product space, the scalars $c_1, \ldots, c_n$ are (of course) real, and so it must be that $T = T^*$ .In other words, if V is a finite-dimensional *real* inner probut space and T is a linear operator for which there is an orthonormal basis of characteristic vectos, then T must be self-adjoint. If V is a complex inner product space, the scalars $c_1, \ldots, c_n$ need not be real, i.e., T need not be self-adjoint. But notice that T must satisfy

$$TT^* = T^*T. \tag{8-17}$$

For, any two diagonal matrices commute, and since T and T^* are both represented by diagonal matrices in the ordered basis $\mathcal{B}$, we have (8-17). It is a rather remarkable fact that in the complex case this condition is also sufficient to imply the existence of an orthonormal basis of characteristic vectors.

Definition. *Let* V *be a finite-dimensional inner product space and* T *a linear operator on* V. *We say that* T *is* **normal** *if it commutes with its adjoint i.e.,* TT* = T*T.

Any self-adjoint operator is normal, as is any unitary operator. Any scalar multiple of a normal operator is normal; however, sums and products of normal operators are not generally normal. Although it is by no means necessary, we shall begin our study of normal operators by considering self-adjoint operators.

Theorem 15. *Let* V *be an inner product space and* T *a self-adjoint linear operator on* V. *Then each characteristic value of* T *is real, and characteristic vectors of* T *associated with distinct characteristic values are orthogonal.*

Proof. Suppose c is a characteristic value of T, i.e., that $T\alpha = c\alpha$ for some non-zero vector α. Then

$$\begin{aligned} c(\alpha|\alpha) &= (c\alpha|\alpha) \\ &= (T\alpha|\alpha) \\ &= (\alpha|T\alpha) \\ &= (\alpha|c\alpha) \\ &= \bar{c}(\alpha|\alpha). \end{aligned}$$

Since $(\alpha|\alpha) \neq 0$, we must have $c = \bar{c}$. Suppose we also have $T\beta = d\beta$ with $\beta \neq 0$. Then

$$\begin{aligned} c(\alpha|\beta) &= (T\alpha|\beta) \\ &= (\alpha|T\beta) \\ &= (\alpha|d\beta) \\ &= \bar{d}(\alpha|\beta) \\ &= d(\alpha|\beta). \end{aligned}$$

If $c \neq d$, then $(\alpha|\beta) = 0$. ∎

It should be pointed out that Theorem 15 says nothing about the existence of characteristic values or characteristic vectors.

Theorem 16. *On a finite-dimensional inner product space of positive dimension, every self-adjoint operator has a (non-zero) characteristic vector.*

Proof. Let V be an inner product space of dimension n, where $n > 0$, and let T be a self-adjoint operator on V. Choose an orthonormal basis $\mathcal{B}$ for V and let $A = [T]_{\mathcal{B}}$. Since $T = T^*$, we have $A = A^*$. Now let W be the space of $n \times 1$ matrices over C, with inner product $(X|Y) = Y^*X$. Then $U(X) = AX$ defines a self-adjoint linear operator U on W. The characteristic polynomial, $\det(xI - A)$, is a polynomial of degree n over the complex numbers; every polynomial over C of positive degree has a root. Thus, there is a complex number c such that $\det(cI - A) = 0$. This means that $A - cI$ is singular, or that there exists a non-zero X such that $AX = cX$. Since the operator U (multiplication by A) is self-adjoint, it follows from Theorem 15 that c is real. If V is a real vector space, we may choose X to have real entries. For then A and $A - cI$ have real entries, and since $A - cI$ is singular, the system $(A - cI)X = 0$ has a non-zero real solution X. It follows that there is a non-zero vector α in V such that $T\alpha = c\alpha$. ▌

There are several comments we should make about the proof.

(1) The proof of the existence of a non-zero X such that $AX = cX$ had nothing to do with the fact that A was Hermitian (self-adjoint). It shows that any linear operator on a finite-dimensional complex vector space has a characteristic vector. In the case of a real inner product space, the self-adjointness of A is used very heavily, to tell us that each characteristic value of A is real and hence that we can find a suitable X with real entries.

(2) The argument shows that the characteristic polynomial of a self-adjoint matrix has real coefficients, in spite of the fact that the matrix may not have real entries.

(3) The assumption that V is finite-dimensional is necessary for the theorem; a self-adjoint operator on an infinite-dimensional inner product space need not have a characteristic value.

