If B is an element of V, let L_B, R_B, and T_B denote the linear operators on V defined by

(a) $L_B(A) = BA$.
(b) $R_B(A) = AB$.
(c) $T_B(A) = BA - AB$.

Consider the three families of operators obtained by letting B vary over all diagonal matrices. Show that each of these families is a commutative self-adjoint algebra and find their spectral resolutions.

8. If B is an arbitrary member of the inner product space in Exercise 7, show that L_B is unitarily equivalent to R_{B^t}.

9. Let V be the inner product space in Exercise 7 and G the group of unitary matrices in V. If B is in G, let C_B denote the linear operator on V defined by

$$C_B(A) = BAB^{-1}.$$

Show that

(a) C_B is a unitary operator on V;
(b) $C_{B_1B_2} = C_{B_1}C_{B_2}$;
(c) there is no unitary transformation U on V such that

$$UL_BU^{-1} = C_B$$

for all B in G.

10. Let $\mathfrak{F}$ be any family of linear operators on a finite-dimensional inner product space V and $\mathfrak{A}$ the self-adjoint algebra generated by $\mathfrak{F}$. Show that

(a) each root of $\mathfrak{A}$ defines a root of $\mathfrak{F}$;
(b) each root r of $\mathfrak{A}$ is a multiplicative linear function on A, i.e.,

$$r(TU) = r(T)r(U)$$
$$r(cT + U) = cr(T) + r(U)$$

for all T and U in $\mathfrak{A}$ and all scalars c.

11. Let $\mathfrak{F}$ be a commuting family of diagonalizable normal operators on a finite-dimensional inner product space V; and let $\mathfrak{A}$ be the self-adjoint algebra generated by $\mathfrak{F}$ and the identity operator I. Show that each root of $\mathfrak{A}$ is different from 0, and that for each root r of $\mathfrak{F}$ there is a unique root s of $\mathfrak{A}$ such that $s(T) = r(T)$ for all T in $\mathfrak{F}$.

12. Let $\mathfrak{F}$ be a commuting family of diagonalizable normal operators on a finite-dimensional inner product space V and A_0 the self-adjoint algebra generated by $\mathfrak{F}$. Let $\mathfrak{A}$ be the self-adjoint algebra generated by $\mathfrak{F}$ and the identity operator I. Show that

(a) $\mathfrak{A}$ is the set of all operators on V of the form $cI + T$ where c is a scalar and T an operator in $\mathfrak{A}_0$
(b) There is at most one root r of $\mathfrak{A}$ such that $r(T) = 0$ for all T in $\mathfrak{A}_0$.
(c) If one of the roots of $\mathfrak{A}$ is 0 on $\mathfrak{A}_0$, the projections $P_1, \ldots, P_k$ in the resolution of the identity defined by $\mathfrak{F}$ may be indexed in such a way that $\mathfrak{A}_0$ consists of all operators on V of the form

$$T = \sum_{j=2}^{k} c_j P_j$$

where $c_2, \ldots, c_k$ are arbitrary scalars.

(d) $\mathfrak{A} = \mathfrak{A}_0$ if and only if for each root r of $\mathfrak{A}$ there exists an operator T in $\mathfrak{A}_0$ such that $r(T) \neq 0$.

9.6. Further Properties of Normal Operators

In Section 8.5 we developed the basic properties of self-adjoint and normal operators, using the simplest and most direct methods possible. In Section 9.5 we considered various aspects of spectral theory. Here we prove some results of a more technical nature which are mainly about normal operators on real spaces.

We shall begin by proving a sharper version of the primary decomposition theorem of Chapter 6 for normal operators. It applies to both the real and complex cases.

