on the space of continuous functions on the interval $[0, 1]$. Is the space of polynomial functions invariant under $T$? The space of differentiable functions? The space of functions which vanish at $x = \frac{1}{2}$?

**10.** Let $A$ be a $3 \times 3$ matrix with real entries. Prove that, if $A$ is not similar over $R$ to a triangular matrix, then $A$ is similar over $C$ to a diagonal matrix.

**11.** True or false? If the triangular matrix $A$ is similar to a diagonal matrix, then $A$ is already diagonal.

**12.** Let $T$ be a linear operator on a finite-dimensional vector space over an algebraically closed field $F$. Let $f$ be a polynomial over $F$. Prove that $c$ is a characteristic value of $f(T)$ if and only if $c = f(t)$, where $t$ is a characteristic value of $T$.

**13.** Let $V$ be the space of $n \times n$ matrices over $F$. Let $A$ be a fixed $n \times n$ matrix over $F$. Let $T$ and $U$ be the linear operators on $V$ defined by

$$\begin{aligned} T(B) &= AB \\ U(B) &= AB - BA. \end{aligned}$$

(a) True or false? If $A$ is diagonalizable (over $F$), then $T$ is diagonalizable.
(b) True or false? If $A$ is diagonalizable, then $U$ is diagonalizable.

## 6.5. Simultaneous Triangulation; Simultaneous Diagonalization

Let $V$ be a finite-dimensional space and let $\mathfrak{F}$ be a family of linear operators on $V$. We ask when we can simultaneously triangulate or diagonalize the operators in $\mathfrak{F}$, i.e., find one basis $\mathfrak{B}$ such that all of the matrices $[T]_{\mathfrak{B}}$, $T$ in $\mathfrak{F}$, are triangular (or diagonal). In the case of diagonalization, it is necessary that $\mathfrak{F}$ be a commuting family of operators: $UT = TU$ for all $T$, $U$ in $\mathfrak{F}$. That follows from the fact that all diagonal matrices commute. Of course, it is also necessary that each operator in $\mathfrak{F}$ be a diagonalizable operator. In order to simultaneously triangulate, each operator in $\mathfrak{F}$ must be triangulable. It is not necessary that $\mathfrak{F}$ be a commuting family; however, that condition is sufficient for simultaneous triangulation (if each $T$ can be individually triangulated). These results follow from minor variations of the proofs of Theorems 5 and 6.

The subspace $W$ is **invariant under** (the family of operators) $\mathfrak{F}$ if $W$ is invariant under each operator in $\mathfrak{F}$.

***Lemma.*** *Let $\mathfrak{F}$ be a commuting family of triangulable linear operators on V. Let W be a proper subspace of V which is invariant under $\mathfrak{F}$. There exists a vector $\alpha$ in V such that*

(a) *$\alpha$ is not in W;*
(b) *for each T in $\mathfrak{F}$, the vector $T\alpha$ is in the subspace spanned by $\alpha$ and W.*

*Proof.* It is no loss of generality to assume that $\mathfrak{F}$ contains only a finite number of operators, because of this observation. Let $\{T_1, \ldots, T_r\}$

be a maximal linearly independent subset of $\mathcal{F}$, i.e., a basis for the subspace spanned by $\mathcal{F}$. If $\alpha$ is a vector such that (b) holds for each $T_i$, then (b) will hold for every operator which is a linear combination of $T_1, \ldots, T_r$.

By the lemma before Theorem 5 (this lemma for a single operator), we can find a vector $\beta_1$ (not in $W$) and a scalar $c_1$ such that $(T_1 - c_1I)\beta_1$ is in $W$. Let $V_1$ be the collection of all vectors $\beta$ in $V$ such that $(T_1 - c_1I)\beta$ is in $W$. Then $V_1$ is a subspace of $V$ which is properly larger than $W$. Furthermore, $V_1$ is invariant under $\mathcal{F}$, for this reason. If $T$ commutes with $T_1$, then

$$(T_1 - c_1I)(T\beta) = T(T_1 - c_1I)\beta.$$

If $\beta$ is in $V_1$, then $(T_1 - c_1I)\beta$ is in $W$. Since $W$ is invariant under each $T$ in $\mathcal{F}$, we have $T(T_1 - c_1I)\beta$ in $W$, i.e., $T\beta$ in $V_1$, for all $\beta$ in $V_1$ and all $T$ in $\mathcal{F}$.

