i)^k z) = i\,\mathrm{Im}\, z$. So $2\sum_k i^k\,\mathrm{Re}((-i)^k z) = 4z$, and $\sum_k i^k \|u+i^k v\|^2 = 4(u, v)$, as asserted.

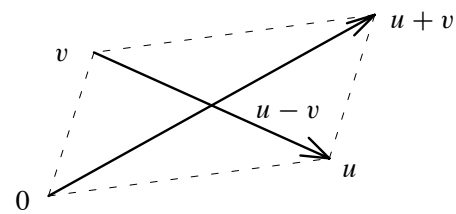



FIGURE 3.1. Geometric interpretation of the parallelogram law: the sum of the squared lengths of the four sides of a parallelogram equals the sum of the squared lengths of the diagonals.

**Proposition 3.1** (Schwarz inequality). In any inner-product space $V$, $|(u, v)| \le \|u\|\|v\|$ for all $u$ and $v$ in $V$.

REMARK. The proof is written so as to use properties (i) through (iv) in the definition of inner product but not (v), a situation often encountered with integrals.

PROOF. Possibly replacing $u$ by $e^{i\theta}u$ for some real $\theta$, we may assume that $(u, v)$ is real. In the case that $\|v\| \ne 0$, the law of cosines gives

$$\big|u - \|v\|^{-2}(u, v)v\big|^2 = \|u\|^2 - 2\|v\|^{-2}|(u, v)|^2 + \|v\|^{-4}|(u, v)|^2\|v\|^2.$$

The left side is $\ge 0$, and the right side simplifies to $\|u\|^2 - \|v\|^{-2}|(u, v)|^2$. Thus the inequality follows in this case.

In the case that $\|v\| = 0$, it is enough to prove that $(u, v) = 0$ for all $u$. If $c$ is a scalar, then we have

$$\|u + cv\|^2 = \|u\|^2 + 2\,\mathrm{Re}\big(c(u, v)\big) + |c|^2\|v\|^2 = \|u\|^2 + 2\,\mathrm{Re}\big(c(u, v)\big).$$

The left side is $\ge 0$ as $c$ varies, but the right side is $< 0$ for a suitable choice of $c$ unless $(u, v) = 0$. This completes the proof. □

**Proposition 3.2.** In any inner-product space $V$, the norm satisfies

(a) $\|v\| \ge 0$ for all $v$ in $V$, with equality if and only if $v = 0$,
(b) $\|cv\| = |c|\|v\|$ for all $v$ in $V$ and all scalars $c$,
(c) $\|u + v\| \le \|u\| + \|v\|$ for all $u$ and $v$ in $V$.

PROOF. Conclusion (a) is immediate from properties (iv) and (v) of an inner product, and (b) follows since $\|cv\|^2 = (cv, cv) = c\bar{c}(v, v) = |c|^2\|v\|^2$. Finally we use the law of cosines and the Schwarz inequality (Proposition 3.1) to write $\|u+v\|^2 = \|u\|^2+2\,\mathrm{Re}(u, v)+\|v\|^2 \le \|u\|^2+2\|u\|\|v\|+\|v\|^2 = (\|u\|+\|v\|)^2$. Taking the square root of both sides yields (c). □

Two vectors $u$ and $v$ in $V$ are said to be **orthogonal** if $(u, v) = 0$, and one sometimes writes $u \perp v$ in this case. The notation is a reminder of the interpretation in the case of dot product—that dot product 0 means that the cosine of the angle between the two vectors is 0 and the vectors are therefore perpendicular. An **orthogonal set** in $V$ is a set of vectors such that each pair is orthogonal.

The nonzero members of an orthogonal set are linearly independent. In fact, if $\{v_1, \dots, v_k\}$ is an orthogonal set of nonzero vectors and some linear combination has $c_1v_1 + \cdots + c_kv_k = 0$, then the inner product of this relation with $v_j$ gives $0 = (c_1v_1 + \cdots + c_kv_k,\ v_j) = c_j\|v_j\|^2$, and we see that $c_j = 0$ for each $j$.

A **unit vector** in $V$ is a vector $u$ with $\|u\| = 1$. If $v$ is any nonzero vector, then $v/\|v\|$ is a unit vector. An **orthonormal set** in $V$ is an orthogonal set of unit vectors. Under the assumption that $V$ is finite-dimensional, an **orthonormal basis** of $V$ is an orthonormal set that is a vector-space basis.[5]

EXAMPLES.

(1) In $\mathbb{R}^n$ or $\mathbb{C}^n$, the standard basis $\{e_1, \dots, e_n\}$ is an orthonormal set.

(2) Let $V$ be the complex inner-product space of all complex finite linear combinations, for $n$ from $-N$ to $+N$, of the functions $x \mapsto e^{inx}$ on the closed interval $[-\pi, \pi]$, the inner product being $(f, g) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx$. With respect to this inner product, the functions $e^{inx}$ form an orthonormal set.

A simple but important exercise in an inner-product space is to resolve a vector into the sum of a multiple of a given unit vector and a vector orthogonal to the given unit vector. This exercise is solved as follows: If $v$ is given and $u$ is a unit vector, then $v$ decomposes as

$$v = (v, u)u + \big(v - (v, u)u\big).$$

Here $(v, u)u$ is a multiple of $u$, and the two components are orthogonal since $\big(u,\ v - (v, u)u\big) = (u, v) - \overline{(v, u)}(u, u) = (u, v) - (u, v) = 0$. This decomposition is unique since if $v = v_1 + v_2$ with $v_1 = cu$ and $(v_2, u) = 0$, then the inner product of $v = v_1 + v_2$ with $u$ yields $(v, u) = (cu, u) + (v_2, u) = c$. Hence

[5]In the infinite-dimensional theory the term "orthonormal basis" is used for an orthonormal set that spans $V$ when *limits* of finite sums are allowed, in addition to finite sums themselves; when $V$ is infinite-dimensional, an orthonormal basis is never large enough to be a vector-space basis.

$c$ must be $(v, u)$, $v_1$ must be $(v, u)u$, and $v_2$ must be $v - (v, u)u$. Figure 3.2 illustrates the decomposition, and Proposition 3.3 generalizes it by replacing the multiples of a single unit vector by the span of a finite orthonormal set.

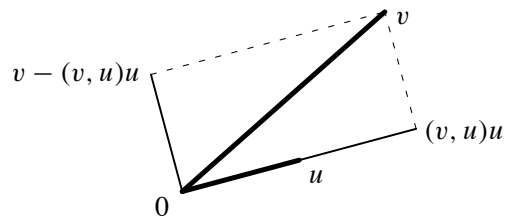



FIGURE 3.2. Resolution of $v$ into a component $(v, u)u$ parallel to a unit vector $u$ and a component orthogonal to $u$.

**Proposition 3.3.** Let $V$ be an inner-product space. If $\{u_1, \dots, u_k\}$ is an orthonormal set in $V$ and if $v$ is given in $V$, then there exists a unique decomposition

$$v = c_1 u_1 + \cdots + c_k u_k + v^{\perp}$$

with $v^{\perp}$ orthogonal to $u_j$ for $1 \le j \le k$. In this decomposition $c_j = (v, u_j)$.

REMARK. The proof illustrates a technique that arises often in mathematics. We seek to prove an existence–uniqueness theorem, and we begin by making calculations toward uniqueness that narrow down the possibilities. We are led to some formulas or conditions, and we use these to define the object in question and thereby prove existence. Although it may not be so clear except in retrospect, this was the technique that lay behind proving the equivalence of various conditions for the invertibility of a square matrix in Section I.6. The technique occurred again in defining and working with determinants in Section II.7.

PROOF OF UNIQUENESS. Taking the inner product of both sides with $u_j$, we obtain $(v, u_j) = (c_1 u_1 + \cdots + c_k u_k + v^{\perp}, u_j) = c_j$ for each $j$. Then $c_j = (v, u_j)$ is forced, and $v^{\perp}$ must be given by $v - (v, u_1)u_1 - \cdots - (v, u_k)u_k$. □

PROOF OF EXISTENCE. Putting $c_j = (v, u_j)$, we need check only that the difference $v - (v, u_1)u_1 - \cdots - (v, u_k)u_k$ is orthogonal to each $u_j$ with $1 \le j \le k$. Direct calculation gives

$$\left(v - \textstyle\sum_i (v, u_i)u_i,\ u_j\right) = (v, u_j) - \textstyle\sum_i ((v, u_i)u_i, u_j) = (v, u_j) - (v, u_j) = 0,$$

and the proof is complete. □

**Corollary 3.4** (Bessel's inequality). Let $V$ be an inner-product space. If $\{u_1, \dots, u_k\}$ is an orthonormal set in $V$ and if $v$ is given in $V$, then $\sum_{j=1}^{k} |(v, u_j)|^2 \le \|v\|^2$ with equality if and only if $v$ is in $\operatorname{span}\{u_1, \dots, u_k\}$.

PROOF. Using Proposition 3.3, write $v = \sum_{j=1}^{k} (v, u_j)u_j + v^\perp$ with $v^\perp$ orthogonal to $u_1, \dots, u_k$. Then

$$\begin{aligned}
\|v\|^2 &= \left(\textstyle\sum_{i=1}^{k} (v, u_i)u_i + v^\perp,\ \ \sum_{j=1}^{k} (v, u_j)u_j + v^\perp\right) \\
&= \textstyle\sum_{i,j} (v, u_i)\overline{(v, u_j)}(u_i, u_j) + \left(\sum_i (v, u_i)u_i, v^\perp\right) \\
&\quad + \left(v^\perp, \textstyle\sum_j (v, u_j)u_j\right) + \|v^\perp\|^2 \\
&= \textstyle\sum_{i,j} (v, u_i)\overline{(v, u_j)}\delta_{ij} + 0 + 0 + \|v^\perp\|^2 \\
&= \textstyle\sum_{j=1}^{k} |(v, u_j)|^2 + \|v^\perp\|^2.
\end{aligned}$$

From Proposition 3.3 we know that $v$ is in $\operatorname{span}\{u_1, \dots, u_k\}$ if and only if $v^\perp = 0$, and the corollary follows. □

We shall now impose the condition of finite dimensionality in order to obtain suitable kinds of orthonormal sets. The argument will enable us to give a basis-free interpretation of Proposition 3.3 and Corollary 3.4, and we shall obtain equivalent conditions for the vector $v^\perp$ in Proposition 3.3 and Corollary 3.4 to be 0 for every $v$.

