of T. Let $c_1, \ldots, c_k$ be the distinct characteristic values of T. For each i, let W_i be the space of characteristic vectors associated with the characteristic value c_i, and let $\mathcal{B}_i$ be an ordered basis for W_i. The lemma before Theorem 2 tells us that $\mathcal{B}' = (\mathcal{B}_1, \ldots, \mathcal{B}_k)$ is an ordered basis for W. In particular,

$$\dim W = \dim W_1 + \cdots + \dim W_k.$$

Let $\mathcal{B}' = \{\alpha_1, \ldots, \alpha_r\}$ so that the first few α's form the basis $\mathcal{B}_1$, the next few $\mathcal{B}_2$, and so on. Then

$$T\alpha_i = t_i\alpha_i, \qquad i = 1, \ldots, r$$

where $(t_1, \ldots, t_r) = (c_1, c_1, \ldots, c_1, \ldots, c_k, c_k, \ldots, c_k)$ with c_i repeated $\dim W_i$ times.

Now W is invariant under T, since for each α in W we have

$$\begin{aligned} \alpha &= x_1\alpha_1 + \cdots + x_r\alpha_r \\ T\alpha &= t_1x_1\alpha_1 + \cdots + t_rx_r\alpha_r. \end{aligned}$$

Choose any other vectors $\alpha_{r+1}, \ldots, \alpha_n$ in V such that $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ is a basis for V. The matrix of T relative to $\mathcal{B}$ has the block form (6-10), and the matrix of the restriction operator T_W relative to the basis $\mathcal{B}'$ is

$$B = \begin{bmatrix} t_1 & 0 & \cdots & 0 \\ 0 & t_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & t_r \end{bmatrix}.$$

The characteristic polynomial of B (i.e., of T_W) is

$$g = (x - c_1)^{e_1} \cdots (x - c_k)^{e_k}$$

where $e_i = \dim W_i$. Furthermore, g divides f, the characteristic polynomial for T. Therefore, the multiplicity of c_i as a root of f is at least $\dim W_i$.

All of this should make Theorem 2 transparent. It merely says that T is diagonalizable if and only if $r = n$, if and only if $e_1 + \cdots + e_k = n$. It does not help us too much with the non-diagonalizable case, since we don't know the matrices C and D of (6-10).

Definition. *Let* W *be an invariant subspace for* T *and let* α *be a vector in* V. *The* **T-conductor of α into** W *is the set* $S_T(\alpha; W)$, *which consists of all polynomials* g *(over the scalar field) such that* $g(T)\alpha$ *is in* W.

Since the operator T will be fixed throughout most discussions, we shall usually drop the subscript T and write $S(\alpha; W)$. The authors usually call that collection of polynomials the 'stuffer' (*das einstopfende Ideal*). 'Conductor' is the more standard term, preferred by those who envision a less aggressive operator $g(T)$, gently leading the vector α into W. In the special case $W = \{0\}$ the conductor is called the ***T*-annihilator of α.**

Lemma. *If* W *is an invariant subspace for* T, *then* W *is invariant under every polynomial in* T. *Thus, for each* α *in* V, *the conductor* S(α; W) *is an ideal in the polynomial algebra* F[x].

Proof. If β is in W, then $T\beta$ is in W. Consequently, $T(T\beta) = T^2\beta$ is in W. By induction, $T^k\beta$ is in W for each k. Take linear combinations to see that $f(T)\beta$ is in W for every polynomial f.

The definition of $S(\alpha; W)$ makes sense if W is any subset of V. If W is a subspace, then $S(\alpha; W)$ is a subspace of $F[x]$, because

$$(cf + g)(T) = cf(T) + g(T).$$

If W is also invariant under T, let g be a polynomial in $S(\alpha; W)$, i.e., let $g(T)\alpha$ be in W. If f is any polynomial, then $f(T)[g(T)\alpha]$ will be in W. Since

$$(fg)(T) = f(T)g(T)$$

fg is in $S(\alpha; W)$. Thus the conductor absorbs multiplication by any polynomial. ∎

The unique monic generator of the ideal $S(\alpha; W)$ is also called the ***T*-conductor of α into *W*** (the ***T*-annihilator** in case $W = \{0\}$). The T-conductor of α into W is the monic polynomial g of least degree such that $g(T)\alpha$ is in W. A polynomial f is in $S(\alpha; W)$ if and only if g divides f. Note that the conductor $S(\alpha; W)$ always contains the minimal polynomial for T; hence, *every* T*-conductor divides the minimal polynomial for* T.