Example 29. Let V be the vector space of continuous complex-valued (or real-valued) continuous functions on the unit interval, $0 \leq t \leq 1$, with the inner product

$$(f|g) = \int_0^1 f(t)\overline{g(t)}\, dt.$$

The operator 'multiplication by t,' $(Tf)(t) = tf(t)$, is self-adjoint. Let us suppose that $Tf = cf$. Then

$$(t - c)f(t) = 0, \qquad 0 \leq t \leq 1$$

and so $f(t) = 0$ for $t \neq c$. Since f is continuous, $f = 0$. Hence T has no characteristic values (vectors).

Theorem 17. *Let V be a finite-dimensional inner product space, and let T be any linear operator on V. Suppose W is a subspace of V which is invariant under T. Then the orthogonal complement of W is invariant under* T*.

Proof. We recall that the fact that W is invariant under T does not mean that each vector in W is left fixed by T; it means that if α is in W then $T\alpha$ is in W. Let β be in $W^\perp$. We must show that $T^*\beta$ is in $W^\perp$, that is, that $(\alpha|T^*\beta) = 0$ for every α in W. If α is in W, then $T\alpha$ is in W, so $(T\alpha|\beta) = 0$. But $(T\alpha|\beta) = (\alpha|T^*\beta)$. ∎

Theorem 18. *Let V be a finite-dimensional inner product space, and let T be a self-adjoint linear operator on V. Then there is an orthonormal basis for V, each vector of which is a characteristic vector for T.*

Proof. We are assuming $\dim V > 0$. By Theorem 16, T has a characteristic vector α. Let $\alpha_1 = \alpha/||\alpha||$ so that α_1 is also a characteristic vector for T and $||\alpha_1|| = 1$. If $\dim V = 1$, we are done. Now we proceed by induction on the dimension of V. Suppose the theorem is true for inner product spaces of dimension less than $\dim V$. Let W be the one-dimensional subspace spanned by the vector α_1. The statement that α_1 is a characteristic vector for T simply means that W is invariant under T. By Theorem 17, the orthogonal complement $W^\perp$ is invariant under $T^* = T$. Now $W^\perp$, with the inner product from V, is an inner product space of dimension one less than the dimension of V. Let U be the linear operator induced on $W^\perp$ by T, that is, the restriction of T to $W^\perp$. Then U is self-adjoint, and by the induction hypothesis, $W^\perp$ has an orthonormal basis $\{\alpha_2, \ldots, \alpha_n\}$ consisting of characteristic vectors for U. Now each of these vectors is also a characteristic vector for T, and since $V = W \oplus W^\perp$, we conclude that $\{\alpha_1, \ldots, \alpha_n\}$ is the desired basis for V. ∎

Corollary. *Let A be an n × n Hermitian (self-adjoint) matrix. Then there is a unitary matrix P such that* $P^{-1}AP$ *is diagonal (A is unitarily equivalent to a diagonal matrix). If A is a real symmetric matrix, there is a real orthogonal matrix P such that* $P^{-1}AP$ *is diagonal.*

Proof. Let V be $C^{n\times 1}$, with the standard inner product, and let T be the linear operator on V which is represented by A in the standard ordered basis. Since $A = A^*$, we have $T = T^*$. Let $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ be an ordered orthonormal basis for V, such that $T\alpha_j = c_j\alpha_j$, $j = 1, \ldots, n$. If $D = [T]_{\mathcal{B}}$, then D is the diagonal matrix with diagonal entries $c_1, \ldots, c_n$. Let P be the matrix with column vectors $\alpha_1, \ldots, \alpha_n$. Then $D = P^{-1}AP$.

In case each entry of A is real, we can take V to be R^n, with the standard inner product, and repeat the argument. In this case, P will be a unitary matrix with real entries, i.e., a real orthogonal matrix. ▌

Combining Theorem 18 with our comments at the beginning of this section, we have the following: If V is a finite-dimensional *real* inner product space and T is a linear operator on V, then V has an orthonormal basis of characteristic vectors for T if and only if T is self-adjoint. Equivalently, if A is an $n \times n$ matrix with *real* entries, there is a real orthogonal matrix P such that P^tAP is diagonal if and only if $A = A^t$. There is no such result for complex symmetric matrices. In other words, for complex matrices there is a significant difference between the conditions $A = A^t$ and $A = A^*$.