Theorem 17. *Let* T *be a normal operator on a finite-dimensional inner product space* V. *Let* p *be the minimal polynomial for* T *and* $p_1, \cdots, p_k$ *its distinct monic prime factors. Then each* p_j *occurs with multiplicity* 1 *in the factorization of* p *and has degree* 1 *or* 2. *Suppose* W_j *is the null space of* $p_j(T)$. *Then*

(i) W_j *is orthogonal to* W_i *when* $i \neq j$;

(ii) $V = W_1 \oplus \cdots \oplus W_k$;

(iii) W_j *is invariant under* T, *and* p_j *is the minimal polynomial for the restriction of* T *to* W_j;

(iv) *for every* j, *there is a polynomial* e_j *with coefficients in the scalar field such that* $e_j(T)$ *is the orthogonal projection of* V *on* W_j.

In the proof we use certain basic facts which we state as lemmas.

Lemma 1. *Let* N *be a normal operator on an inner product space* W. *Then the null space of* N *is the orthogonal complement of its range.*

Proof. Suppose $(\alpha|N\beta) = 0$ for all β in W. Then $(N^*\alpha|\beta) = 0$ for all β; hence $N^*\alpha = 0$. By Theorem 19 of Chapter 8, this implies $N\alpha = 0$. Conversely, if $N\alpha = 0$, then $N^*\alpha = 0$, and

$$(N^*\alpha|\beta) = (\alpha|N\beta) = 0$$

for all β in W. ∎

Lemma 2. *If* N *is a normal operator and* α *is a vector such that* $N^2\alpha = 0$, *then* $N\alpha = 0$.

Proof. Suppose N is normal and that $N^2\alpha = 0$. Then $N\alpha$ lies in the range of N and also lies in the null space of N. By Lemma 1, this implies $N\alpha = 0$. ∎

Lemma 3. *Let* T *be a normal operator and* f *any polynomial with coefficients in the scalar field. Then* f(T) *is also normal.*

Proof. Suppose $f = a_0 + a_1x + \cdots + a_nx^n$. Then

$$f(T) = a_0I + a_1T + \cdots + a_nT^n$$

and

$$f(T)^* = \bar{a}_0I + \bar{a}_1T^* + \cdots + \bar{a}_n(T^*)^n.$$

Since $T^*T = TT^*$, it follows that $f(T)$ commutes with $f(T)^*$. ∎

Lemma 4. *Let* T *be a normal operator and* f, g *relatively prime polynomials with coefficients in the scalar field. Suppose α and β are vectors such that* $f(T)\alpha = 0$ *and* $g(T)\beta = 0$. *Then* $(\alpha|\beta) = 0$.

Proof. There are polynomials a and b with coefficients in the scalar field such that $af + bg = 1$. Thus

$$a(T)f(T) + b(T)g(T) = I$$

and $\alpha = g(T)b(T)\alpha$. It follows that

$$(\alpha|\beta) = (g(T)b(T)\alpha|\beta) = (b(T)\alpha|g(T)^*\beta).$$

By assumption $g(T)\beta = 0$. By Lemma 3, $g(T)$ is normal. Therefore, by Theorem 19 of Chapter 8, $g(T)^*\beta = 0$; hence $(\alpha|\beta) = 0$. ∎

Proof of Theorem 17. Recall that the minimal polynomial for T is the monic polynomial of least degree among all polynomials f such that $f(T) = 0$. The existence of such polynomials follows from the assumption that V is finite-dimensional. Suppose some prime factor p_j of p is repeated. Then $p = p_j^2g$ for some polynomial g. Since $p(T) = 0$, it follows that

$$(p_j(T))^2g(T)\alpha = 0$$

for every α in V. By Lemma 3, $p_j(T)$ is normal. Thus Lemma 2 implies

$$p_j(T)g(T)\alpha = 0$$

for every α in V. But this contradicts the assumption that p has least degree among all f such that $f(T) = 0$. Therefore, $p = p_1 \cdots p_k$. If V is a complex inner product space each p_j is necessarily of the form

$$p_j = x - c_j$$

with c_j real or complex. On the other hand, if V is a real inner product space, then $p_j = x_j - c_j$ with c_j in R or

$$p_j = (x - c)(x - \bar{c})$$

where c is a non-real complex number.