If an ordered set of $k$ linearly independent vectors in the inner-product space $V$ is given, the above proposition suggests a way of adjusting the set so that it becomes orthonormal. Let us write the formulas here and carry out the verification via Proposition 3.3 in the proof of Proposition 3.5 below. The method of adjusting the set so as to make it orthonormal is called the **Gram–Schmidt orthogonalization process**. The given linearly independent set is denoted by $\{v_1, \dots, v_k\}$, and we define

$$\begin{aligned}
u_1 &= \frac{v_1}{\|v_1\|}, \\
u_2' &= v_2 - (v_2, u_1)u_1, \\
u_2 &= \frac{u_2'}{\|u_2'\|}, \\
u_3' &= v_3 - (v_3, u_1)u_1 - (v_3, u_2)u_2, \\
u_3 &= \frac{u_3'}{\|u_3'\|}, \\
&\ \ \vdots \\
u_k' &= v_k - (v_k, u_1)u_1 - \cdots - (v_k, u_{k-1})u_{k-1}, \\
u_k &= \frac{u_k'}{\|u_k'\|}.
\end{aligned}$$

**Proposition 3.5.** If $\{v_1, \ldots, v_k\}$ is a linearly independent set in an inner-product space $V$, then the Gram–Schmidt orthogonalization process replaces $\{v_1, \ldots, v_k\}$ by an orthonormal set $\{u_1, \ldots, u_k\}$ such that $\operatorname{span}\{v_1, \ldots, v_j\} = \operatorname{span}\{u_1, \ldots, u_j\}$ for all $j$.

PROOF. We argue by induction on $j$. The base case is $j = 1$, and the result is evident in this case. Assume inductively that $u_1, \ldots, u_{j-1}$ are well defined and orthonormal and that $\operatorname{span}\{v_1, \ldots, v_{j-1}\} = \operatorname{span}\{u_1, \ldots, u_{j-1}\}$. Proposition 3.3 shows that $u'_j$ is orthogonal to $u_1, \ldots, u_{j-1}$. If $u'_j = 0$, then $v_j$ has to be in $\operatorname{span}\{u_1, \ldots, u_{j-1}\} = \operatorname{span}\{v_1, \ldots, v_{j-1}\}$, and we have a contradiction to the assumed linear independence of $\{v_1, \ldots, v_k\}$. Thus $u'_j \neq 0$, and $\{u_1, \ldots, u_j\}$ is a well-defined orthonormal set. This set must be linearly independent, and hence its linear span is a $j$-dimensional vector subspace of the linear span of $\{v_1, \ldots, v_j\}$. By Corollary 2.4, the two linear spans coincide. This completes the induction and the proof. □

**Corollary 3.6.** If $V$ is a finite-dimensional inner-product space, then any orthonormal set in a vector subspace $S$ of $V$ can be extended to an orthonormal basis of $S$.

PROOF. Extend the given orthonormal set to a basis of $S$ by Corollary 2.3b. Then apply the Gram–Schmidt orthogonalization process. The given vectors do not get changed by the process, as we see from the formulas for the vectors $u'_j$ and $u_j$, and hence the result is an extension of the given orthonormal set to an orthonormal basis. □

**Corollary 3.7.** If $S$ is a vector subspace of a finite-dimensional inner-product space $V$, then $S$ has an orthonormal basis.

PROOF. This is the special case of Corollary 3.6 in which the given orthonormal set is empty. □

The set of all vectors orthogonal to a subset $M$ of the inner-product space $V$ is denoted by $M^\perp$. In symbols,

$$M^\perp = \{u \in V \mid (u, v) = 0 \text{ for all } v \in M\}.$$

We see by inspection that $M^\perp$ is a vector subspace. Moreover, $M \cap M^\perp = 0$ since any $u$ in $M \cap M^\perp$ must have $(u, u) = 0$. The interest in the vector subspace $M^\perp$ comes from the following proposition.

**Theorem 3.8** (Projection Theorem). If $S$ is a vector subspace of the finite-dimensional inner-product space $V$, then every $v$ in $V$ decomposes uniquely as $v = v_1 + v_2$ with $v_1$ in $S$ and $v_2$ in $S^\perp$. In other words, $V = S \oplus S^\perp$.

REMARKS. Because of this proposition, $S^\perp$ is often called the **orthogonal complement** of the vector subspace $S$.

PROOF. Uniqueness follows from the fact that $S \cap S^\perp = 0$. For existence, use of Corollaries 3.7 and 3.6 produces an orthonormal basis $\{u_1, \dots, u_r\}$ of $S$ and extends it to an orthonormal basis $\{u_1, \dots, u_n\}$ of $V$. The vectors $u_j$ for $j > r$ are orthogonal to each $u_i$ with $i \le r$ and hence are in $S^\perp$. If $v$ is given in $S$, we can write $v = \sum_{j=1}^n u_j$ as $v = v_1 + v_2$ with $v_1 = \sum_{i=1}^r (v, u_i)u_i$ and $v_2 = \sum_{j=r+1}^n (v, u_j)u_j$, and this decomposition for all $v$ shows that $V = S + S^\perp$. □

**Corollary 3.9.** If $S$ is a vector subspace of the finite-dimensional inner-product space $V$, then

(a) $\dim V = \dim S + \dim S^\perp$,
(b) $S^{\perp\perp} = S$.

PROOF. Conclusion (a) is immediate from the direct-sum decomposition $V = S \oplus S^\perp$ of Theorem 3.8. For (b), the definition of orthogonal complement gives $S \subseteq S^{\perp\perp}$. On the other hand, application of (a) twice shows that $S$ and $S^{\perp\perp}$ have the same finite dimension. By Corollary 2.4, $S^{\perp\perp} = S$. □

Section II.6 introduced "projection" mappings in the setting of any direct sum of two vector spaces, and we shall use those mappings in connection with the decomposition $V = S \oplus S^\perp$ of Theorem 3.8. We make one adjustment in working with the projections, changing their ranges from the image, namely $S$ or $S^\perp$, to the larger space $V$. In effect, a linear map $p_1$ or $p_2$ as in Section II.6 will be replaced by $i_1 p_1$ or $i_2 p_2$.

Specifically let $E : V \to V$ be the linear map that is the identity on $S$ and is 0 on $S^\perp$. Then $E$ is called the **orthogonal projection** of $V$ on $S$. The linear map $I - E$ is the identity on $S^\perp$ and is 0 on $S$. Since $S = S^{\perp\perp}$, $I - E$ is the orthogonal projection of $V$ on $S^\perp$. It is the linear map that picks out the $S^\perp$ component relative to the direct-sum decomposition $V = S^\perp \oplus S^{\perp\perp}$. Proposition 3.3 and Corollary 3.4 can be restated in terms of orthogonal projections.

**Corollary 3.10.** Let $V$ be a finite-dimensional inner-product space, let $S$ be a vector subspace of $V$, let $\{u_1, \dots, u_k\}$ be an orthonormal basis of $S$, and let $E$ be the orthogonal projection of $V$ on $S$. If $v$ is in $V$, then

$$E(v) = \sum_{j=1}^k (v, u_j)u_j$$

and

$$\|E(v)\|^2 = \sum_{j=1}^k |(v, u_j)|^2.$$

The vector $v^\perp$ in the expansion $v = \sum_{j=1}^k (v, u_j)u_j + v^\perp$ of Proposition 3.3 is equal to $(I - E)v$, and the equality of norms

$$\|v\|^2 = \sum_{j=1}^k |(v, u_j)|^2 + \|v^\perp\|^2$$

has the interpretations that

$$\|v\|^2 = \|E(v)\|^2 + \|(I - E)v\|^2$$

and that equality holds in Bessel's inequality if and only if $E(v) = v$.

PROOF. Write $v = \sum_{j=1}^k (v, u_j)u_j + v^\perp$ as in Proposition 3.3. Then each $u_j$ is in $S$, and the vector $v^\perp$, being orthogonal to each member of a basis of $S$, is in $S^\perp$. This proves the formula for $E(v)$, and the formula for $\|E(v)\|^2$ follows by applying Corollary 3.4 to $v - v^\perp$.

Reassembling $v$, we now have $v = E(v) + v^\perp$, and hence $v^\perp = v - E(v) = (I - E)v$. Finally the decomposition $v = E(v) + (I - E)(v)$ is into orthogonal terms, and the Pythagorean Theorem shows that $\|v\|^2 = \|E(v)\|^2 + \|(I - E)v\|^2$. □

**Theorem 3.11** (Parseval's equality). If $V$ is a finite-dimensional inner-product space, then the following conditions on an orthonormal set $\{u_1, \dots, u_m\}$ are equivalent:

(a) $\{u_1, \dots, u_m\}$ is a vector-space basis of $V$, hence an orthonormal basis,
(b) the only vector orthogonal to all of $u_1, \dots, u_m$ is 0,
(c) $v = \sum_{j=1}^m (v, u_j)u_j$ for all $v$ in $V$,
(d) $\|v\|^2 = \sum_{j=1}^m |(v, u_j)|^2$ for all $v$ in $V$,
(e) $(v, w) = \sum_{j=1}^m (v, u_j)\overline{(w, u_j)}$ for all $v$ and $w$ in $V$.

PROOF. Let $S = \operatorname{span}\{u_1, \dots, u_m\}$, and let $E$ be the orthogonal projection of $V$ on $S$. If (a) holds, then $S = V$ and $S^\perp = 0$. Thus (b) holds.

If (b) holds, then $S^\perp = 0$ and $E$ is the identity. Thus (c) holds by Corollary 3.10.

If (c) holds, then Corollary 3.4 shows that (d) holds.

