6. Let T be the linear operator on R^4 which is represented in the standard ordered basis by the matrix

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \end{bmatrix}.$$

Under what conditions on a, b, and c is T diagonalizable?

7. Let T be a linear operator on the n-dimensional vector space V, and suppose that T has n *distinct* characteristic values. Prove that T is diagonalizable.

8. Let A and B be $n \times n$ matrices over the field F. Prove that if $(I - AB)$ is invertible, then $I - BA$ is invertible and

$$(I - BA)^{-1} = I + B(I - AB)^{-1}A.$$

9. Use the result of Exercise 8 to prove that, if A and B are $n \times n$ matrices over the field F, then AB and BA have precisely the same characteristic values in F.

10. Suppose that A is a 2×2 matrix with real entries which is symmetric ($A^t = A$). Prove that A is similar over R to a diagonal matrix.

11. Let N be a 2×2 complex matrix such that $N^2 = 0$. Prove that either $N = 0$ or N is similar over C to

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

12. Use the result of Exercise 11 to prove the following: If A is a 2×2 matrix with complex entries, then A is similar over C to a matrix of one of the two types

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \qquad \begin{bmatrix} a & 0 \\ 1 & a \end{bmatrix}.$$

13. Let V be the vector space of all functions from R into R which are continuous, i.e., the space of continuous real-valued functions on the real line. Let T be the linear operator on V defined by

$$(Tf)(x) = \int_0^x f(t)\, dt.$$

Prove that T has no characteristic values.

14. Let A be an $n \times n$ *diagonal* matrix with characteristic polynomial

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

where $c_1, \ldots, c_k$ are distinct. Let V be the space of $n \times n$ matrices B such that $AB = BA$. Prove that the dimension of V is $d_1^2 + \cdots + d_k^2$.

15. Let V be the space of $n \times n$ matrices over F. Let A be a fixed $n \times n$ matrix over F. Let T be the linear operator 'left multiplication by A' on V. Is it true that A and T have the same characteristic values?

6.3. Annihilating Polynomials

In attempting to analyze a linear operator T, one of the most useful things to know is the class of polynomials which annihilate T. Specifically,

suppose T is a linear operator on V, a vector space over the field F. If p is a polynomial over F, then $p(T)$ is again a linear operator on V. If q is another polynomial over F, then

$$(p + q)(T) = p(T) + q(T)$$
$$(pq)(T) = p(T)q(T).$$

Therefore, the collection of polynomials p which annihilate T, in the sense that

$$p(T) = 0,$$

is an ideal in the polynomial algebra $F[x]$. It may be the zero ideal, i.e., it may be that T is not annihilated by any non-zero polynomial. But, that cannot happen if the space V is finite-dimensional.

Suppose T is a linear operator on the n-dimensional space V. Look at the first $(n^2 + 1)$ powers of T:

$$I, T, T^2, \ldots, T^{n^2}.$$

This is a sequence of $n^2 + 1$ operators in $L(V, V)$, the space of linear operators on V. The space $L(V, V)$ has dimension n^2. Therefore, that sequence of $n^2 + 1$ operators must be linearly dependent, i.e., we have

$$c_0 I + c_1 T + \cdots + c_{n^2} T^{n^2} = 0$$

for some scalars c_i, not all zero. So, the ideal of polynomials which annihilate T contains a non-zero polynomial of degree n^2 or less.

According to Theorem 5 of Chapter 4, every polynomial ideal consists of all multiples of some fixed monic polynomial, the generator of the ideal. Thus, there corresponds to the operator T a monic polynomial p with this property: If f is a polynomial over F, then $f(T) = 0$ if and only if $f = pg$, where g is some polynomial over F.

Definition. *Let* T *be a linear operator on a finite-dimensional vector space* V *over the field* F. *The* **minimal polynomial** *for* T *is the (unique) monic generator of the ideal of polynomials over* F *which annihilate* T.

