which sends γ_1 into W_0, we see that g_1 is the monic polynomial of least degree in the ideal $S(V; W_0)$. By the same argument, p_1 is the generator of that ideal, so $p_1 = g_1$.

If f is a polynomial and W is a subspace of V, we shall employ the shorthand fW for the set of all vectors $f\alpha$ with α in W. We have left to the exercises the proofs of the following three facts.

1. $fZ(\alpha; T) = Z(f\alpha; T)$.
2. If $V = V_1 \oplus \cdots \oplus V_k$, where each V_i is invariant under T, then $fV = fV_1 \oplus \cdots \oplus fV_k$.
3. If α and γ have the same T-annihilator, then $f\alpha$ and $f\gamma$ have the same T-annihilator and (therefore)

$$\dim Z(f\alpha; T) = \dim Z(f\gamma; T).$$

Now, we proceed by induction to show that $r = s$ and $p_i = g_i$ for $i = 2, \ldots, r$. The argument consists of counting dimensions in the right way. We shall give the proof that if $r \geq 2$ then $p_2 = g_2$, and from that the induction should be clear. Suppose that $r \geq 2$. Then

$$\dim W_0 + \dim Z(\alpha_1; T) < \dim V.$$

Since we know that $p_1 = g_1$, we know that $Z(\alpha_1; T)$ and $Z(\gamma_1; T)$ have the same dimension. Therefore,

$$\dim W_0 + \dim Z(\gamma_1; T) < \dim V$$

which shows that $s \geq 2$. Now it makes sense to ask whether or not $p_2 = g_2$. From the two decompositions of V, we obtain two decompositions of the subspace p_2V:

$$\begin{aligned} p_2V &= p_2W_0 \oplus Z(p_2\alpha_1; T) \\ p_2V &= p_2W_0 \oplus Z(p_2\gamma_1; T) \oplus \cdots \oplus Z(p_2\gamma_s; T). \end{aligned} \tag{7-14}$$

We have made use of facts (1) and (2) above and we have used the fact that $p_2\alpha_i = 0$, $i \geq 2$. Since we know that $p_1 = g_1$, fact (3) above tells us that $Z(p_2\alpha_1; T)$ and $Z(p_2\gamma_1; T)$ have the same dimension. Hence, it is apparent from (7-14) that

$$\dim Z(p_2\gamma_i; T) = 0, \qquad i \geq 2.$$

We conclude that $p_2\gamma_2 = 0$ and g_2 divides p_2. The argument can be reversed to show that p_2 divides g_2. Therefore $p_2 = g_2$. ∎

Corollary. *If* T *is a linear operator on a finite-dimensional vector space, then every* T*-admissible subspace has a complementary subspace which is also invariant under* T.

Proof. Let W_0 be an admissible subspace of V. If $W_0 = V$, the complement we seek is $\{0\}$. If W_0 is proper, apply Theorem 3 and let

$$W_0' = Z(\alpha_1; T) \oplus \cdots \oplus Z(\alpha_r; T).$$

Then W_0' is invariant under T and $V = W_0 \oplus W_0'$. ∎

Corollary. *Let* T *be a linear operator on a finite-dimensional vector space* V.

(a) *There exists a vector* α *in* V *such that the* T*-annihilator of* α *is the minimal polynomial for* T.

(b) T *has a cyclic vector if and only if the characteristic and minimal polynomials for* T *are identical.*

Proof. If $V = \{0\}$, the results are trivially true. If $V \neq \{0\}$, let

$$V = Z(\alpha_1; T) \oplus \cdots \oplus Z(\alpha_r; T) \tag{7-15}$$

where the T-annihilators $p_1, \ldots, p_r$ are such that p_{k+1} divides p_k, $1 \leq k \leq r - 1$. As we noted in the proof of Theorem 3, it follows easily that p_1 is the minimal polynomial for T, i.e., the T-conductor of V into $\{0\}$. We have proved (a).

We saw in Section 7.1 that, if T has a cyclic vector, the minimal polynomial for T coincides with the characteristic polynomial. The content of (b) is in the converse. Choose any α as in (a). If the degree of the minimal polynomial is dim V, then $V = Z(\alpha; T)$. ∎

Theorem 4 (Generalized Cayley-Hamilton Theorem). *Let* T *be a linear operator on a finite-dimensional vector space* V. *Let* p *and* f *be the minimal and characteristic polynomials for* T, *respectively.*

(i) p *divides* f.