Now $W$ is a proper subspace of $V_1$. Let $U_2$ be the linear operator on $V_1$ obtained by restricting $T_2$ to the subspace $V_1$. The minimal polynomial for $U_2$ divides the minimal polynomial for $T_2$. Therefore, we may apply the lemma before Theorem 5 to that operator and the invariant subspace $W$. We obtain a vector $\beta_2$ in $V_1$ (not in $W$) and a scalar $c_2$ such that $(T_2 - c_2I)\beta_2$ is in $W$. Note that

(a) $\beta_2$ is not in $W$;
(b) $(T_1 - c_1I)\beta_2$ is in $W$;
(c) $(T_2 - c_2I)\beta_2$ is in $W$.

Let $V_2$ be the set of all vectors $\beta$ in $V_1$ such that $(T_2 - c_2I)\beta$ is in $W$. Then $V_2$ is invariant under $\mathcal{F}$. Apply the lemma before Theorem 5 to $U_3$, the restriction of $T_3$ to $V_2$. If we continue in this way, we shall reach a vector $\alpha = \beta_r$ (not in $W$) such that $(T_j - c_jI)\alpha$ is in $W$, $j = 1, \ldots, r$. ∎

***Theorem 7.*** *Let $V$ be a finite-dimensional vector space over the field $F$. Let $\mathcal{F}$ be a commuting family of triangulable linear operators on $V$. There exists an ordered basis for $V$ such that every operator in $\mathcal{F}$ is represented by a triangular matrix in that basis.*

*Proof.* Given the lemma which we just proved, this theorem has the same proof as does Theorem 5, if one replaces $T$ by $\mathcal{F}$. ∎

***Corollary.*** *Let $\mathcal{F}$ be a commuting family of $n \times n$ matrices over an algebraically closed field $F$. There exists a non-singular $n \times n$ matrix $P$ with entries in $F$ such that $P^{-1}AP$ is upper-triangular, for every matrix $A$ in $\mathcal{F}$.*

***Theorem 8.*** *Let $\mathcal{F}$ be a commuting family of diagonalizable linear operators on the finite-dimensional vector space $V$. There exists an ordered basis for $V$ such that every operator in $\mathcal{F}$ is represented in that basis by a diagonal matrix.*

*Proof.* We could prove this theorem by adapting the lemma before Theorem 7 to the diagonalizable case, just as we adapted the lemma before Theorem 5 to the diagonalizable case in order to prove Theorem 6. However, at this point it is easier to proceed by induction on the dimension of $V$.

If $\dim V = 1$, there is nothing to prove. Assume the theorem for vector spaces of dimension less than $n$, and let $V$ be an $n$-dimensional space. Choose any $T$ in $\mathfrak{F}$ which is not a scalar multiple of the identity. Let $c_1, \ldots, c_k$ be the distinct characteristic values of $T$, and (for each $i$) let $W_i$ be the null space of $T - c_iI$. Fix an index $i$. Then $W_i$ is invariant under every operator which commutes with $T$. Let $\mathfrak{F}_i$ be the family of linear operators on $W_i$ obtained by restricting the operators in $\mathfrak{F}$ to the (invariant) subspace $W_i$. Each operator in $\mathfrak{F}_i$ is diagonalizable, because its minimal polynomial divides the minimal polynomial for the corresponding operator in $\mathfrak{F}$. Since $\dim W_i < \dim V$, the operators in $\mathfrak{F}_i$ can be simultaneously diagonalized. In other words, $W_i$ has a basis $\mathcal{B}_i$ which consists of vectors which are simultaneously characteristic vectors for every operator in $\mathfrak{F}_i$.

Since $T$ is diagonalizable, the lemma before Theorem 2 tells us that $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_k)$ is a basis for $V$. That is the basis we seek. ▌

## *Exercises*

**1.** Find an invertible real matrix $P$ such that $P^{-1}AP$ and $P^{-1}BP$ are both diagonal, where $A$ and $B$ are the real matrices

(a) $$A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & -8 \\ 0 & -1 \end{bmatrix}$$

(b) $$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}.$$

**2.** Let $\mathfrak{F}$ be a commuting family of $3 \times 3$ complex matrices. How many linearly independent matrices can $\mathfrak{F}$ contain? What about the $n \times n$ case?

**3.** Let $T$ be a linear operator on an $n$-dimensional space, and suppose that $T$ has $n$ distinct characteristic values. Prove that any linear operator which commutes with $T$ is a polynomial in $T$.

