For each i, let

$$f_i = \frac{p}{p_i^{r_i}} = \prod_{j \neq i} p_j^{r_j}.$$

Since $p_1, \ldots, p_k$ are distinct prime polynomials, the polynomials $f_1, \ldots, f_k$ are relatively prime (Theorem 10, Chapter 4). Thus there are polynomials $g_1, \ldots, g_k$ such that

$$\sum_{i=1}^{n} f_i g_i = 1.$$

Note also that if $i \neq j$, then $f_i f_j$ is divisible by the polynomial p, because $f_i f_j$ contains each $p_m^{r_m}$ as a factor. We shall show that the polynomials $h_i = f_i g_i$ behave in the manner described in the first paragraph of the proof.

Let $E_i = h_i(T) = f_i(T)g_i(T)$. Since $h_1 + \cdots + h_k = 1$ and p divides $f_i f_j$ for $i \neq j$, we have

$$\begin{aligned} E_1 + \cdots + E_k &= I \\ E_i E_j &= 0, \qquad \text{if } i \neq j. \end{aligned}$$

Thus the E_i are projections which correspond to some direct-sum decomposition of the space V. We wish to show that the range of E_i is exactly the subspace W_i. It is clear that each vector in the range of E_i is in W_i, for if α is in the range of E_i, then $\alpha = E_i\alpha$ and so

$$\begin{aligned} p_i(T)^{r_i}\alpha &= p_i(T)^{r_i}E_i\alpha \\ &= p_i(T)^{r_i}f_i(T)g_i(T)\alpha \\ &= 0 \end{aligned}$$

because $p^{r_i}f_i g_i$ is divisible by the minimal polynomial p. Conversely, suppose that α is in the null space of $p_i(T)^{r_i}$. If $j \neq i$, then $f_j g_j$ is divisible by $p_i^{r_i}$ and so $f_j(T)g_j(T)\alpha = 0$, i.e., $E_j\alpha = 0$ for $j \neq i$. But then it is immediate that $E_i\alpha = \alpha$, i.e., that α is in the range of E_i. This completes the proof of statement (i).

It is certainly clear that the subspaces W_i are invariant under T. If T_i is the operator induced on W_i by T, then evidently $p_i(T_i)^{r_i} = 0$, because by definition $p_i(T)^{r_i}$ is 0 on the subspace W_i. This shows that the minimal polynomial for T_i divides $p_i^{r_i}$. Conversely, let g be any polynomial such that $g(T_i) = 0$. Then $g(T)f_i(T) = 0$. Thus gf_i is divisible by the minimal polynomial p of T, i.e., $p_i^{r_i}f_i$ divides gf_i. It is easily seen that $p_i^{r_i}$ divides g. Hence the minimal polynomial for T_i is $p_i^{r_i}$. ∎

Corollary. *If* $E_1, \ldots, E_k$ *are the projections associated with the primary decomposition of* T, *then each* E_i *is a polynomial in* T, *and accordingly if a linear operator* U *commutes with* T *then* U *commutes with each of the* E_i, *i.e., each subspace* W_i *is invariant under* U.

In the notation of the proof of Theorem 12, let us take a look at the special case in which the minimal polynomial for T is a product of first-degree polynomials, i.e., the case in which each p_i is of the form $p_i = x - c_i$. Now the range of E_i is the null space W_i of $(T - c_iI)^{r_i}$. Let us put $D = c_1E_1 + \cdots + c_kE_k$. By Theorem 11, D is a diagonalizable operator which we shall call the **diagonalizable part** of T. Let us look at the operator $N = T - D$. Now

$$T = TE_1 + \cdots + TE_k$$
$$D = c_1E_1 + \cdots + c_kE_k$$

so

$$N = (T - c_1I)E_1 + \cdots + (T - c_kI)E_k.$$

The reader should be familiar enough with projections by now so that he sees that

$$N^2 = (T - c_1I)^2E_1 + \cdots + (T - c_kI)^2E_k$$

and in general that

$$N^r = (T - c_1I)^rE_1 + \cdots + (T - c_kI)^rE_k.$$

When $r \geq r_i$ for each i, we shall have $N^r = 0$, because the operator $(T - c_iI)^r$ will then be 0 on the range of E_i.