If (d) holds, we use polarization to prove (e). Let $k$ be in $\{0, 2\}$ if $\mathbb{F} = \mathbb{R}$, or in $\{0, 1, 2, 3\}$ if $\mathbb{F} = \mathbb{C}$. Conclusion (d) gives us

$$\|v + i^k w\|^2 = \sum_{j=1}^m |(v + i^k w, u_j)|^2 = \|v\|^2 + \sum_{j=1}^m 2\operatorname{Re}\big((v, u_j)\overline{i^k(w, u_j)}\big) + \|w\|^2.$$

Multiplying by $i^k$ and summing over $k$, we obtain

$$4(v, w) = 2\sum_{j=1}^m \sum_k i^k \operatorname{Re}\big((-i)^k (v, u_j)\overline{(w, u_j)}\big).$$

In the proof of polarization, we saw that $2\sum_k i^k \operatorname{Re}((-i)^k z) = 4z$. Hence $4(v, w) = 4\sum_{j=1}^m (v, u_j)\overline{(w, u_j)}$. This proves (e).

If (e) holds, we take $w = v$ in (e) and apply Corollary 3.10 to see that $\|E(v)\|^2 = \|v\|^2$ for all $v$. Then $\|(I - E)v\|^2 = 0$ for all $v$, and $E(v) = v$ for all $v$. Hence $S = V$, and $\{u_1, \dots, u_m\}$ is a basis. This proves (a). □

**Theorem 3.12** (Riesz Representation Theorem). If $\ell$ is a linear functional on the finite-dimensional inner-product space $V$, then there exists a unique $v$ in $V$ with $\ell(u) = (u, v)$ for all $u$ in $V$.

PROOF. Uniqueness is immediate by subtracting two such expressions, since if $(u, v) = 0$ for all $u$, then the special case $u = v$ gives $(v, v) = 0$ and $v = 0$. Let us prove existence. If $\ell = 0$, take $v = 0$. Otherwise let $S = \ker \ell$. Corollary 2.15 shows that $\dim S = \dim V - 1$, and Corollary 3.9a then shows that $\dim S^\perp = 1$. Let $w$ be a nonzero vector in $S^\perp$. This vector $w$ must have $\ell(w) \neq 0$ since $S \cap S^\perp = 0$, and we let $v$ be the member of $S^\perp$ given by

$$v = \frac{\overline{\ell(w)}}{\|w\|^2}\, w.$$

For any $u$ in $V$, we have $\ell\big(u - \frac{\ell(u)}{\ell(w)}\, w\big) = 0$, and hence $u - \frac{\ell(u)}{\ell(w)}\, w$ is in $S$. Since $v$ is in $S^\perp$, $u - \frac{\ell(u)}{\ell(w)}\, w$ is orthogonal to $v$. Thus

$$(u, v) = \Big(\frac{\ell(u)}{\ell(w)}\, w,\ v\Big) = \Big(\frac{\ell(u)}{\ell(w)}\, w,\ \frac{\overline{\ell(w)}}{\|w\|^2}\, w\Big) = \ell(u)\, \frac{\ell(w)}{\ell(w)}\, \frac{\|w\|^2}{\|w\|^2} = \ell(u).$$

This proves existence. □

## 2. Adjoints

Throughout this section, $V$ will denote a finite-dimensional inner-product space with inner product $(\,\cdot\,,\,\cdot\,)$ and with scalars from $\mathbb{F}$, with $\mathbb{F}$ equal to $\mathbb{R}$ or $\mathbb{C}$. We shall study aspects of linear maps $L : V \to V$ related to the inner product on $V$. The starting point is to associate to any such $L$ another linear map $L^* : V \to V$ known as the "adjoint" of $V$, and then to investigate some of its properties. A tool in this investigation will be the scalar-valued function on $V \times V$ given by $(u, v) \mapsto (L(u), v)$, which captures the information in any matrix of $L$ without requiring the choice of an ordered basis. This function determines $L$ uniquely because an equality $(L(u), v) = (L'(u), v)$ for all $u$ and $v$ implies $(L(u) - L'(u), v) = 0$ for all $u$ and $v$, in particular for $v = L(u) - L'(u)$; thus $\|L(u) - L'(u)\|^2 = 0$ and $L(u) = L'(u)$ for all $u$.

**Proposition 3.13.** Let $L : V \to V$ be a linear map on the finite-dimensional inner-product space $V$. For each $u$ in $V$, there exists a unique vector $L^*(u)$ in $V$ such that

$$(L(v), u) = (v, L^*(u)) \qquad \text{for all } v \text{ in } V.$$

As $u$ varies, this formula defines $L^*$ as a linear map from $V$ to $V$.

REMARK. The linear map $L^* : V \to V$ is called the **adjoint** of $L$.

PROOF. The function $v \mapsto (L(v), u)$ is a linear functional on $V$, and Theorem 3.12 shows that it is given by the inner product with a unique vector of $V$. Thus we define $L^*(u)$ to be the unique vector of $V$ with $(L(v), u) = (v, L^*(u))$ for all $v$ in $V$.

If $c$ is a scalar, then the uniqueness and the computation $(v, L^*(cu)) = (L(v), cu) = \bar{c}(L(v), u) = \bar{c}(v, L^*(u)) = (v, cL^*(u))$ yield $L^*(cu) = cL^*(u)$. Similarly the uniqueness and the computation

$$\begin{aligned}(v, L^*(u_1 + u_2)) &= (L(v), u_1 + u_2) = (L(v), u_1) + (L(v), u_2) \\ &= (v, L^*(u_1)) + (v, L^*(u_2)) = (v, L^*(u_1) + L^*(u_2))\end{aligned}$$

yield $L^*(u_1 + u_2) = L^*(u_1) + L^*(u_2)$. Therefore $L^*$ is linear. □

The passage $L \mapsto L^*$ to the adjoint is a function from $\mathrm{Hom}_{\mathbb{F}}(V, V)$ to itself that is conjugate linear, and it reverses the order of multiplication: $(L_1 L_2)^* = L_2^* L_1^*$. Since the formula $(L(v), u) = (v, L^*(u))$ in the proposition is equivalent to the formula $(u, L(v)) = (L^*(u), v)$, we see that $L^{**} = L$.

All of the results in Section II.3 concerning the association of matrices to linear maps are applicable here, but our interest now will be in what happens when the bases we use are orthonormal. Recall from Section II.3 that if $\Gamma = (u_1, \dots, u_n)$ and $\Delta = (v_1, \dots, v_n)$ are any ordered bases of $V$, then the matrix $A = \begin{pmatrix} L \\ \Delta\Gamma \end{pmatrix}$ associated to the linear map $L : V \to V$ has $A_{ij} = \begin{pmatrix} L(u_j) \\ \Delta \end{pmatrix}_i$.

**Lemma 3.14.** If $L : V \to V$ is a linear map on the finite-dimensional inner-product space $V$ and if $\Gamma = (u_1, \dots, u_n)$ and $\Delta = (v_1, \dots, v_n)$ are ordered orthonormal bases of $V$, then the the matrix $A = \begin{pmatrix} L \\ \Delta\Gamma \end{pmatrix}$ has $A_{ij} = (L(u_j), v_i)$.

PROOF. Applying Theorem 3.11c, we have

$$\begin{aligned}A_{ij} &= \begin{pmatrix} L(u_j) \\ \Delta \end{pmatrix}_i = \begin{pmatrix} \sum_{i'} (L(u_j), v_{i'}) v_{i'} \\ \Delta \end{pmatrix}_i \\ &= \sum_{i'} (L(u_j), v_{i'}) \begin{pmatrix} v_{i'} \\ \Delta \end{pmatrix}_i = \sum_{i'} (L(u_j), v_{i'}) \delta_{ii'} = (L(u_j), v_i). \qquad \square\end{aligned}$$

**Proposition 3.15.** If $L : V \to V$ is a linear map on the finite-dimensional inner-product space $V$ and if $\Gamma = (u_1, \dots, u_n)$ and $\Delta = (v_1, \dots, v_n)$ are ordered orthonormal bases of $V$, then the matrices $A = \begin{pmatrix} L \\ \Delta\Gamma \end{pmatrix}$ and $A^* = \begin{pmatrix} L^* \\ \Gamma\Delta \end{pmatrix}$ of $L$ and its adjoint are related by $A^*_{ij} = \overline{A_{ji}}$.

PROOF. Lemma 3.14 and the definition of $L^*$ give $A^*_{ij} = (L^*(v_j), u_i) = (v_j, L(u_i)) = \overline{(L(u_i), v_j)} = \overline{A_{ji}}$. □

Accordingly, we define $A^* = \overline{A}^t$ for any square matrix $A$, sometimes calling $A^*$ the **adjoint**[6] of $A$.

A linear map $L : V \to V$ is called **self-adjoint** if $L^* = L$. Correspondingly a square matrix $A$ is **self-adjoint** if $A^* = A$. It is more common, however, to say that a matrix with $A^* = A$ is **symmetric** if $\mathbb{F} = \mathbb{R}$ or **Hermitian**[7] if $\mathbb{F} = \mathbb{C}$. A real Hermitian matrix is symmetric, and the term "Hermitian" is thus applicable also when $\mathbb{F} = \mathbb{R}$.

Any Hermitian matrix $A$ arises from a self-adjoint linear map $L$. Namely, we take $V$ to be $\mathbb{F}^n$ with the usual inner product, and we let $\Gamma$ and $\Delta$ each be the standard ordered basis $\Sigma = (e_1, \dots, e_n)$. This basis is orthonormal, and we define $L$ by the matrix product $L(v) = Av$ for any column vector $v$. We know that $\begin{pmatrix} L \\ \Sigma\Sigma \end{pmatrix} = A$. Since $A^* = A$, we conclude from Proposition 3.15 that $L^* = L$. Thus we are free to deduce properties of Hermitian matrices from properties of self-adjoint linear maps.

Self-adjoint linear maps will be of special interest to us. Nontrivial examples of self-adjoint linear maps, constructed without simply writing down Hermitian matrices, may be produced by the following proposition.

**Proposition 3.16.** If $V$ is a finite-dimensional inner-product space and $S$ is a vector subspace of $V$, then the orthogonal projection $E : V \to V$ of $V$ on $S$ is self-adjoint.