As the first illustration of how to use the conductor $S(\alpha; W)$, we shall characterize triangulable operators. The linear operator T is called **triangulable** if there is an ordered basis in which T is represented by a triangular matrix.

Lemma. *Let* V *be a finite-dimensional vector space over the field* F. *Let* T *be a linear operator on* V *such that the minimal polynomial for* T *is a product of linear factors*

$$p = (x - c_1)^{r_1} \cdots (x - c_k)^{r_k}, \qquad c_i \textit{ in } F.$$

Let W *be a proper* (W $\neq$ V) *subspace of* V *which is invariant under* T. *There exists a vector* α *in* V *such that*

(a) α *is not in* W;

(b) $(T - cI)\alpha$ *is in* W, *for some characteristic value* c *of the operator* T.

Proof. What (a) and (b) say is that the T-conductor of α into W is a linear polynomial. Let β be any vector in V which is not in W. Let g be the T-conductor of β into W. Then g divides p, the minimal polynomial for T. Since β is not in W, the polynomial g is not constant. Therefore,

$$g = (x - c_1)^{e_1} \cdots (x - c_k)^{e_k}$$

where at least one of the integers e_i is positive. Choose j so that $e_j > 0$. Then $(x - c_j)$ divides g:

$$g = (x - c_j)h.$$

By the definition of g, the vector $\alpha = h(T)\beta$ cannot be in W. But

$$\begin{aligned}(T - c_jI)\alpha &= (T - c_jI)h(T)\beta \\ &= g(T)\beta\end{aligned}$$

is in W. ▌

Theorem 5. *Let V be a finite-dimensional vector space over the field F and let T be a linear operator on V. Then T is triangulable if and only if the minimal polynomial for T is a product of linear polynomials over F.*

Proof. Suppose that the minimal polynomial factors

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

By repeated application of the lemma above, we shall arrive at an ordered basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ in which the matrix representing T is upper-triangular:

$$[T]_{\mathcal{B}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix}. \tag{6-11}$$

Now (6-11) merely says that

$$T\alpha_j = a_{1j}\alpha_1 + \cdots + a_{jj}\alpha_j, \qquad 1 \leq j \leq n \tag{6-12}$$

that is, $T\alpha_j$ is in the subspace spanned by $\alpha_1, \ldots, \alpha_j$. To find $\alpha_1, \ldots, \alpha_n$, we start by applying the lemma to the subspace $W = \{0\}$, to obtain the vector α_1. Then apply the lemma to W_1, the space spanned by α_1, and we obtain α_2. Next apply the lemma to W_2, the space spanned by α_1 and α_2. Continue in that way. One point deserves comment. After $\alpha_1, \ldots, \alpha_i$ have been found, it is the triangular-type relations (6-12) for $j = 1, \ldots, i$ which ensure that the subspace spanned by $\alpha_1, \ldots, \alpha_i$ is invariant under T.

If T is triangulable, it is evident that the characteristic polynomial for T has the form

$$f = (x - c_1)^{d_1} \cdots (x - c_k)^{d_k}, \qquad c_i \text{ in } F.$$

Just look at the triangular matrix (6-11). The diagonal entries $a_{11}, \ldots, a_{1n}$ are the characteristic values, with c_i repeated d_i times. But, if f can be so factored, so can the minimal polynomial p, because it divides f. ▌

Corollary. *Let F be an algebraically closed field, e.g., the complex number field. Every n × n matrix over F is similar over F to a triangular matrix.*

Theorem 6. *Let V be a finite-dimensional vector space over the field F and let T be a linear operator on V. Then T is diagonalizable if and only if the minimal polynomial for T has the form*

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

where $c_1, \ldots, c_k$ *are distinct elements of* F.

