(c) Find the Jordan form of A.

(d) Find a direct-sum decomposition of R^3 into T-cyclic subspaces as in Theorem 3. (*Hint:* One way to do this is to use the results in (b) and an appropriate generalization of the ideas discussed in Example 4.)

7.5. Summary; Semi-Simple Operators

In the last two chapters, we have been dealing with a single linear operator T on a finite-dimensional vector space V. The program has been to decompose T into a direct sum of linear operators of an elementary nature, for the purpose of gaining detailed information about how T 'operates' on the space V. Let us review briefly where we stand.

We began to study T by means of characteristic values and characteristic vectors. We introduced diagonalizable operators, the operators which can be completely described in terms of characteristic values and vectors. We then observed that T might not have a single characteristic vector. Even in the case of an algebraically closed scalar field, when every linear operator does have at least one characteristic vector, we noted that the characteristic vectors of T need not span the space.

We then proved the cyclic decomposition theorem, expressing any linear operator as the direct sum of operators with a cyclic vector, with no assumption about the scalar field. If U is a linear operator with a cyclic vector, there is a basis $\{\alpha_1, \ldots, \alpha_n\}$ with

$$\begin{aligned} U\alpha_j &= \alpha_{j+1}, \qquad j = 1, \ldots, n-1 \\ U\alpha_n &= -c_0\alpha_1 - c_1\alpha_2 - \cdots - c_{n-1}\alpha_n. \end{aligned}$$

The action of U on this basis is then to shift each α_j to the next vector α_{j+1}, except that $U\alpha_n$ is some prescribed linear combination of the vectors in the basis. Since the general linear operator T is the direct sum of a finite number of such operators U, we obtained an explicit and reasonably elementary description of the action of T.

We next applied the cyclic decomposition theorem to nilpotent operators. For the case of an algebraically closed scalar field, we combined this with the primary decomposition theorem to obtain the Jordan form. The Jordan form gives a basis $\{\alpha_1, \ldots, \alpha_n\}$ for the space V such that, for each j, either $T\alpha_j$ is a scalar multiple of α_j or $T\alpha_j = c\alpha_j + \alpha_{j+1}$. Such a basis certainly describes the action of T in an explicit and elementary manner.

The importance of the rational form (or the Jordan form) derives from the fact that it exists, rather than from the fact that it can be computed in specific cases. Of course, if one is given a specific linear operator T and can compute its cyclic or Jordan form, that is the thing to do; for, having such a form, one can reel off vast amounts of information about T. Two different types of difficulties arise in the computation of such standard forms. One difficulty is, of course, the length of the computations. The other difficulty is that there may not be any method for doing the computations, even if one has the necessary time and patience. The second difficulty arises in, say, trying to find the Jordan form of a complex matrix. There simply is no well-defined method for factoring the characteristic polynomial, and thus one is stopped at the outset. The rational form does not suffer from this difficulty. As we showed in Section 7.4, there is a well-defined method for finding the rational form of a given $n \times n$ matrix; however, such computations are usually extremely lengthy.

In our summary of the results of these last two chapters, we have not yet mentioned one of the theorems which we proved. This is the theorem which states that if T is a linear operator on a finite-dimensional vector space over an algebraically closed field, then T is uniquely expressible as the sum of a diagonalizable operator and a nilpotent operator which commute. This was proved from the primary decomposition theorem and certain information about diagonalizable operators. It is not as deep a theorem as the cyclic decomposition theorem or the existence of the Jordan form, but it does have important and useful applications in certain parts of mathematics. In concluding this chapter, we shall prove an analogous theorem, without assuming that the scalar field is algebraically closed. We begin by defining the operators which will play the role of the diagonalizable operators.

Definition. *Let V be a finite-dimensional vector space over the field F, and let T be a linear operator on V. We say that T is* **semi-simple** *if every T-invariant subspace has a complementary T-invariant subspace.*

What we are about to prove is that, with some restriction on the field F, every linear operator T is uniquely expressible in the form $T = S + N$, where S is semi-simple, N is nilpotent, and $SN = NS$. First, we are going to characterize semi-simple operators by means of their minimal polynomials, and this characterization will show us that, when F is algebraically closed, an operator is semi-simple if and only if it is diagonalizable.

Lemma. *Let T be a linear operator on the finite-dimensional vector space V, and let $V = W_1 \oplus \cdots \oplus W_k$ be the primary decomposition for T. In other words, if p is the minimal polynomial for T and $p = p_1^{r_1} \cdots p_k^{r_k}$ is the prime factorization of p, then W_j is the null space of $p_j(T)^{r_j}$. Let W be any subspace of V which is invariant under T. Then*

$$W = (W \cap W_1) \oplus \cdots \oplus (W \cap W_k)$$

Proof. For the proof we need to recall a corollary to our proof of the primary decomposition theorem in Section 6.8. If $E_1, \ldots, E_k$ are the projections associated with the decomposition $V = W_1 \oplus \cdots \oplus W_k$, then each E_j is a polynomial in T. That is, there are polynomials $h_1, \ldots, h_k$ such that $E_j = h_j(T)$.