PROOF. Let $v = v_1 + v_2$ and $u = u_1 + u_2$ be the decompositions of two members of $V$ according to $V = S \oplus S^\perp$. Then we have $(v, E^*(u)) = (E(v), u) = (v_1, u_1 + u_2) = (v_1, u_1) = (v, u_1) = (v, E(u))$, and the proposition follows by the uniqueness in Proposition 3.13. □

[6]The name "adjoint" happens to coincide with the name for a different notion that arose in connection with Cramer's rule in Section II.7. The two notions never seem to arise at the same time, and thus no confusion need occur.

[7]The term "Hermitian" is used also for a class of linear maps in the infinite-dimensional case, but care is needed because the terms "Hermitian" and "self-adjoint" mean different things in the infinite-dimensional case.

To understand Proposition 3.16 in terms of matrices, take an ordered orthonormal basis $(u_1, \ldots, u_r)$ of $S$, and extend it to an ordered orthonormal basis $\Gamma = (u_1, \ldots, u_n)$ of $V$. Then

$$E(u_j) = \begin{cases} u_j & \text{for } j \le r, \\ 0 & \text{for } j > r, \end{cases}$$

and hence $\begin{pmatrix} E(u_j) \\ \Gamma \end{pmatrix}$ equals the $j^{\text{th}}$ standard basis vector $e_j$ if $j \le r$ and equals 0 if $j > r$. Consequently the matrix $\begin{pmatrix} E \\ \Gamma\Gamma \end{pmatrix}$ is diagonal with 1's in the first $r$ diagonal entries and 0's elsewhere. This matrix is equal to its conjugate transpose, as it must be according to Propositions 3.15 and 3.16.

**Proposition 3.17.** If $V$ is a finite-dimensional inner-product space and $L : V \to V$ is a self-adjoint linear map, then $(L(v), v)$ is in $\mathbb{R}$ for every $v$ in $V$, and consequently every eigenvalue of $L$ is in $\mathbb{R}$. Conversely if $\mathbb{F} = \mathbb{C}$ and if $L : V \to V$ is a linear map such that $(L(v), v)$ is in $\mathbb{R}$ for every $v$ in $V$, then $L$ is self-adjoint.

REMARK. The hypothesis $\mathbb{F} = \mathbb{C}$ is essential in the converse. In fact, the $90^\circ$ rotation $L$ of $\mathbb{R}^2$ whose matrix in the standard basis is $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is not self-adjoint but does have $L(v) \cdot v = 0$ for every $v$ in $\mathbb{R}^2$.

PROOF. If $L = L^*$, then $(L(v), v) = (v, L^*(v)) = (v, L(v)) = \overline{(L(v), v)}$, and hence $(L(v), v)$ is real-valued. If $v$ is an eigenvector with eigenvalue $\lambda$, then substitution of $L(v) = \lambda v$ into $(L(v), v) = \overline{(L(v), v)}$ gives $\lambda \|v\|^2 = \bar{\lambda} \|v\|^2$. Since $v \neq 0$, $\lambda$ must be real.

For the converse we begin with the special case that $(L(w), w) = 0$ for all $w$. For $0 \le k \le 3$, we then have

$$(-i)^k (L(u), v) + i^k (L(v), u) = (L(u + i^k v), u + i^k v) - (L(u), u) - (L(v), v) = 0.$$

Taking $k = 0$ gives $(L(u), v) + (L(v), u) = 0$, while taking $k = 1$ gives $(L(u), v) - (L(v), u) = 0$. Hence $(L(u), v) = 0$ for all $u$ and $v$. Since the function $(u, v) \mapsto L(u, v)$ determines $L$, we obtain $L = 0$.

In the general case, $(L(v), v)$ real-valued implies that $(L(v), v) = (L^*(v), v)$ for all $v$. Therefore $((L - L^*)(v), v) = 0$ for all $v$, and the special case shows that $L - L^* = 0$. This completes the proof. $\square$

We conclude this section by examining one further class of linear maps having a special relationship with their adjoints.

**Proposition 3.18.** If $V$ is a finite-dimensional inner-product space, then the following conditions on a linear map $L : V \to V$ are equivalent:

(a) $L^*L = I$,
(b) $L$ carries some orthonormal basis of $V$ to an orthonormal basis,
(c) $L$ carries each orthonormal basis of $V$ to an orthonormal basis,
(d) $(L(u), L(v)) = (u, v)$ for all $u$ and $v$ in $V$,
(e) $\|L(v)\| = \|v\|$ for all $v$ in $V$.

REMARK. A linear map satisfying these equivalent conditions is said to be **orthogonal** if $\mathbb{F} = \mathbb{R}$ and **unitary** if $\mathbb{F} = \mathbb{C}$.

PROOF. We prove that (a), (d), and (e) are equivalent and that (b), (c), and (d) are equivalent.

If (a) holds and $u$ and $v$ are given in $V$, then $(L(u), L(v)) = (L^*L(u), v) = (I(u), v) = (u, v)$, and (d) holds. If (d) holds, then setting $u = v$ shows that (e) holds. If (e) holds, we use polarization twice to write

$$\begin{aligned}(L(u), L(v)) &= \textstyle\sum_k \frac{1}{4} i^k \|L(u) + i^k L(v)\|^2 = \sum_k \frac{1}{4} i^k \|L(u + i^k v)\|^2 \\ &= \textstyle\sum_k \frac{1}{4} i^k \|u + i^k v\|^2 = (u, v).\end{aligned}$$

Then $((L^*L - I)(u), v) = 0$ for all $u$ and $v$, and we conclude that (a) holds.

Since (b) is a special case of (c) and (c) is a special case of (d), proving that (b) implies (d) will prove that (b), (c), and (d) are equivalent. Thus let $\{u_1, \dots, u_n\}$ be an orthonormal basis of $V$ such that $\{L(u_1), \dots, L(u_n)\}$ is an orthonormal basis, and let $u$ and $v$ be given. Then

$$\begin{aligned}(L(u), L(v)) &= \Big(L\big(\textstyle\sum_i (u, u_i)u_i\big),\ L\big(\sum_j (v, u_j)u_j\big)\Big) \\ &= \textstyle\sum_{i,j} (u, u_i)\overline{(v, u_j)}(L(u_i), L(u_j)) \\ &= \textstyle\sum_{i,j} (u, u_i)\overline{(v, u_j)}\delta_{ij} = \sum_i (u, u_i)\overline{(v, u_i)} = (u, v),\end{aligned}$$

the last equality following from Parseval's equality (Theorem 3.11). □

As with self-adjointness, we use the geometrically meaningful definition for linear maps to obtain a definition for matrices: a square matrix $A$ with $A^*A = I$ is said to be **orthogonal** if $\mathbb{F} = \mathbb{R}$ and **unitary** if $\mathbb{F} = \mathbb{C}$. The condition is that $A$ is invertible and its inverse equals its adjoint. In terms of individual entries, the condition is that $\sum_k A^*_{ik} A_{kj} = \delta_{ij}$, hence that $\sum_k \overline{A}_{ki} A_{kj} = \delta_{ij}$. This is the condition that the columns of $A$ form an orthonormal basis relative to the usual inner product on $\mathbb{R}^n$ or $\mathbb{C}^n$. A real unitary matrix is orthogonal.

If $A$ is an orthogonal or unitary matrix, we can construct a corresponding orthogonal or unitary linear map on $\mathbb{R}^n$ or $\mathbb{C}^n$ relative to the standard ordered

basis $\Sigma$. Namely, we define $L(v) = Av$, and Proposition 3.15 shows that $L$ is orthogonal or unitary: $L^*L(v) = A^*Av = Iv = v$. Proposition 3.19 below gives a converse.

Let us notice that an orthogonal or unitary matrix $A$ necessarily has $|\det A| = 1$. In fact, the formula $A^* = (\overline{A})^t$ implies that $\det A^* = \overline{\det A}$. Then

$$1 = \det I = \det A^*A = \det A^* \det A = \overline{\det A} \det A = |\det A|^2.$$

An orthogonal matrix thus has determinant $\pm 1$, while we conclude for a unitary matrix only that the determinant is a complex number of absolute value 1.

EXAMPLES.

(1) The 2-by-2 orthogonal matrices of determinant $+1$ are all matrices of the form $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$. The 2-by-2 orthogonal matrices of determinant $-1$ are the product of $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and the 2-by-2 orthogonal matrices of determinant $+1$.

(2) The 2-by-2 unitary matrices of determinant $+1$ are all matrices of the form $\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}$ with $|\alpha|^2 + |\beta|^2 = 1$; these may be regarded as parametrizing the points of the unit sphere $S^3$ of $\mathbb{R}^4$. The 2-by-2 unitary matrices of arbitrary determinant are the products of all matrices $\begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}$ and the 2-by-2 unitary matrices of determinant $+1$.

**Proposition 3.19.** If $V$ is a finite-dimensional inner-product space, if $\Gamma = (u_1, \dots, u_n)$ and $\Delta = (v_1, \dots, v_n)$ are ordered orthonormal bases of $V$, and if $L : V \to V$ is a linear map that is orthogonal if $\mathbb{F} = \mathbb{R}$ and unitary if $\mathbb{F} = \mathbb{C}$, then the matrix $A = \begin{pmatrix} L \\ \Delta\Gamma \end{pmatrix}$ is orthogonal or unitary.

PROOF. Proposition 3.15 and Theorem 2.16 give $A^*A = \begin{pmatrix} L^* \\ \Gamma\Delta \end{pmatrix} \begin{pmatrix} L \\ \Delta\Gamma \end{pmatrix} = \begin{pmatrix} I \\ \Delta\Delta \end{pmatrix}$, and the right side is the identity matrix, as required. □

One consequence of Proposition 3.19 is that any matrix $\begin{pmatrix} I \\ \Delta\Gamma \end{pmatrix}$ relative to two ordered orthonormal bases is orthogonal or unitary, since the identity function $I : V \to V$ is certainly orthogonal or unitary. Thus a change from writing the matrix of a linear map $L$ in one ordered orthonormal basis $\Gamma$ to writing the matrix of $L$ in another ordered orthonormal basis $\Delta$ is implemented by the formula $\begin{pmatrix} L \\ \Gamma\Gamma \end{pmatrix} = C^{-1} \begin{pmatrix} L \\ \Delta\Delta \end{pmatrix} C$, where $C$ is the orthogonal or unitary matrix $\begin{pmatrix} I \\ \Delta\Gamma \end{pmatrix}$.