The name 'minimal polynomial' stems from the fact that the generator of a polynomial ideal is characterized by being the monic polynomial of minimum degree in the ideal. That means that the minimal polynomial p for the linear operator T is uniquely determined by these three properties:

(1) p is a monic polynomial over the scalar field F.
(2) $p(T) = 0$.
(3) No polynomial over F which annihilates T has smaller degree than p has.

If A is an $n \times n$ matrix over F, we define the **minimal polynomial** for A in an analogous way, as the unique monic generator of the ideal of all polynomials over F which annihilate A. If the operator T is represented in

some ordered basis by the matrix A, then T and A have the same minimal polynomial. That is because $f(T)$ is represented in the basis by the matrix $f(A)$, so that $f(T) = 0$ if and only if $f(A) = 0$.

From the last remark about operators and matrices it follows that similar matrices have the same minimal polynomial. That fact is also clear from the definitions because

$$f(P^{-1}AP) = P^{-1}f(A)P$$

for every polynomial f.

There is another basic remark which we should make about minimal polynomials of matrices. Suppose that A is an $n \times n$ matrix with entries in the field F. Suppose that F_1 is a field which contains F as a subfield. (For example, A might be a matrix with rational entries, while F_1 is the field of real numbers. Or, A might be a matrix with real entries, while F_1 is the field of complex numbers.) We may regard A either as an $n \times n$ matrix over F or as an $n \times n$ matrix over F_1. On the surface, it might appear that we obtain two different minimal polynomials for A. Fortunately that is not the case; and we must see why. What is the definition of the minimal polynomial for A, regarded as an $n \times n$ matrix over the field F? We consider all monic polynomials with coefficients in F which annihilate A, and we choose the one of least degree. If f is a monic polynomial over F:

$$(6\text{-}4) \qquad\qquad f = x^k + \sum_{j=0}^{k-1} a_j x^j$$

then $f(A) = 0$ merely says that we have a linear relation between the powers of A:

$$(6\text{-}5) \qquad\quad A^k + a_{k-1}A^{k-1} + \cdots + a_1 A + a_0 I = 0.$$

The degree of the minimal polynomial is the least positive integer k such that there is a linear relation of the form (6-5) between the powers I, $A, \ldots, A^k$. Furthermore, by the uniqueness of the minimal polynomial, there is for that k one and only one relation of the form (6-5); i.e., once the minimal k is determined, there are unique scalars $a_0, \ldots, a_{k-1}$ in F such that (6-5) holds. They are the coefficients of the minimal polynomial.

Now (for each k) we have in (6-5) a system of n^2 linear equations for the 'unknowns' $a_0, \ldots, a_{k-1}$. Since the entries of A lie in F, the coefficients of the system of equations (6-5) are in F. Therefore, if the system has a solution with $a_0, \ldots, a_{k-1}$ in F_1 it has a solution with $a_0, \ldots, a_{k-1}$ in F. (See the end of Section 1.4.) It should now be clear that the two minimal polynomials are the same.

What do we know thus far about the minimal polynomial for a linear operator on an n-dimensional space? Only that its degree does not exceed n^2. That turns out to be a rather poor estimate, since the degree cannot exceed n. We shall prove shortly that the operator is annihilated by its characteristic polynomial. First, let us observe a more elementary fact.

Theorem 3. *Let* T *be a linear operator on an* n-*dimensional vector space* V [*or, let* A *be an* n × n *matrix*]. *The characteristic and minimal polynomials for* T [*for* A] *have the same roots, except for multiplicities.*

Proof. Let p be the minimal polynomial for T. Let c be a scalar. What we want to show is that $p(c) = 0$ if and only if c is a characteristic value of T.