(ii) p *and* f *have the same prime factors, except for multiplicities.*

(iii) *If*

$$\mathrm{p} = \mathrm{f}_1^{r_1} \cdots \mathrm{f}_k^{r_k} \tag{7-16}$$

is the prime factorization of p, *then*

$$\mathrm{f} = \mathrm{f}_1^{d_1} \cdots \mathrm{f}_k^{d_k} \tag{7-17}$$

where d_i *is the nullity of* $\mathrm{f}_i(\mathrm{T})^{r_i}$ *divided by the degree of* f_i.

Proof. We disregard the trivial case $V = \{0\}$. To prove (i) and (ii), consider a cyclic decomposition (7-15) of V obtained from Theorem 3. As we noted in the proof of the second corollary, $p_1 = p$. Let U_i be the restriction of T to $Z(\alpha_i; T)$. Then U_i has a cyclic vector and so p_i is both the minimal polynomial and the characteristic polynomial for U_i. Therefore, the characteristic polynomial f is the product $f = p_1 \cdots p_r$. That is evident from the block form (6-14) which the matrix of T assumes in a suitable basis. Clearly $p_1 = p$ divides f, and this proves (i). Obviously any prime divisor of p is a prime divisor of f. Conversely, a prime divisor of $f = p_1 \cdots p_r$ must divide one of the factors p_i, which in turn divides p_1.

Let (7-16) be the prime factorization of p. We employ the primary decomposition theorem (Theorem 12 of Chapter 6). It tells us that, if V_i is the null space of $f_i(T)^{r_i}$, then

(7-18) $$V = V_1 \oplus \cdots \oplus V_k$$

and $f_i^{r_i}$ is the minimal polynomial of the operator T_i, obtained by restricting T to the (invariant) subspace V_i. Apply part (ii) of the present theorem to the operator T_i. Since its minimal polynomial is a power of the prime f_i, the characteristic polynomial for T_i has the form $f_i^{d_i}$, where $d_i \geq r_i$. Obviously

$$d_i = \frac{\dim V_i}{\deg f_i}$$

and (almost by definition) $\dim V_i = \text{nullity } f_i(T)^{r_i}$. Since T is the direct sum of the operators $T_1, \ldots, T_k$, the characteristic polynomial f is the product

$$f = f_1^{d_1} \cdots f_k^{d_k}. \quad \blacksquare$$

Corollary. *If* T *is a nilpotent linear operator on a vector space of dimension* n, *then the characteristic polynomial for* T *is* x^n.

Now let us look at the matrix analogue of the cyclic decomposition theorem. If we have the operator T and the direct-sum decomposition of Theorem 3, let $\mathcal{B}_i$ be the 'cyclic ordered basis'

$$\{\alpha_i, T\alpha_i, \ldots, T^{k_i-1}\alpha_i\}$$

for $Z(\alpha_i; T)$. Here k_i denotes the dimension of $Z(\alpha_i; T)$, that is, the degree of the annihilator p_i. The matrix of the induced operator T_i in the ordered basis $\mathcal{B}_i$ is the companion matrix of the polynomial p_i. Thus, if we let $\mathcal{B}$ be the ordered basis for V which is the union of the $\mathcal{B}_i$ arranged in the order $\mathcal{B}_1, \ldots, \mathcal{B}_r$, then the matrix of T in the ordered basis $\mathcal{B}$ will be

(7-19) $$A = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & A_r \end{bmatrix}$$

where A_i is the $k_i \times k_i$ companion matrix of p_i. An $n \times n$ matrix A, which is the direct sum (7-19) of companion matrices of non-scalar monic polynomials $p_1, \ldots, p_r$ such that p_{i+1} divides p_i for $i = 1, \ldots, r-1$, will be said to be in **rational form.** The cyclic decomposition theorem tells us the following concerning matrices.

Theorem 5. *Let* F *be a field and let* B *be an* n × n *matrix over* F. *Then* B *is similar over the field* F *to one and only one matrix which is in rational form.*

Proof. Let T be the linear operator on F^n which is represented by B in the standard ordered basis. As we have just observed, there is some ordered basis for F^n in which T is represented by a matrix A in rational form. Then B is similar to this matrix A. Suppose B is similar over F to

another matrix C which is in rational form. This means simply that there is some ordered basis for F^n in which the operator T is represented by the matrix C. If C is the direct sum of companion matrices C_i of monic polynomials $g_1, \ldots, g_s$ such that g_{i+1} divides g_i for $i = 1, \ldots, s-1$, then it is apparent that we shall have non-zero vectors $\beta_1, \ldots, \beta_s$ in V with T-annihilators $g_1, \ldots, g_s$ such that

$$V = Z(\beta_1; T) \oplus \cdots \oplus Z(\beta_s; T).$$

But then by the uniqueness statement in the cyclic decomposition theorem, the polynomials g_i are identical with the polynomials p_i which define the matrix A. Thus $C = A$. ∎

The polynomials $p_1, \ldots, p_r$ are called the **invariant factors** for the matrix B. In Section 7.4, we shall describe an algorithm for calculating the invariant factors of a given matrix B. The fact that it is possible to compute these polynomials by means of a finite number of rational operations on the entries of B is what gives the rational form its name.