Another consequence of Proposition 3.19 is that the matrix $\begin{pmatrix} L \\ \Gamma\Gamma \end{pmatrix}$ of an orthogonal or unitary linear map $L$ in an ordered orthonormal basis $\Gamma$ is an orthogonal or unitary matrix. We have defined $\det L$ to be the determinant of $\begin{pmatrix} L \\ \Gamma\Gamma \end{pmatrix}$ relative to *any* $\Gamma$, and we conclude that $|\det L| = 1$.

## 3. Spectral Theorem

In this section we deal with the geometric structure of certain kinds of linear maps from finite-dimensional inner-product spaces into themselves. We shall see that linear maps that are self-adjoint or unitary, among other possible conditions, have bases of eigenvectors in the sense of Section II.8. Moreover, such a basis may be taken to be orthonormal. When an ordered basis of eigenvectors is used for expressing the linear map as a matrix, the result is that the matrix is diagonal. Thus these linear maps have an especially uncomplicated structure. In terms of matrices, the result is that a Hermitian or unitary matrix $A$ is similar to a diagonal matrix $D$, and the matrix $C$ with $D = C^{-1}AC$ may be taken to be unitary. We begin with a lemma.

**Lemma 3.20.** If $L : V \to V$ is a self-adjoint linear map on an inner-product space $V$, then $v \mapsto (L(v), v)$ is real-valued, every eigenvalue of $L$ is real, eigenvectors under $L$ for distinct eigenvalues are orthogonal, and every vector subspace $S$ of $V$ with $L(S) \subseteq S$ has $L(S^\perp) \subseteq S^\perp$.

PROOF. The first two conclusions are contained in Proposition 3.17. If $v_1$ and $v_2$ are eigenvectors of $L$ with distinct real eigenvalues $\lambda_1$ and $\lambda_2$, then

$$(\lambda_1 - \lambda_2)(v_1, v_2) = (\lambda_1 v_1, v_2) - (v_1, \lambda_2 v_2) = (L(v_1), v_2) - (v_1, L(v_2)) = 0.$$

Since $\lambda_1 \neq \lambda_2$, we must have $(v_1, v_2) = 0$. If $S$ is a vector subspace with $L(S) \subseteq S$, then also $L(S^\perp) \subseteq S^\perp$ because $s \in S$ and $s^\perp \in S^\perp$ together imply

$$0 = (L(s), s^\perp) = (s, L(s^\perp)). \qquad \square$$

**Theorem 3.21** (Spectral Theorem). Let $L : V \to V$ be a self-adjoint linear map on an inner-product space $V$. Then $V$ has an orthonormal basis of eigenvectors of $L$. In addition, for each scalar $\lambda$, let

$$V_\lambda = \{v \in V \mid L(v) = \lambda v\},$$

so that $V_\lambda$ when nonzero is the eigenspace of $L$ for the eigenvalue $\lambda$. Then the eigenvalues of $L$ are all real, the vector subspaces $V_\lambda$ are mutually orthogonal,

and any orthonormal basis of $V$ of eigenvectors of $L$ is the union of orthonormal bases of the $V_\lambda$'s. Correspondingly if $A$ is any Hermitian $n$-by-$n$ matrix, then there exists a unitary matrix $C$ such that $C^{-1}AC$ is diagonal with real entries. If the matrix $A$ has real entries, then $C$ may be taken to be an orthogonal matrix.

PROOF. Lemma 3.20 shows that the eigenvalues of $L$ are all real and that the vector subspaces $V_\lambda$ are mutually orthogonal.

To proceed further, we first assume that $\mathbb{F} = \mathbb{C}$. Applying the Fundamental Theorem of Algebra (Theorem 1.18) to the characteristic polynomial of $L$, we see that $L$ has at least one eigenvalue, say $\lambda_1$. Then $L(V_{\lambda_1}) \subseteq V_{\lambda_1}$, and Lemma 3.20 shows that $L((V_{\lambda_1})^\perp) \subseteq (V_{\lambda_1})^\perp$. The vector subspace $(V_{\lambda_1})^\perp$ is an inner-product space, and the claim is that $L\big|_{(V_{\lambda_1})^\perp}$ is self-adjoint. In fact, if $v_1$ and $v_2$ are in $(V_{\lambda_1})^\perp$, then

$$\begin{aligned}\big((L\big|_{(V_{\lambda_1})^\perp})^*(v_1), v_2\big) &= \big(v_1, L\big|_{(V_{\lambda_1})^\perp}(v_2)\big) = (v_1, L(v_2)) \\ &= (L(v_1), v_2) = \big(L\big|_{(V_{\lambda_1})^\perp}(v_1), v_2\big),\end{aligned}$$

and the claim is proved. Since $\lambda_1$ is an eigenvalue of $L$, $\dim(V_{\lambda_1})^\perp < \dim V$. Therefore we can now set up an induction that ultimately exhibits $V$ as an orthogonal direct sum $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$. If $v$ is an eigenvector of $L$ with eigenvalue $\lambda'$, then either $\lambda' = \lambda_j$ for some $j$ in this decomposition, in which case $v$ is in $V_{\lambda_j}$, or $\lambda'$ is not equal to any $\lambda_j$, in which case $v$, by the lemma, is orthogonal to all vectors in $V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$, hence to all vectors in $V$; being orthogonal to all vectors in $V$, $v$ must be 0. Choosing an orthonormal basis for each $V_{\lambda_j}$ and taking their union provides an orthonormal basis of eigenvectors and completes the proof for $L$ when $\mathbb{F} = \mathbb{C}$.

Next assume that $A$ is a Hermitian $n$-by-$n$ matrix. We define a linear map $L : \mathbb{C}^n \to \mathbb{C}^n$ by $L(v) = Av$, and we know from Proposition 3.15 that $L$ is self-adjoint. The case just proved shows that $L$ has an ordered orthonormal basis $\Gamma$ of eigenvectors, all the eigenvalues being real. If $\Sigma$ denotes the standard ordered basis of $\mathbb{C}^n$, then $D = \begin{pmatrix} L \\ \Gamma\Gamma \end{pmatrix}$ is diagonal with real entries and is equal to

$$\begin{pmatrix} I \\ \Gamma\Sigma \end{pmatrix} \begin{pmatrix} L \\ \Sigma\Sigma \end{pmatrix} \begin{pmatrix} I \\ \Sigma\Gamma \end{pmatrix} = C^{-1}AC,$$

where $C = \begin{pmatrix} L \\ \Sigma\Gamma \end{pmatrix}$. The matrix $C$ is unitary by Proposition 3.19, and the formula $D = C^{-1}AC$ shows that $A$ is as asserted.

Now let us return to $L$ and suppose that $\mathbb{F} = \mathbb{R}$. The idea is to use the same argument as above in the case that $\mathbb{F} = \mathbb{C}$, but we need a substitute for

the use of the Fundamental Theorem of Algebra. Fixing any orthonormal basis of $V$, let $A$ be the matrix of $L$. Then $A$ is Hermitian with real entries. The previous paragraph shows that any Hermitian matrix, whether or not real, has a characteristic polynomial that splits as a product $\prod_{j=1}^{m} (\lambda - r_j)^{m_j}$ with all $r_j$ real. Consequently $L$ has this property as well. Thus any self-adjoint $L$ when $\mathbb{F} = \mathbb{R}$ has an eigenvalue. Returning to the argument for $L$ above when $\mathbb{F} = \mathbb{C}$, we readily see that it now applies when $\mathbb{F} = \mathbb{R}$.

Finally if $A$ is a Hermitian matrix with real entries, then we can define a self-adjoint linear map $L : \mathbb{R}^n \to \mathbb{R}^n$ by $L(v) = Av$, obtain an orthonormal basis of eigenvectors for $L$, and argue as above to obtain $D = C^{-1}AC$, where $D$ is diagonal and $C$ is unitary. The matrix $C$ has columns that are eigenvectors in $\mathbb{R}^n$ of the associated $L$, and these have real entries. Thus $C$ is orthogonal. $\square$

An important application of the Spectral Theorem is to the formation of a square root for any "positive semidefinite" linear map. We say that a linear map $L : V \to V$ on a finite-dimensional inner-product space is **positive semidefinite** if $L^* = L$ and $(L(v), v) \geq 0$ for all $v$ in $V$. If $\mathbb{F} = \mathbb{C}$, then the condition $L^* = L$ is redundant, according to Proposition 3.17, but that fact will not be important for us. Similarly an $n$-by-$n$ matrix $A$ is **positive semidefinite** if $A^* = A$ and $\bar{x}^t Ax \geq 0$ for all column vectors $x$. An example of a positive semidefinite $n$-by-$n$ matrix is any matrix $A = B^*B$, where $B$ is an arbitrary $k$-by-$n$ matrix. In fact, if $x$ is in $\mathbb{F}^n$, then $\bar{x}^t B^* Bx = (\overline{Bx})^t (Bx)$, and the right side is $\geq 0$, being a sum of absolute values squared.

**Corollary 3.22.** Let $L : V \to V$ be a positive semidefinite linear map on a finite-dimensional inner-product space, and let $A$ be an $n$-by-$n$ Hermitian matrix. Then

(a) $L$ or $A$ is positive semidefinite if and only if all of its eigenvalues are $\geq 0$.
(b) whenever $L$ or $A$ is positive semidefinite, $L$ or $A$ is invertible if and only if $(L(v), v) > 0$ for all $v \neq 0$ or $\bar{x}^t Ax > 0$ for all $x \neq 0$.
(c) whenever $L$ or $A$ is positive semidefinite, $L$ or $A$ has a unique positive semidefinite square root.

REMARKS. A positive semidefinite linear map or matrix satisfying the condition in (b) is said to be **positive definite**, and the content of (b) is that a positive semidefinite linear map or matrix is positive definite if and only if it is invertible.