Definition. *Let* N *be a linear operator on the vector space* V. *We say that* N *is* **nilpotent** *if there is some positive integer* r *such that* $N^r = 0$.

Theorem 13. *Let* T *be a linear operator on the finite-dimensional vector space* V *over the field* F. *Suppose that the minimal polynomial for* T *decomposes over* F *into a product of linear polynomials. Then there is a diagonalizable operator* D *on* V *and a nilpotent operator* N *on* V *such that*

(i) $T = D + N$,
(ii) $DN = ND$.

The diagonalizable operator D *and the nilpotent operator* N *are uniquely determined by* (i) *and* (ii) *and each of them is a polynomial in* T.

Proof. We have just observed that we can write $T = D + N$ where D is diagonalizable and N is nilpotent, and where D and N not only commute but are polynomials in T. Now suppose that we also have $T = D' + N'$ where D' is diagonalizable, N' is nilpotent, and $D'N' = N'D'$. We shall prove that $D = D'$ and $N = N'$.

Since D' and N' commute with one another and $T = D' + N'$, we see that D' and N' commute with T. Thus D' and N' commute with any polynomial in T; hence they commute with D and with N. Now we have

$$D + N = D' + N'$$

or

$$D - D' = N' - N$$

and all four of these operators commute with one another. Since D and D' are both diagonalizable and they commute, they are simultaneously diagonalizable, and $D - D'$ is diagonalizable. Since N and N' are both nilpotent and they commute, the operator $(N' - N)$ is nilpotent; for, using the fact that N and N' commute

$$(N' - N)^r = \sum_{j=0}^{r} \binom{r}{j} (N')^{r-j}(-N)^j$$

and so when r is sufficiently large every term in this expression for $(N' - N)^r$ will be 0. (Actually, a nilpotent operator on an n-dimensional space must have its nth power 0; if we take $r = 2n$ above, that will be large enough. It then follows that $r = n$ is large enough, but this is not obvious from the above expression.) Now $D - D'$ is a diagonalizable operator which is also nilpotent. Such an operator is obviously the zero operator; for since it is nilpotent, the minimal polynomial for this operator is of the form x^r for some $r \leq m$; but then since the operator is diagonalizable, the minimal polynomial cannot have a repeated root; hence $r = 1$ and the minimal polynomial is simply x, which says the operator is 0. Thus we see that $D = D'$ and $N = N'$. ▌

Corollary. *Let* V *be a finite-dimensional vector space over an algebraically closed field* F, *e.g., the field of complex numbers. Then every linear operator* T *on* V *can be written as the sum of a diagonalizable operator* D *and a nilpotent operator* N *which commute. These operators* D *and* N *are unique and each is a polynomial in* T.

From these results, one sees that the study of linear operators on vector spaces over an algebraically closed field is essentially reduced to the study of nilpotent operators. For vector spaces over non-algebraically closed fields, we still need to find some substitute for characteristic values and vectors. It is a very interesting fact that these two problems can be handled simultaneously and this is what we shall do in the next chapter.

In concluding this section, we should like to give an example which illustrates some of the ideas of the primary decomposition theorem. We have chosen to give it at the end of the section since it deals with differential equations and thus is not purely linear algebra.

EXAMPLE 14. In the primary decomposition theorem, it is not necessary that the vector space V be finite dimensional, nor is it necessary for parts (i) and (ii) that p be the minimal polynomial for T. If T is a linear operator on an arbitrary vector space and *if* there is a monic polynomial p such that $p(T) = 0$, then parts (i) and (ii) of Theorem 12 are valid for T with the proof which we gave.

Let n be a positive integer and let V be the space of all n times continuously differentiable functions f on the real line which satisfy the differential equation

$$\frac{d^n f}{dt^n} + a_{n-1}\frac{d^{n-1}f}{dt^{n-1}} + \cdots + a_1\frac{df}{dt} + a_0 f = 0 \tag{6-18}$$

where $a_0, \ldots, a_{n-1}$ are some fixed constants. If C_n denotes the space of n times continuously differentiable functions, then the space V of solutions of this differential equation is a subspace of C_n. If D denotes the differentiation operator and p is the polynomial

$$p = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

then V is the null space of the operator $p(D)$, because (6-18) simply says $p(D)f = 0$. Therefore, V is invariant under D. Let us now regard D as a linear operator on the subspace V. Then $p(D) = 0$.