Having disposed of the self-adjoint case, we now return to the study of normal operators in general. We shall prove the analogue of Theorem 18 for normal operators, in the *complex* case. There is a reason for this restriction. A normal operator on a real inner product space may not have any non-zero characteristic vectors. This is true, for example, of all but two rotations in R^2.

Theorem 19. *Let* V *be a finite-dimensional inner product space and* T *a normal operator on* V*. Suppose* α *is a vector in* V*. Then* α *is a characteristic vector for* T *with characteristic value* c *if and only if* α *is a characteristic vector for* T* *with characteristic value* $\bar{c}$*.*

Proof. Suppose U is any normal operator on V. Then $||U\alpha|| = ||U^*\alpha||$. For using the condition $UU^* = U^*U$ one sees that

$$\begin{aligned} ||U\alpha||^2 = (U\alpha|U\alpha) &= (\alpha|U^*U\alpha) \\ &= (\alpha|UU^*\alpha) = (U^*\alpha|U^*\alpha) = ||U^*\alpha||^2. \end{aligned}$$

If c is any scalar, the operator $U = T - cI$ is normal. For $(T - cI)^* = T^* - \bar{c}I$, and it is easy to check that $UU^* = U^*U$. Thus

$$||(T - cI)\alpha|| = ||(T^* - cI)\alpha||$$

so that $(T - cI)\alpha = 0$ if and only if $(T^* - \bar{c}I)\alpha = 0$. ▌

Definition. *A complex* n × n *matrix* A *is called* **normal** *if* AA* = A*A.

It is not so easy to understand what normality of matrices or operators really means; however, in trying to develop some feeling for the concept, the reader might find it helpful to know that a triangular matrix is normal if and only if it is diagonal.

Theorem 20. *Let* V *be a finite-dimensional inner product space,* T *a linear operator on* V*, and* $\mathcal{B}$ *an orthonormal basis for* V*. Suppose that the* *matrix* A *of* T *in the basis* $\mathcal{B}$ *is upper triangular. Then* T *is normal if and only if* A *is a diagonal matrix.*

Proof. Since $\mathcal{B}$ is an orthonormal basis, A^* is the matrix of T^* in $\mathcal{B}$. If A is diagonal, then $AA^* = A^*A$, and this implies $TT^* = T^*T$. Conversely, suppose T is normal, and let $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$. Then, since A is upper-triangular, $T\alpha_1 = A_{11}\alpha_1$. By Theorem 19 this implies, $T^*\alpha_1 = \bar{A}_{11}\alpha_1$. On the other hand,

$$\begin{aligned} T^*\alpha_1 &= \sum_j (A^*)_{j1}\alpha_j \\ &= \sum_j \bar{A}_{1j}\alpha_j. \end{aligned}$$

Therefore, $A_{1j} = 0$ for every $j > 1$. In particular, $A_{12} = 0$, and since A is upper-triangular, it follows that

$$T\alpha_2 = A_{22}\alpha_2.$$

Thus $T^*\alpha_2 = \bar{A}_{22}\alpha_2$ and $A_{2j} = 0$ for all $j \neq 2$. Continuing in this fashion, we find that A is diagonal. ∎

Theorem 21. *Let* V *be a finite-dimensional complex inner product space and let* T *be any linear operator on* V. *Then there is an orthonormal basis for* V *in which the matrix of* T *is upper triangular.*

Proof. Let n be the dimension of V. The theorem is true when $n = 1$, and we proceed by induction on n, assuming the result is true for linear operators on complex inner product spaces of dimension $n - 1$. Since V is a finite-dimensional complex inner product space, there is a unit vector α in V and a scalar c such that

$$T^*\alpha = c\alpha.$$

Let W be the orthogonal complement of the subspace spanned by α and let S be the restriction of T to W. By Theorem 17, W is invariant under T. Thus S is a linear operator on W. Since W has dimension $n - 1$, our inductive assumption implies the existence of an orthonormal basis $\{\alpha_1, \ldots, \alpha_{n-1}\}$ for W in which the matrix of S is upper-triangular; let $\alpha_n = \alpha$. Then $\{\alpha_1, \ldots, \alpha_n\}$ is an orthonormal basis for V in which the matrix of T is upper-triangular. ∎

This theorem implies the following result for matrices.

Corollary. *For every complex* $n \times n$ *matrix* A *there is a unitary matrix* U *such that* $U^{-1}AU$ *is upper-triangular.*

Now combining Theorem 21 and Theorem 20, we immediately obtain the following analogue of Theorem 18 for normal operators.