PROOF. We apply the Spectral Theorem (Theorem 3.21). For each conclusion the result for a matrix $A$ is a special case of the result for the linear map $L$, and it is enough to treat only $L$. In (a), let $(u_1, \ldots, u_n)$ be an ordered basis of eigen-

vectors with respective eigenvalues $\lambda_1, \dots, \lambda_n$, not necessarily distinct. Then $(L(u_j), u_j) = \lambda_j$ shows the necessity of having $\lambda_j \geq 0$, while the computation

$$\begin{aligned}(L(v), v) &= \big(L\big(\textstyle\sum_i (v, u_i)u_i\big),\ \textstyle\sum_j (v, u_j)u_j\big)\\ &= \big(\textstyle\sum_i \lambda_i(v, u_i)u_i,\ \textstyle\sum_j (v, u_j)u_j\big)\\ &= \textstyle\sum_i \lambda_i |(v, u_i)|^2\end{aligned}$$

shows the sufficiency.

In (b), if $L$ fails to be invertible, then 0 is an eigenvalue for some eigenvector $v \neq 0$, and $v$ has $(L(v), v) = 0$. Conversely if $L$ is invertible, then all the eigenvalues $\lambda_i$ are $> 0$ by (a), and the computation in (a) yields

$$(L(v), v) = \sum_i \lambda_i |(v, u_i)|^2 \geq \big(\min_j \lambda_j\big) \sum_i |(v, u_i)|^2 = \big(\min_j \lambda_j\big)\|v\|^2,$$

the last step following from Parseval's equality (Theorem 3.11).

For existence in (c), the Spectral Theorem says that there exists an ordered orthonormal basis $\Gamma = (u_1, \dots, u_n)$ of eigenvectors of $L$, say with respective eigenvalues $\lambda_1, \dots, \lambda_n$. The eigenvalues are all $\geq 0$ by (a). The linear extension of the function $P$ with $P(u_j) = \lambda_j^{1/2} u_j$ is given by

$$P(v) = \sum_{j=1}^{n} \lambda_j^{1/2} (v, u_j) u_j,$$

and it has

$$P^2(v) = \textstyle\sum_j \lambda_j (v, u_j) u_j = \sum_j (v, u_j) L(u_j) = L\big(\sum_j (v, u_j) u_j\big) = L(v).$$

Thus $P^2 = L$. Relative to $\Gamma$, we have

$$\begin{pmatrix} P \\ \Gamma\Gamma \end{pmatrix}_{ij} = \big((P(u_j), u_1)u_1 + \cdots + (P(u_j), u_n)u_n\big)_i = (P(u_j), u_i) = \lambda_j^{1/2} \delta_{ij},$$

and this is a Hermitian matrix; Proposition 3.15 therefore shows that $P^* = P$. Finally

$$(P(v), v) = \big(\textstyle\sum_i \lambda_i^{1/2} (v, u_i) u_i, \sum_j (v, u_j) u_j\big) = \lambda_i^{1/2} |(v, u_i)|^2 \geq 0,$$

and thus $P$ is positive semidefinite. This proves existence.

For uniqueness in (c), let $P$ satisfy $P^* = P$ and $P^2 = L$, and suppose $P$ is positive semidefinite. Choose an orthonormal basis of eigenvectors $u_1, \dots, u_n$ of $P$, say with eigenvalues $c_1, \dots, c_n$, all $\geq 0$. Then $L(u_j) = P^2(u_j) = c_j^2 u_j$, and we see that $u_1, \dots, u_n$ form an orthonormal basis of eigenvectors of $L$ with eigenvalues $c_j^2$. On the space where $L$ acts as the scalar $\lambda_i$, $P$ must therefore act as the scalar $\lambda_i^{1/2}$. We conclude that $P$ is unique. $\square$

The technique of proof of (c) allows one, more generally, to define $f(L)$ for any function $f : \mathbb{R} \to \mathbb{C}$ whenever $L$ is self-adjoint. Actually, the function $f$ needs to be defined only on the set of eigenvalues of $L$ for the definition to make sense.

At the end of this section, we shall use the existence of the square root in (c) to obtain the so-called "polar decomposition" of square matrices. But before doing that, let us mine three additional easy consequences of the Spectral Theorem. The first deals with several self-adjoint linear maps rather than one, and the other two apply that conclusion to deal with single linear maps that are not necessarily self-adjoint.

**Corollary 3.23.** Let $V$ be a finite-dimensional inner-product space, and let $L_1, \dots, L_m$ be self-adjoint linear maps from $V$ to $V$ that commute in the sense that $L_i L_j = L_j L_i$ for all $i$ and $j$. Then $V$ has an orthonormal basis of simultaneous eigenvectors of $L_1, \dots, L_m$. In addition, for each $m$-tuple of scalars $\lambda_1, \dots, \lambda_m$, let

$$V_{\lambda_1,\dots,\lambda_m} = \{v \in V \mid L_j(v) = \lambda_j v \ \text{ for } 1 \le j \le m\}$$

consist of 0 and the simultaneous eigenvectors of $L_1, \dots, L_m$ corresponding to $\lambda_1, \dots, \lambda_m$. Then all the eigenvalues $\lambda_j$ are real, the vector subspaces $V_{\lambda_1,\dots,\lambda_m}$ are mutually orthogonal, and any orthonormal basis of $V$ of simultaneous eigenvectors of $L_1, \dots, L_m$ is the union of orthonormal bases of the $V_{\lambda_1,\dots,\lambda_m}$'s. Correspondingly if $A_1, \dots, A_m$ are commuting Hermitian $n$-by-$n$ matrices, then there exists a unitary matrix $C$ such that $C^{-1} A_j C$ is diagonal with real entries for all $j$. If all the matrices $A_j$ have real entries, then $C$ may be taken to be an orthogonal matrix.

PROOF. This follows by iterating the Spectral Theorem (Theorem 3.21). In fact, let $\{V_{\lambda_1}\}$ be the system of vector subspaces produced by the theorem for $L_1$. For each $j$, the commutativity of the linear maps $L_i$ forces

$$L_1(L_i(v)) = L_i(L_1(v)) = L_i(\lambda_1 v) = \lambda_1 L_i(v) \qquad \text{for } v \in V_{\lambda_1},$$

and thus $L_i(V_{\lambda_1}) \subseteq V_{\lambda_1}$. The restrictions of $L_1, \dots, L_m$ to $V_{\lambda_1}$ are self-adjoint and commute. Let $\{V_{\lambda_1,\lambda_2}\}$ be the system of vector subspaces produced by the Spectral Theorem for $L_2\big|_{V_{\lambda_1}}$. Each of these, by the commutativity, is carried into itself by $L_3, \dots, L_m$, and the restrictions of $L_3, \dots, L_m$ to $V_{\lambda_1,\lambda_2}$ form a commuting family of self-adjoint linear maps. Continuing in this way, we arrive at the decomposition asserted by the corollary for $L_1, \dots, L_m$. The assertion of the corollary about commuting Hermitian matrices is a special case, in the same way that the assertions in Theorem 3.21 about matrices were special cases of the assertions about linear maps. $\square$

A linear map $L : V \to V$, not necessarily self-adjoint, is said to be **normal** if $L$ commutes with its adjoint: $LL^* = L^*L$.

**Corollary 3.24.** Suppose that $\mathbb{F} = \mathbb{C}$, and let $L : V \to V$ be a normal linear map on the finite-dimensional inner-product space $V$. Then $V$ has an orthonormal basis of eigenvectors of $L$. In addition, for each complex scalar $\lambda$, let

$$V_\lambda = \{v \in V \mid L(v) = \lambda v\},$$

so that $V_\lambda$ when nonzero is the eigenspace of $L$ for the eigenvalue $\lambda$. Then the vector subspaces $V_\lambda$ are mutually orthogonal, and any orthonormal basis of $V$ of eigenvectors of $L$ is the union of orthonormal bases of the $V_\lambda$'s. Correspondingly if $A$ is any $n$-by-$n$ complex matrix such that $AA^* = A^*A$, then there exists a unitary matrix $C$ such that $C^{-1}AC$ is diagonal.

REMARK. The corollary fails if $\mathbb{F} = \mathbb{R}$: for the linear map $L : \mathbb{R}^2 \to \mathbb{R}^2$ with $L(v) = Av$ and $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $L^* = L^{-1}$ commutes with $L$, but $L$ has no eigenvectors in $\mathbb{R}^2$ since the characteristic polynomial $\lambda^2 + 1$ has no first-degree factors with real coefficients.

PROOF. The point is that $L = \left(\frac{1}{2}(L+L^*)\right) + i\left(\frac{1}{2i}(L-L^*)\right)$ and that $\frac{1}{2}(L+L^*)$ and $\frac{1}{2i}(L - L^*)$ are self-adjoint. If $L$ commutes with $L^*$, then $T_1 = \frac{1}{2}(L + L^*)$ and $T_2 = \frac{1}{2i}(L - L^*)$ commute with each other. We apply Corollary 3.23 to the commuting self-adjoint linear maps $T_1$ and $T_2$. The vector subspace $V_{\alpha,\beta}$ produced by Corollary 3.23 coincides with the vector subspace $V_{\alpha+i\beta}$ defined in the present corollary, and the result for $L$ follows. The result for matrices is a special case. □

**Corollary 3.25.** Suppose that $\mathbb{F} = \mathbb{C}$, and let $L : V \to V$ be a unitary linear map on the finite-dimensional inner-product space $V$. Then $V$ has an orthonormal basis of eigenvectors of $L$. In addition, for each complex scalar $\lambda$, let

$$V_\lambda = \{v \in V \mid L(v) = \lambda v\},$$

so that $V_\lambda$ when nonzero is the eigenspace of $L$ for the eigenvalue $\lambda$. Then the eigenvalues of $L$ all have absolute value 1, the vector subspaces $V_\lambda$ are mutually orthogonal, and any orthonormal basis of $V$ of eigenvectors of $L$ is the union of orthonormal bases of the $V_\lambda$'s. Correspondingly if $A$ is any $n$-by-$n$ unitary matrix, then there exists a unitary matrix $C$ such that $C^{-1}AC$ is diagonal; the diagonal entries of $C^{-1}AC$ all have absolute value 1.