Now let W be a subspace which is invariant under T. If α is any vector in W, then $\alpha = \alpha_1 + \cdots + \alpha_k$, where α_j is in W_j. Now $\alpha_j = E_j\alpha = h_j(T)\alpha$, and since W is invariant under T, each α_j is also in W. Thus each vector α in W is of the form $\alpha = \alpha_1 + \cdots + \alpha_k$, where α_j is in the intersection $W \cap W_j$. This expression is unique, since $V = W_1 \oplus \cdots \oplus W_k$. Therefore

$$W = (W \cap W_1) \oplus \cdots \oplus (W \cap W_k). \quad \blacksquare$$

Lemma. *Let* T *be a linear operator on* V, *and suppose that the minimal polynomial for* T *is irreducible over the scalar field* F. *Then* T *is semi-simple.*

Proof. Let W be a subspace of V which is invariant under T. We must prove that W has a complementary T-invariant subspace. According to a corollary of Theorem 3, it will suffice to prove that if f is a polynomial and β is a vector in V such that $f(T)\beta$ is in W, then there is a vector α in W with $f(T)\beta = f(T)\alpha$. So suppose β is in V and f is a polynomial such that $f(T)\beta$ is in W. If $f(T)\beta = 0$, we let $\alpha = 0$ and then α is a vector in W with $f(T)\beta = f(T)\alpha$. If $f(T)\beta \neq 0$, the polynomial f is not divisible by the minimal polynomial p of the operator T. Since p is prime, this means that f and p are relatively prime, and there exist polynomials g and h such that $fg + ph = 1$. Because $p(T) = 0$, we then have $f(T)g(T) = I$. From this it follows that the vector β must itself be in the subspace W; for

$$\begin{aligned}\beta &= g(T)f(T)\beta \\ &= g(T)(f(T)\beta)\end{aligned}$$

while $f(T)\beta$ is in W and W is invariant under T. Take $\alpha = \beta$. $\blacksquare$

Theorem 11. *Let* T *be a linear operator on the finite-dimensional vector space* V. *A necessary and sufficient condition that* T *be semi-simple is that the minimal polynomial* p *for* T *be of the form* $\mathrm{p} = \mathrm{p}_1 \cdots \mathrm{p}_k$, *where* $\mathrm{p}_1, \ldots, \mathrm{p}_k$ *are distinct irreducible polynomials over the scalar field* F.

Proof. Suppose T is semi-simple. We shall show that no irreducible polynomial is repeated in the prime factorization of the minimal polynomial p. Suppose the contrary. Then there is some non-scalar monic polynomial g such that g^2 divides p. Let W be the null space of the operator $g(T)$. Then W is invariant under T. Now $p = g^2h$ for some polynomial h. Since g is not a scalar polynomial, the operator $g(T)h(T)$ is not the zero operator, and there is some vector β in V such that $g(T)h(T)\beta \neq 0$, i.e., $(gh)\beta \neq 0$. Now $(gh)\beta$ is in the subspace W, since $g(gh\beta) = g^2h\beta = p\beta = 0$. But there is no vector α in W such that $gh\beta = gh\alpha$; for, if α is in W

$$(gh)\alpha = (hg)\alpha = h(g\alpha) = h(0) = 0.$$

Thus, W cannot have a complementary T-invariant subspace, contradicting the hypothesis that T is semi-simple.

Now suppose the prime factorization of p is $p = p_1 \cdots p_k$, where $p_1, \ldots, p_k$ are distinct irreducible (non-scalar) monic polynomials. Let W be a subspace of V which is invariant under T. We shall prove that W has a complementary T-invariant subspace. Let $V = W_1 \oplus \cdots \oplus W_k$ be the primary decomposition for T, i.e., let W_j be the null space of $p_j(T)$. Let T_j be the linear operator induced on W_j by T, so that the minimal polynomial for T_j is the prime p_j. Now $W \cap W_j$ is a subspace of W_j which is invariant under T_j (or under T). By the last lemma, there is a subspace V_j of W_j such that $W_j = (W \cap W_j) \oplus V_j$ and V_j is invariant under T_j (and hence under T). Then we have

$$\begin{aligned} V &= W_1 \oplus \cdots \oplus W_k \\ &= (W \cap W_1) \oplus V_1 \oplus \cdots \oplus (W \cap W_k) \oplus V_k \\ &= (W \cap W_1) + \cdots + (W \cap W_k) \oplus V_1 \oplus \cdots \oplus V_k. \end{aligned}$$

By the first lemma above, $W = (W \cap W_1) \oplus \cdots \oplus (W \cap W_k)$, so that if $W' = V_1 \oplus \cdots \oplus V_k$, then $V = W \oplus W'$ and W' is invariant under T. ▮

Corollary. *If* T *is a linear operator on a finite-dimensional vector space over an algebraically closed field, then* T *is semi-simple if and only if* T *is diagonalizable.*

Proof. If the scalar field F is algebraically closed, the monic primes over F are the polynomials $x - c$. In this case, T is semi-simple if and only if the minimal polynomial for T is $p = (x - c_1) \cdots (x - c_k)$, where $c_1, \ldots, c_k$ are distinct elements of F. This is precisely the criterion for T to be diagonalizable, which we established in Chapter 6. ▮

We should point out that T is semi-simple if and only if there is some polynomial f, which is a product of distinct primes, such that $f(T) = 0$. This is only superficially different from the condition that the minimal polynomial be a product of distinct primes.

We turn now to expressing a linear operator as the sum of a semi-simple operator and a nilpotent operator which commute. In this, we shall restrict the scalar field to a subfield of the complex numbers. The informed reader will see that what is important is that the field F be a field of characteristic zero, that is, that for each positive integer n the sum $1 + \cdots + 1$ (n times) in F should not be 0. For a polynomial f over F, we denote by $f^{(k)}$ the kth formal derivative of f. In other words, $f^{(k)} = D^k f$, where D is the differentiation operator on the space of polynomials. If g is another polynomial, $f(g)$ denotes the result of substituting g in f, i.e., the polynomial obtained by applying f to the element g in the linear algebra $F[x]$.