Proof. We have noted earlier that, if T is diagonalizable, its minimal polynomial is a product of distinct linear factors (see the discussion prior to Example 4). To prove the converse, let W be the subspace spanned by all of the characteristic vectors of T, and suppose $W \neq V$. By the lemma used in the proof of Theorem 5, there is a vector α not in W and a characteristic value c_j of T such that the vector

$$\beta = (T - c_jI)\alpha$$

lies in W. Since β is in W,

$$\beta = \beta_1 + \cdots + \beta_k$$

where $T\beta_i = c_i\beta_i$, $1 \leq i \leq k$, and therefore the vector

$$h(T)\beta = h(c_1)\beta_1 + \cdots + h(c_k)\beta_k$$

is in W, for every polynomial h.

Now $p = (x - c_j)q$, for some polynomial q. Also

$$q - q(c_j) = (x - c_j)h.$$

We have

$$q(T)\alpha - q(c_j)\alpha = h(T)(T - c_jI)\alpha = h(T)\beta.$$

But $h(T)\beta$ is in W and, since

$$0 = p(T)\alpha = (T - c_jI)q(T)\alpha$$

the vector $q(T)\alpha$ is in W. Therefore, $q(c_j)\alpha$ is in W. Since α is not in W, we have $q(c_j) = 0$. That contradicts the fact that p has distinct roots. ∎

At the end of Section 6.7, we shall give a different proof of Theorem 6. In addition to being an elegant result, Theorem 6 is useful in a computational way. Suppose we have a linear operator T, represented by the matrix A in some ordered basis, and we wish to know if T is diagonalizable. We compute the characteristic polynomial f. *If* we can factor f:

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

we have two different methods for determining whether or not T is diagonalizable. One method is to see whether (for each i) we can find d_i independent characteristic vectors associated with the characteristic value c_i. The other method is to check whether or not $(T - c_1I) \cdots (T - c_kI)$ is the zero operator.

Theorem 5 provides a different proof of the Cayley-Hamilton theorem. That theorem is easy for a triangular matrix. Hence, via Theorem 5, we obtain the result for any matrix over an algebraically closed field. Any field is a subfield of an algebraically closed field. If one knows that result, one obtains a proof of the Cayley-Hamilton theorem for matrices over any field. If we at least admit into our discussion the Fundamental Theorem of Algebra (the complex number field is algebraically closed), then Theorem 5 provides a proof of the Cayley-Hamilton theorem for complex matrices, and that proof is independent of the one which we gave earlier.

Exercises

1. Let T be the linear operator on R^2, the matrix of which in the standard ordered basis is

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

(a) Prove that the only subspaces of R^2 invariant under T are R^2 and the zero subspace.

(b) If U is the linear operator on C^2, the matrix of which in the standard ordered basis is A, show that U has 1-dimensional invariant subspaces.

2. Let W be an invariant subspace for T. Prove that the minimal polynomial for the restriction operator T_W divides the minimal polynomial for T, without referring to matrices.

3. Let c be a characteristic value of T and let W be the space of characteristic vectors associated with the characteristic value c. What is the restriction operator T_W?

4. Let

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

Is A similar over the field of real numbers to a triangular matrix? If so, find such a triangular matrix.

5. Every matrix A such that $A^2 = A$ is similar to a diagonal matrix.

6. Let T be a diagonalizable linear operator on the n-dimensional vector space V, and let W be a subspace which is invariant under T. Prove that the restriction operator T_W is diagonalizable.

7. Let T be a linear operator on a finite-dimensional vector space over the field of complex numbers. Prove that T is diagonalizable if and only if T is annihilated by some polynomial over C which has distinct roots.

8. Let T be a linear operator on V. If every subspace of V is invariant under T, then T is a scalar multiple of the identity operator.

9. Let T be the indefinite integral operator

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