PROOF. This is a special case of Corollary 3.24 since a unitary linear map $L$ has $LL^* = I = L^*L$. The eigenvalues all have absolute value 1 as a consequence of Proposition 3.18e. □

Now we come to the **polar decomposition** of linear maps and of matrices. When $\mathbb{F} = \mathbb{C}$, this is a generalization of the polar decomposition $z = e^{i\theta}r$ of complex numbers. When $\mathbb{F} = \mathbb{R}$, it generalizes the decomposition $x = (\operatorname{sgn} x)|x|$ of real numbers.

**Theorem 3.26** (polar decomposition). If $L : V \to V$ is a linear map on a finite-dimensional inner-product space, then $L$ decomposes as $L = UP$, where $P$ is positive semidefinite and $U$ is orthogonal if $\mathbb{F} = \mathbb{R}$ and unitary if $\mathbb{F} = \mathbb{C}$. The linear map $P$ is unique, and $U$ is unique if $L$ is invertible. Correspondingly any $n$-by-$n$ matrix $A$ decomposes as $A = UP$, where $P$ is a positive semidefinite matrix and $U$ is an orthogonal matrix if $\mathbb{F} = \mathbb{R}$ and a unitary matrix if $\mathbb{F} = \mathbb{C}$. The matrix $P$ is unique, and $U$ is unique if $A$ is invertible.

REMARKS. As we have already seen in other situations, the motivation for the proof comes from the uniqueness.

PROOF OF UNIQUENESS. Let $L = UP = U'P'$. Then $L^*L = P^2 = P'^2$. The linear map $L^*L$ is positive semidefinite since its adjoint is $(L^*L)^* = L^*L^{**} = L^*L$ and since $(L^*L(v), v) = (L(v), L(v)) \geq 0$. Therefore Corollary 3.22c shows that $L^*L$ has a unique positive semidefinite square root. Hence $P = P'$. If $L$ is invertible, then $P$ is invertible and $L = UP$ implies that $U = LP^{-1}$. The same argument applies in the case of matrices. □

PROOF OF EXISTENCE. If $L$ is given, then we have just seen that $L^*L$ is positive semidefinite. Let $P$ be its unique positive semidefinite square root. The proof is clearer when $L$ is invertible, and we consider that case first. Then we can set $U = LP^{-1}$. Since $U^* = (P^{-1})^*L^* = P^{-1}L^*$, we find that $U^*U = P^{-1}L^*LP^{-1} = P^{-1}P^2P^{-1} = I$, and we conclude that $U$ is unitary.

When $L$ is not necessarily invertible, we argue a little differently with the positive semidefinite square root $P$ of $L^*L$. The kernel $K$ of $P$ is the 0 eigenspace of $P$, and the Spectral Theorem (Theorem 3.21) shows that the image of $P$ is the sum of all the other eigenspaces and is just $K^\perp$. Since $K \cap K^\perp = 0$, $P$ is one-one from $K^\perp$ onto itself. Thus $P(v) \mapsto L(v)$ is a one-one linear map from $K^\perp$ into $V$. Call this function $U$, so that $U(P(v)) = L(v)$. For any $v_1$ and $v_2$ in $V$, we have

$$(L(v_1), L(v_2)) = (L^*L(v_1), v_2) = (P^2(v_1), v_2) = (P(v_1), P(v_2)), \qquad (*)$$

and hence $U : K^\perp \to V$ preserves inner products. Let $\{u_1, \dots, u_k\}$ be an orthonormal basis of $K^\perp$, and let $\{u_{k+1}, \dots, u_n\}$ be an orthonormal basis of $K$. Since $U$ preserves inner products and is linear, $\{U(u_1), \dots, U(u_k)\}$ is an orthonormal basis of $U(K^\perp)$. Extend $\{U(u_1), \dots, U(u_k)\}$ to an orthonormal basis of $V$ by adjoining vectors $v_{k+1}, \dots, v_n$, define $U(u_j) = v_j$ for $k + 1 \leq$

$j \le n$, and write $U$ also for the linear extension to all of $V$. Since $U$ carries one orthonormal basis $\{u_1, \dots, u_n\}$ of $V$ to another, $U$ is unitary. We have $UP = L$ on $K^\perp$, and equation $(*)$ with $v_1 = v_2$ shows that $\ker L = \ker P = K$. Therefore $UP = L$ everywhere. $\square$

## 4. Problems

1. Let $V = M_{nn}(\mathbb{C})$, and define an inner product on $V$ by $\langle A, B\rangle = \mathrm{Tr}(B^*A)$. The norm $\|\cdot\|_{\mathrm{HS}}$ obtained from this inner product is called the **Hilbert–Schmidt norm** of the matrix in question.
   (a) Prove that $\|A\|_{\mathrm{HS}}^2 = \sum_{i,j} |A_{ij}|^2$ for $A$ in $V$.
   (b) Let $E_{ij}$ be the matrix that is 1 in the $(i, j)^{\text{th}}$ entry and is 0 elsewhere. Prove that the set of all $E_{ij}$ is an orthonormal basis of $V$.
   (c) Interpret (a) in the light of (b).
   (d) Prove that the Hilbert–Schmidt norm is given on any matrix $A$ in $V$ by
   $$\|A\|_{\mathrm{HS}}^2 = \textstyle\sum_j \|Au_j\|^2 = \sum_{i,j} |v_i^* A u_j|^2,$$
   where $\{u_1, \dots, u_n\}$ and $\{v_1, \dots, v_n\}$ are any orthonormal bases of $\mathbb{C}^n$ and $v^*$ refers to the conjugate transpose of any member $v$ of $\mathbb{C}^n$.
   (e) Let $W$ be the vector subspace of all diagonal matrices in $V$. Describe explicitly the orthogonal complement $W^\perp$, and find its dimension.
2. Let $V_n$ be the inner-product space over $\mathbb{R}$ of all polynomials on $[0, 1]$ of degree $\le n$ with real coefficients. (The 0 polynomial is to be included.) The Riesz Representation Theorem says that there is a unique polynomial $p_n$ such that $f\left(\frac{1}{2}\right) = \int_0^1 f(x)p_n(x)\,dx$ for all $f$ in $V_n$. Set up a system of linear equations whose solution tells what $p_n$ is.
3. Let $V$ be a finite-dimensional inner-product space, and suppose that $L$ and $M$ are self-adjoint linear maps from $V$ to $V$. Show that $LM$ is self-adjoint if and only if $LM = ML$.
4. Let $V$ be a finite-dimensional inner-product space. If $L : V \to V$ is a linear map with adjoint $L^*$, prove that $\ker L = (\text{image } L^*)^\perp$.
5. Find all 2-by-2 Hermitian matrices $A$ with characteristic polynomial $\lambda^2 + 4\lambda + 6$.
6. Let $V_1$ and $V_2$ be finite-dimensional inner-product spaces over the same $\mathbb{F}$, the inner products being $(\cdot, \cdot)_1$ and $(\cdot, \cdot)_2$.
   (a) Using the case when $V_1 = V_2$ as a model, define the adjoint of a linear map $L : V_1 \to V_2$, proving its existence. The adjoint is to be a linear map $L^* : V_2 \to V_1$.

(b) If $\Gamma$ is an orthonormal basis of $V_1$ and $\Delta$ is an orthonormal basis of $V_2$, prove that the matrices of $L$ and $L^*$ in these bases are conjugate transposes of one another.

7. Suppose that a finite-dimensional inner-product space $V$ is a direct sum $V = S \oplus T$ of vector subspaces. Let $E : V \to V$ be the linear map that is the identity on $S$ and is 0 on $T$.
   (a) Prove that $V = S^\perp \oplus T^\perp$.
   (b) Prove that $E^* : V \to V$ is the linear map that is the identity on $T^\perp$ and is 0 on $S^\perp$.

8. **(Iwasawa decomposition)** Let $g$ be an invertible $n$-by-$n$ complex matrix. Apply the Gram–Schmidt orthogonalization process to the basis $\{ge_1, \dots, ge_n\}$, where $\{e_1, \dots, e_n\}$ is the standard basis, and let the resulting orthonormal basis be $\{v_1, \dots, v_n\}$. Define an invertible $n$-by-$n$ matrix $k$ such that $k^{-1}v_j = e_j$ for $1 \le j \le n$. Prove that $k^{-1}g$ is upper triangular with positive diagonal entries, and conclude that $g = k(k^{-1}g)$ exhibits $g$ as the product of a unitary matrix and an upper triangular matrix whose diagonal entries are positive.

9. Let $A$ be an $n$-by-$n$ positive definite matrix.
   (a) Prove that $\det A > 0$.
   (b) Prove for any subset of integers $1 \le i_1 < i_2 < \cdots < i_k \le n$ that the submatrix of $A$ built from rows and columns indexed by $(i_1, \dots, i_k)$ is positive definite.

10. Prove that if $A$ is a positive definite $n$-by-$n$ matrix, then there exists an $n$-by-$n$ upper-triangular matrix $B$ with positive diagonal entries such that $A = B^*B$.

11. The most general 2-by-2 Hermitian matrix is of the form $A = \begin{pmatrix} a & b \\ \bar{b} & d \end{pmatrix}$ with $a$ and $d$ real and with $b$ complex. Find a diagonal matrix $D$ and a unitary matrix $U$ such that $D = U^{-1}AU$.

12. In the previous problem,
   (a) what conditions on $A$ make $A$ positive definite?
   (b) when $A$ is positive definite, how can its positive definite square root be computed explicitly?

13. Prove that if an $n$-by-$n$ real symmetric matrix $A$ has $v^tAv = 0$ for all $v$ in $\mathbb{R}^n$, then $A = 0$.

14. Let $L : \mathbb{C}^n \to \mathbb{C}^n$ be a self-adjoint linear map. Show for each $x \in \mathbb{C}^n$ that there is some $y \in \mathbb{C}^n$ such that $(I - L)^2(y) = (I - L)(x)$.

15. In the polar decomposition $L = UP$, prove that if $P$ and $U$ commute, then $L$ is normal.

16. Let $V$ be an $n$-dimensional inner-product space over $\mathbb{R}$. What is the largest possible dimension of a commuting family of self-adjoint linear maps $L : V \to V$?