EXAMPLE 2. Suppose that V is a two-dimensional vector space over the field F and T is a linear operator on V. The possibilities for the cyclic subspace decomposition for T are very limited. For, if the minimal polynomial for T has degree 2, it is equal to the characteristic polynomial for T and T has a cyclic vector. Thus there is some ordered basis for V in which T is represented by the companion matrix of its characteristic polynomial. If, on the other hand, the minimal polynomial for T has degree 1, then T is a scalar multiple of the identity operator. If $T = cI$, then for any two linear independent vectors α_1 and α_2 in V we have

$$V = Z(\alpha_1; T) \oplus Z(\alpha_2; T)$$
$$p_1 = p_2 = x - c.$$

For matrices, this analysis says that every 2×2 matrix over the field F is similar over F to exactly one matrix of the types

$$\begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix}, \qquad \begin{bmatrix} 0 & -c_0 \\ 1 & -c_1 \end{bmatrix}.$$

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

$$A = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}$$

in the standard ordered basis. We have computed earlier that the characteristic polynomial for T is $f = (x-1)(x-2)^2$ and the minimal polynomial for T is $p = (x-1)(x-2)$. Thus we know that in the cyclic decomposition for T the first vector α_1 will have p as its T-annihilator.

Since we are operating in a three-dimensional space, there can be only one further vector, α_2. It must generate a cyclic subspace of dimension 1, i.e., it must be a characteristic vector for T. Its T-annihilator p_2 must be $(x - 2)$, because we must have $pp_2 = f$. Notice that this tells us immediately that the matrix A is similar to the matrix

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

that is, that T is represented by B in some ordered basis. How can we find suitable vectors α_1 and α_2? Well, we know that any vector which generates a T-cyclic subspace of dimension 2 is a suitable α_1. So let's just try ϵ_1. We have

$$T\epsilon_1 = (5, -1, 3)$$

which is not a scalar multiple of ϵ_1; hence $Z(\epsilon_1; T)$ has dimension 2. This space consists of all vectors $a\epsilon_1 + b(T\epsilon_1)$:

$$a(1, 0, 0) + b(5, -1, 3) = (a + 5b, -b, 3b)$$

or, all vectors (x_1, x_2, x_3) satisfying $x_3 = -3x_2$. Now what we want is a vector α_2 such that $T\alpha_2 = 2\alpha_2$ and $Z(\alpha_2; T)$ is disjoint from $Z(\epsilon_1; T)$. Since α_2 is to be a characteristic vector for T, the space $Z(\alpha_2; T)$ will simply be the one-dimensional space spanned by α_2, and so what we require is that α_2 not be in $Z(\epsilon_1; T)$. If $\alpha = (x_1, x_2, x_3)$, one can easily compute that $T\alpha = 2\alpha$ if and only if $x_1 = 2x_2 + 2x_3$. Thus $\alpha_2 = (2, 1, 0)$ satisfies $T\alpha_2 = 2\alpha_2$ and generates a T-cyclic subspace disjoint from $Z(\epsilon_1; T)$. The reader should verify directly that the matrix of T in the ordered basis

$$\{(1, 0, 0), (5, -1, 3), (2, 1, 0)\}$$

is the matrix B above.

EXAMPLE 4. Suppose that T is a diagonalizable linear operator on V. It is interesting to relate a cyclic decomposition for T to a basis which diagonalizes the matrix of T. Let $c_1, \ldots, c_k$ be the distinct characteristic values of T and let V_i be the space of characteristic vectors associated with the characteristic value c_i. Then

$$V = V_1 \oplus \cdots \oplus V_k$$

and if $d_i = \dim V_i$ then

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

is the characteristic polynomial for T. If α is a vector in V, it is easy to relate the cyclic subspace $Z(\alpha; T)$ to the subspaces $V_1, \ldots, V_k$. There are unique vectors $\beta_1, \ldots, \beta_k$ such that β_i is in V_i and

$$\alpha = \beta_1 + \cdots + \beta_k.$$