Now let $f_j = p/p_j$. Then, since $f_1, \ldots, f_k$ are relatively prime, there exist polynomials g_j with coefficients in the scalar field such that

$$1 = \sum_j f_j g_j. \tag{9-16}$$

We briefly indicate how such g_j may be constructed. If $p_j = x - c_j$, then $f_j(c_j) \neq 0$, and for g_j we take the scalar polynomial $1/f_j(c_j)$. When every p_j is of this form, the $f_j g_j$ are the familiar Lagrange polynomials associated with $c_1, \ldots, c_k$, and (9-16) is clearly valid. Suppose some $p_j = (x - c)(x - \bar{c})$ with c a non-real complex number. Then V is a real inner product space, and we take

$$g_j = \frac{x - \bar{c}}{s} + \frac{x - c}{\bar{s}}$$

where $s = (c - \bar{c})f_j(c)$. Then

$$g_j = \frac{(s + \bar{s})x - (cs + \bar{c}s)}{s\bar{s}}$$

so that g_j is a polynomial with real coefficients. If p has degree n, then

$$1 - \sum_j f_j g_j$$

is a polynomial with real coefficients of degree at most $n - 1$; moreover, it vanishes at each of the n (complex) roots of p, and hence is identically 0.

Now let α be an arbitrary vector in V. Then by (9-16)

$$\alpha = \sum_j f_j(T) g_j(T)\alpha$$

and since $p_j(T)f_j(T) = 0$, it follows that $f_j(T)g_j(T)\alpha$ is in W_j for every j. By Lemma 4, W_j is orthogonal to W_i whenever $i \neq j$. Therefore, V is the orthogonal direct sum of $W_1, \ldots, W_k$. If β is any vector in W_j, then

$$p_j(T)T\beta = Tp_j(T)\beta = 0;$$

thus W_j is invariant under T. Let T_j be the restriction of T to W_j. Then $p_j(T_j) = 0$, so that p_j is divisible by the minimal polynomial for T_j. Since p_j is irreducible over the scalar field, it follows that p_j is the minimal polynomial for T_j.

Next, let $e_j = f_j g_j$ and $E_j = e_j(T)$. Then for every vector α in V, $E_j\alpha$ is in W_j, and

$$\alpha = \sum_j E_j\alpha.$$

Thus $\alpha - E_i\alpha = \sum_{j \neq i} E_j\alpha$; since W_j is orthogonal to W_i when $j \neq i$, this implies that $\alpha - E_i\alpha$ is in $W_i^{\perp}$. It now follows from Theorem 4 of Chapter 8 that E_i is the orthogonal projection of V on W_i. ∎

Definition. *We call the subspaces* W_j $(1 \leq j \leq k)$ *the* **primary components of V under T.**

Corollary. *Let T be a normal operator on a finite-dimensional inner product space V and $W_1, \ldots, W_k$ the primary components of V under T. Suppose W is a subspace of V which is invariant under T. Then*

$$W = \sum_j W \cap W_j.$$

Proof. Clearly W contains $\sum_j W \cap W_j$. On the other hand, W, being invariant under T, is invariant under every polynomial in T. In particular, W is invariant under the orthogonal projection E_j of V on W_j. If α is in W, it follows that $E_j\alpha$ is in $W \cap W_j$, and, at the same time, $\alpha = \sum_j E_j\alpha$. Therefore W is contained in $\sum_j W \cap W_j$. ∎

Theorem 17 shows that every normal operator T on a finite-dimensional inner product space is canonically specified by a finite number of normal operators T_j, defined on the primary components W_j of V under T, each of whose minimal polynomials is irreducible over the field of scalars. To complete our understanding of normal operators it is necessary to study normal operators of this special type.