If we are discussing differentiable complex-valued functions, then C_n and V are complex vector spaces, and $a_0, \ldots, a_{n-1}$ may be any complex numbers. We now write

$$p = (x - c_1)^{r_1} \cdots (x - c_k)^{r_k}$$

where $c_1, \ldots, c_k$ are distinct complex numbers. If W_j is the null space of $(D - c_jI)^{r_j}$, then Theorem 12 says that

$$V = W_1 \oplus \cdots \oplus W_k.$$

In other words, if f satisfies the differential equation (6-18), then f is uniquely expressible in the form

$$f = f_1 + \cdots + f_k$$

where f_j satisfies the differential equation $(D - c_jI)^{r_j}f_j = 0$. Thus, the study of the solutions to the equation (6-18) is reduced to the study of the space of solutions of a differential equation of the form

$$(D - cI)^r f = 0. \tag{6-19}$$

This reduction has been accomplished by the general methods of linear algebra, i.e., by the primary decomposition theorem.

To describe the space of solutions to (6-19), one must know something about differential equations, that is, one must know something about D other than the fact that it is a linear operator. However, one does not need to know very much. It is very easy to establish by induction on r that if f is in C_r then

$$(D - cI)^r f = e^{ct}D^r(e^{-ct}f)$$

that is,

$$\frac{df}{dt} - cf(t) = e^{ct}\frac{d}{dt}(e^{-ct}f), \quad \text{etc.}$$

Thus $(D - cI)^r f = 0$ if and only if $D^r(e^{-ct}f) = 0$. A function g such that $D^r g = 0$, i.e., $d^r g/dt^r = 0$, must be a polynomial function of degree $(r - 1)$ or less:

$$g(t) = b_0 + b_1 t + \cdots + b_{r-1}t^{r-1}.$$

Thus f satisfies (6-19) if and only if f has the form

$$f(t) = e^{ct}(b_0 + b_1 t + \cdots + b_{r-1}t^{r-1}).$$

Accordingly, the 'functions' $e^{ct}, te^{ct}, \ldots, t^{r-1}e^{ct}$ span the space of solutions of (6-19). Since $1, t, \ldots, t^{r-1}$ are linearly independent functions and the exponential function has no zeros, these r functions $t^j e^{ct}$, $0 \leq j \leq r-1$, form a basis for the space of solutions.

Returning to the differential equation (6-18), which is

$$p(D)f = 0$$
$$p = (x - c_1)^{r_1} \cdots (x - c_k)^{r_k}$$

we see that the n functions $t^m e^{c_j t}$, $0 \leq m \leq r_j - 1$, $1 \leq j \leq k$, form a basis for the space of solutions to (6-18). In particular, the space of solutions is finite-dimensional and has dimension equal to the degree of the polynomial p.

Exercises

1. Let T be a linear operator on R^3 which is represented in the standard ordered basis by the matrix

$$\begin{bmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{bmatrix}.$$

Express the minimal polynomial p for T in the form $p = p_1 p_2$, where p_1 and p_2 are monic and irreducible over the field of real numbers. Let W_i be the null space of $p_i(T)$. Find bases $\mathcal{B}_i$ for the spaces W_1 and W_2. If T_i is the operator induced on W_i by T, find the matrix of T_i in the basis $\mathcal{B}_i$ (above).

2. Let T be the linear operator on R^3 which is represented by the matrix

$$\begin{bmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{bmatrix}$$

in the standard ordered basis. Show that there is a diagonalizable operator D on R^3 and a nilpotent operator N on R^3 such that $T = D + N$ and $DN = ND$. Find the matrices of D and N in the standard basis. (Just repeat the proof of Theorem 12 for this special case.)

3. If V is the space of all polynomials of degree less than or equal to n over a field F, prove that the differentiation operator on V is nilpotent.

4. Let T be a linear operator on the finite-dimensional space V with characteristic polynomial

$$f = (x - c_1)^{d_1} \cdots (x - c_k)^{d_k}$$

and minimal polynomial

$$p = (x - c_1)^{r_1} \cdots (x - c_k)^{r_k}.$$

Let W_i be the null space of $(T - c_iI)^{r_i}$.