17. Let $v_1, \dots, v_n$ be an ordered list of vectors in an inner-product space. The associated **Gram matrix** is the Hermitian matrix of inner products given by $G(v_1, \dots, v_n) = [(v_i, v_j)]$, and $\det G(v_1, \dots, v_n)$ is called its **Gram determinant**.
    (a) If $c_1, \dots, c_n$ are in $\mathbb{C}$, let $c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$. Prove that $c^t G(v_1, \dots, v_n)\bar{c} = \|c_1 v_1 + \cdots + c_n v_n\|^2$, and conclude that $G(v_1, \dots, v_n)$ is positive semidefinite.
    (b) Prove that $\det G(v_1, \dots, v_n) \geq 0$ with equality if and only if $v_1, \dots, v_n$ are linearly dependent. (This generalizes the Schwarz inequality.)
    (c) Under what circumstances does equality hold in the Schwarz inequality?

Problems 18–23 introduce the **Legendre polynomials** and establish some of their elementary properties, including their orthogonality under the inner product $\langle P, Q\rangle = \int_{-1}^{1} P(x)Q(x)\,dx$. They form the simplest family of classical orthogonal polynomials. They are uniquely determined by the conditions that the $n^{\text{th}}$ one $P_n$, for $n \geq 0$, is of degree $n$, they are orthogonal under $\langle \cdot, \cdot \rangle$, and they are normalized so that $P_n(1) = 1$. But these conditions are a little hard to work with initially, and instead we adopt the recursive definition $P_0(x) = 1$, $P_1(x) = x$, and

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x) \qquad \text{for } n \geq 1.$$

18. (a) Prove that $P_n(x)$ has degree $n$, that $P_n(-x) = (-1)^n P_n(x)$, and that $P_n(1) = 1$. In particular, $P_n$ is an even function if $n$ is even and is an odd function if $n$ is odd.
    (b) Let $c^{(n)}$ be the constant term of $P_n$ if $n$ is even and the coefficient of $x$ if $n$ is odd, so that $c^{(0)} = c^{(1)} = 1$. Prove that $c^{(n)} = -\frac{n-1}{n}\, c^{(n-2)}$ for $n \geq 2$.
19. This part establishes a useful concrete formula for $P_n(x)$. Let $D = d/dx$ and $X = x^2 - 1$, writing $X' = 2x$, $X'' = 2$, and $X''' = 0$ for the derivatives. Two parts of this problem make use of the Leibniz rule $D^n(fg) = \sum_{k=0}^{n} \binom{n}{k}(D^{n-k}f)(D^k g)$ for higher-order derivatives of a product.
    (a) Verify that $D^2(X^{n+1}) = (2n+1)D(X^n X') - n(2n+1)X''X^n - 4n^2 X^{n-1}$.
    (b) By applying $D^{n-1}$ to the result of (a) and rearranging terms, show that $D^{n+1}(X^{n+1}) = (2n+1)X'D^n(X^n) - 4n^2 D^{n-1}(X^{n-1})$.
    (c) Put $R_n(x) = (2^n n!)^{-1} D^n(X^n)$ for $n \geq 0$. Show that $R_0(x) = 1$, $R_1(x) = x$, and $(n+1)R_{n+1}(x) = (2n+1)xR_n(x) - nR_{n-1}(x)$ for $n \geq 1$.
    (d) **(Rodrigues's formula)** Conclude that $2^n n! P_n(x) = \left(\frac{d}{dx}\right)^n [(x^2-1)^n]$.
20. Using Rodrigues's formula and iterated integration by parts, prove that

$$\int_{-1}^{1} P_m(x)P_n(x)\,dx = 0 \qquad \text{for } m < n.$$

Conclude that $\{P_0, P_1, \dots, P_n\}$ is an orthogonal basis of the inner-product space of polynomials on $[-1, 1]$ of degree $\leq n$ with inner product $\langle \cdot, \cdot \rangle$.

21. Arguing as in the previous problem and taking for granted that $\int_{-1}^{1}(1-x^2)^n\,dx = \frac{2(2^n n!)^2}{(2n+1)!}$, prove that $\langle P_n, P_n\rangle = \left(n+\frac{1}{2}\right)^{-1}$.

22. This problem shows that $P_n(x)$ satisfies a certain second-order differential equation. Let $D = d/dx$. The first two parts of this problem use the Leibniz rule quoted in Problem 19. Let $X = x^2 - 1$ and $K_n = 2^n n!$, so that Rodrigues's formula says that $K_n P_n = D^n(X^n)$.
    (a) Expand $D^{n+1}[(D(X^n))X]$ by the Leibniz rule.
    (b) Observe that $(D(X^n))X = nX^nX'$, and expand $D^{n+1}[(nX^n)X']$ by the Leibniz rule.
    (c) Equating the results of the previous two parts, conclude that $y = P_n(x)$ satisfies the differential equation $(1-x^2)y'' - 2xy' + n(n+1)y = 0$.

23. Let $P_n(x) = \sum_{k=0}^{n} c_k x^k$. Using the differential equation, show that the coefficients $c_k$ satisfy $k(k-1)c_k = [(k-2)(k-1) - n(n+1)]c_{k-2}$ for $k \geq 2$ and that $c_k = 0$ unless $n-k$ is even.

Problems 24–28 concern the complex conjugate of an inner-product space over $\mathbb{C}$. For any finite-dimensional inner-product space $V$, the Riesz Representation Theorem identifies the dual $V'$ with $V$, saying that each member of $V'$ is given by taking the inner product with some member of $V$. When the scalars are real, this identification is linear; thus the Riesz theorem uses the inner product to construct a canonical isomorphism of $V$ onto $V'$. When the scalars are complex, the identification is conjugate linear, and we do not get an isomorphism of $V$ with $V'$. The **complex conjugate** of $V$ provides a substitute result.

24. Let $V$ be a finite-dimensional vector space over $\mathbb{C}$. Define a new complex vector space $\overline{V}$ as follows: The elements of $\overline{V}$ are the elements of $V$, and the definition of addition is unchanged. However, there is a change in the definition of scalar multiplication, in that if $v$ is in $V$, then the product $cv$ in $\overline{V}$ is to equal the product $\bar{c}v$ in $V$. Verify that $\overline{V}$ is indeed a complex vector space.

25. If $V$ is a complex vector space and $L : V \to V$ is a linear map, define $\overline{L} : \overline{V} \to \overline{V}$ to be the same function as $L$. Prove that $\overline{L}$ is linear.

26. Suppose that the complex vector space $V$ is actually a finite-dimensional inner-product space, with inner product $(\cdot\,,\,\cdot)_V$. Define $(u, v)_{\overline{V}} = (v, u)_V$. Verify that $\overline{V}$ is an inner-product space.

27. With $V$ as in the previous problem, show that the Riesz Representation Theorem uses the inner product to set up a canonical isomorphism of $V'$ with $\overline{V}$.

28. With $V$ and $\overline{V}$ as in the two previous problems, let $L : V \to V$ be linear, so that $(\overline{L})^* : \overline{V} \to \overline{V}$ is linear. Under the identification of the previous problem of $\overline{V}$ with $V'$, show that $(\overline{L})^*$ corresponds to the contragredient $L^t$ as defined in Section II.4.

Problems 29–32 use inner-product spaces to obtain a decomposition of polynomials in several variables. A real-valued polynomial function $p$ in $x_1, \dots, x_n$ is said to be **homogeneous** of degree $N$ if every monomial in $p$ has total degree $N$. Let $V_N$ be the space of real-valued polynomials in $x_1, \dots, x_n$ homogeneous of degree $N$. For any homogeneous polynomial $p$, we define a differential operator $\partial(p)$ with constant coefficients by requiring that $\partial(\,\cdot\,)$ be linear in $(\,\cdot\,)$ and that

$$\partial(x_1^{k_1} \cdots x_n^{k_n}) = \frac{\partial^{k_1+\cdots+k_n}}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}}.$$

For example, if $|x|^2$ stands for $x_1^2 + \cdots + x_n^2$, then $\partial(|x|^2) = \Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}$. If $p$ and $q$ are in the same $V_N$, then $\partial(q)p$ is a constant polynomial, and we define $\langle p, q\rangle$ to be that constant. Then $\langle\,\cdot\,,\,\cdot\,\rangle$ is bilinear.

29. (a) Prove that $\langle\,\cdot\,,\,\cdot\,\rangle$ satisfies $\langle p, q\rangle = \langle q, p\rangle$.
    (b) Prove that $\langle x_1^{k_1} \cdots x_n^{k_n}, x_1^{l_1} \cdots x_n^{l_n}\rangle$ is positive if $(k_1, \dots, k_n) = (l_1, \dots, l_n)$ and is 0 otherwise.
    (c) Deduce that $\langle\,\cdot\,,\,\cdot\,\rangle$ is an inner product on $V_N$.
30. Call $p \in V_N$ **harmonic** if $\partial(|x|^2)p = 0$, and let $H_N$ be the vector subspace of harmonic polynomials. Prove that the orthogonal complement of $|x|^2 V_{N-2}$ in $V_N$ relative to $\langle\,\cdot\,,\,\cdot\,\rangle$ is $H_N$.
31. Deduce from Problem 30 that each $p \in V_N$ decomposes uniquely as

$$p = h_N + |x|^2 h_{N-2} + |x|^4 h_{N-4} + \cdots$$

    with $h_N, h_{N-2}, h_{N-4}, \dots$ homogeneous harmonic of the indicated degrees.
32. For $n = 2$, describe a computational procedure for decomposing the element $x_1^4 + x_2^4$ of $V_4$ as in Problem 31.

Problems 33–34 concern products of $n$-by-$n$ positive semidefinite matrices. They make use of Problem 26 in Chapter II, which says that $\det(\lambda I - CD) = \det(\lambda I - DC)$.

33. Let $A$ and $B$ be positive semidefinite. Using the positive definite square root of $B$, prove that every eigenvalue of $AB$ is $\geq 0$.
34. Let $A$, $B$, and $C$ be positive semidefinite, and suppose that $ABC$ is Hermitian. Under the assumption that $C$ is invertible, introduce the positive definite square root $P$ of $C$. By considering $P^{-1}ABCP^{-1}$, prove that $ABC$ is positive semidefinite.