A normal operator whose minimal polynomial is of degree 1 is clearly just a scalar multiple of the identity. On the other hand, when the minimal polynomial is irreducible and of degree 2 the situation is more complicated.

Example 1. Suppose $r > 0$ and that θ is a real number which is not an integral multiple of π. Let T be the linear operator on R^2 whose matrix in the standard orthonormal basis is

$$A = r\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Then T is a scalar multiple of an orthogonal transformation and hence normal. Let p be the characteristic polynomial of T. Then

$$\begin{aligned} p &= \det(xI - A) \\ &= (x - r\cos\theta)^2 + r^2\sin^2\theta \\ &= x - 2r\cos\theta x + r^2. \end{aligned}$$

Let $a = r\cos\theta$, $b = r\sin\theta$, and $c = a + ib$. Then $b \neq 0$, $c = re^{i\theta}$

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

and $p = (x - c)(x - \bar{c})$. Hence p is irreducible over R. Since p is divisible by the minimal polynomial for T, it follows that p is the minimal polynomial.

This example suggests the following converse.

Theorem 18. *Let T be a normal operator on a finite-dimensional real inner product space V and p its minimal polynomial. Suppose*

$$p = (x - a)^2 + b^2$$

where a and b are real and $b \neq 0$. *Then there is an integer* $s > 0$ *such that* p^s *is the characteristic polynomial for T, and there exist subspaces* $V_1, \ldots, V_s$ *of V such that*

(i) V_j *is orthogonal to* V_i *when* $i \neq j$;
(ii) $V = V_1 \oplus \cdots \oplus V_s$;
(iii) *each* V_j *has an orthonormal basis* $\{\alpha_j, \beta_j\}$ *with the property that*

$$T\alpha_j = a\alpha_j + b\beta_j$$
$$T\beta_j = -b\alpha_j + a\beta_j.$$

In other words, if $r = \sqrt{a^2 + b^2}$ and θ is chosen so that $a = r \cos \theta$ and $b = r \sin \theta$, then V is an orthogonal direct sum of two-dimensional subspaces V_j on each of which T acts as 'r times rotation through the angle θ'.

The proof of Theorem 18 will be based on the following result.

Lemma. *Let V be a real inner product space and S a normal operator on V such that* $S^2 + I = 0$. *Let* α *be any vector in V and* $\beta = S\alpha$. *Then*

$$S^*\alpha = -\beta \tag{9-17}$$
$$S^*\beta = \alpha$$

$(\alpha|\beta) = 0$, *and* $||\alpha|| = ||\beta||$.

Proof. We have $S\alpha = \beta$ and $S\beta = S^2\alpha = -\alpha$. Therefore

$$0 = ||S\alpha - \beta||^2 + ||S\beta + \alpha||^2 = ||S\alpha||^2 - 2(S\alpha|\beta) + ||\beta||^2 + ||S\beta||^2 + 2(S\beta|\alpha) + ||\alpha||^2.$$

Since S is normal, it follows that

$$0 = ||S^*\alpha||^2 - 2(S^*\beta|\alpha) + ||\beta||^2 + ||S^*\beta||^2 + 2(S^*\alpha|\beta) + ||\alpha||^2 = ||S^*\alpha + \beta||^2 + ||S^*\beta - \alpha||^2.$$

This implies (9-17); hence

$$\begin{aligned}(\alpha|\beta) &= (S^*\beta|\beta) = (\beta|S\beta) \\ &= (\beta|-\alpha) \\ &= -(\alpha|\beta)\end{aligned}$$

and $(\alpha|\beta) = 0$. Similarly

$$||\alpha||^2 = (S^*\beta|\alpha) = (\beta|S\alpha) = ||\beta||^2. \quad \blacksquare$$

Proof of Theorem 18. Let $V_1, \ldots, V_s$ be a maximal collection of two-dimensional subspaces satisfying (i) and (ii), and the additional conditions