First, suppose $p(c) = 0$. Then

$$p = (x - c)q$$

where q is a polynomial. Since deg $q <$ deg p, the definition of the minimal polynomial p tells us that $q(T) \neq 0$. Choose a vector β such that $q(T)\beta \neq 0$. Let $\alpha = q(T)\beta$. Then

$$\begin{aligned} 0 &= p(T)\beta \\ &= (T - cI)q(T)\beta \\ &= (T - cI)\alpha \end{aligned}$$

and thus, c is a characteristic value of T.

Now, suppose that c is a characteristic value of T, say, $T\alpha = c\alpha$ with $\alpha \neq 0$. As we noted in a previous lemma,

$$p(T)\alpha = p(c)\alpha.$$

Since $p(T) = 0$ and $\alpha \neq 0$, we have $p(c) = 0$. ▮

Let T be a diagonalizable linear operator and let $c_1, \ldots, c_k$ be the distinct characteristic values of T. Then it is easy to see that the minimal polynomial for T is the polynomial

$$p = (x - c_1) \cdots (x - c_k).$$

If α is a characteristic vector, then one of the operators $T - c_1I, \ldots, T - c_kI$ sends α into 0. Therefore

$$(T - c_1I) \cdots (T - c_kI)\alpha = 0$$

for every characteristic vector α. There is a basis for the underlying space which consists of characteristic vectors of T; hence

$$p(T) = (T - c_1I) \cdots (T - c_kI) = 0.$$

What we have concluded is this. If T is a diagonalizable linear operator, then the minimal polynomial for T is a product of distinct linear factors. As we shall soon see, that property characterizes diagonalizable operators.

EXAMPLE 4. Let's try to find the minimal polynomials for the operators in Examples 1, 2, and 3. We shall discuss them in reverse order. The operator in Example 3 was found to be diagonalizable with characteristic polynomial

$$f = (x - 1)(x - 2)^2.$$

From the preceding paragraph, we know that the minimal polynomial for T is
$$p = (x - 1)(x - 2).$$
The reader might find it reassuring to verify directly that
$$(A - I)(A - 2I) = 0.$$

In Example 2, the operator T also had the characteristic polynomial $f = (x - 1)(x - 2)^2$. But, this T is not diagonalizable, so we don't know that the minimal polynomial is $(x - 1)$ $(x - 2)$. What do we know about the minimal polynomial in this case? From Theorem 3 we know that its roots are 1 and 2, with some multiplicities allowed. Thus we search for p among polynomials of the form $(x - 1)^k(x - 2)^l$, $k \geq 1, l \geq 1$. Try $(x - 1)$ $(x - 2)$:

$$(A - I)(A - 2I) = \begin{bmatrix} 2 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & -1 \\ 2 & 2 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 & -1 \\ 2 & 0 & -1 \\ 4 & 0 & -2 \end{bmatrix}.$$

Thus, the minimal polynomial has degree at least 3. So, next we should try either $(x - 1)^2(x - 2)$ or $(x - 1)(x - 2)^2$. The second, being the characteristic polynomial, would seem a less random choice. One can readily compute that $(A - I)(A - 2I)^2 = 0$. Thus the minimal polynomial for T is its characteristic polynomial.

In Example 1 we discussed the linear operator T on R^2 which is represented in the standard basis by the matrix

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

The characteristic polynomial is $x^2 + 1$, which has no real roots. To determine the minimal polynomial, forget about T and concentrate on A. As a complex 2×2 matrix, A has the characteristic values i and $-i$. Both roots must appear in the minimal polynomial. Thus the minimal polynomial is divisible by $x^2 + 1$. It is trivial to verify that $A^2 + I = 0$. So the minimal polynomial is $x^2 + 1$.

Theorem 4 (Cayley-Hamilton). *Let* T *be a linear operator on a finite dimensional vector space* V. *If* f *is the characteristic polynomial for* T, *then* f(T) = 0; *in other words, the minimal polynomial divides the characteristic polynomial for* T.

Proof. Later on we shall give two proofs of this result independent of the one to be given here. The present proof, although short, may be difficult to understand. Aside from brevity, it has the virtue of providing