**4.** Let $A$, $B$, $C$, and $D$ be $n \times n$ complex matrices which commute. Let $E$ be the $2n \times 2n$ matrix

$$E = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

Prove that $\det E = \det (AD - BC)$.

**5.** Let $F$ be a field, $n$ a positive integer, and let $V$ be the space of $n \times n$ matrices over $F$. If $A$ is a fixed $n \times n$ matrix over $F$, let $T_A$ be the linear operator on $V$ defined by $T_A(B) = AB - BA$. Consider the family of linear operators $T_A$ obtained by letting $A$ vary over all diagonal matrices. Prove that the operators in that family are simultaneously diagonalizable.

## 6.6. Direct-Sum Decompositions

As we continue with our analysis of a single linear operator, we shall formulate our ideas in a slightly more sophisticated way—less in terms of matrices and more in terms of subspaces. When we began this chapter, we described our goal this way: To find an ordered basis in which the matrix of $T$ assumes an especially simple form. Now, we shall describe our goal as follows: To decompose the underlying space $V$ into a sum of invariant subspaces for $T$ such that the restriction operators on those subspaces are simple.

**Definition.** *Let* $W_1, \ldots, W_k$ *be subspaces of the vector space* V. *We say that* $W_1, \ldots, W_k$ *are* **independent** *if*

$$\alpha_1 + \cdots + \alpha_k = 0, \qquad \alpha_i \text{ in } W_i$$

*implies that each* $\alpha_i$ *is* 0.

For $k = 2$, the meaning of independence is $\{0\}$ intersection, i.e., $W_1$ and $W_2$ are independent if and only if $W_1 \cap W_2 = \{0\}$. If $k > 2$, the independence of $W_1, \ldots, W_k$ says much more than $W_1 \cap \cdots \cap W_k = \{0\}$. It says that each $W_j$ intersects the sum of the other subspaces $W_i$ only in the zero vector.

The significance of independence is this. Let $W = W_1 + \cdots + W_k$ be the subspace spanned by $W_1, \ldots, W_k$. Each vector $\alpha$ in $W$ can be expressed as a sum

$$\alpha = \alpha_1 + \cdots + \alpha_k, \qquad \alpha_i \text{ in } W_i.$$

*If* $W_1, \ldots, W_k$ are independent, then that expression for $\alpha$ is unique; for if

$$\alpha = \beta_1 + \cdots + \beta_k, \qquad \beta_i \text{ in } W_i$$

then $0 = (\alpha_1 - \beta_1) + \cdots + (\alpha_k - \beta_k)$, hence $\alpha_i - \beta_i = 0$, $i = 1, \ldots, k$. Thus, when $W_1, \ldots, W_k$ are independent, we can operate with the vectors in $W$ as $k$-tuples $(\alpha_1, \ldots, \alpha_k)$, $\alpha_i$ in $W_i$, in the same way as we operate with vectors in $R^k$ as $k$-tuples of numbers.

**Lemma.** *Let* V *be a finite-dimensional vector space. Let* $W_1, \ldots, W_k$ *be subspaces of* V *and let* $W = W_1 + \cdots + W_k$. *The following are equivalent.*

(a) $W_1, \ldots, W_k$ *are independent.*

(b) *For each* $j$, $2 \leq j \leq k$, *we have*

$$W_j \cap (W_1 + \cdots + W_{j-1}) = \{0\}.$$

(c) *If* $\mathcal{B}_i$ *is an ordered basis for* $W_i$, $1 \leq i \leq k$, *then the sequence* $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_k)$ *is an ordered basis for* W.

*Proof.* Assume (a). Let $\alpha$ be a vector in the intersection $W_j \cap (W_1 + \cdots + W_{j-1})$. Then there are vectors $\alpha_1, \ldots, \alpha_{j-1}$ with $\alpha_i$ in $W_i$ such that $\alpha = \alpha_1 + \cdots + \alpha_{j-1}$. Since

$$\alpha_1 + \cdots + \alpha_{j-1} + (-\alpha) + 0 + \cdots + 0 = 0$$

and since $W_1, \ldots, W_k$ are independent, it must be that $\alpha_1 = \alpha_2 = \cdots = \alpha_{j-1} = \alpha = 0$.

Now, let us observe that (b) implies (a). Suppose

$$0 = \alpha_1 + \cdots + \alpha_k, \qquad \alpha_i \text{ in } W_i.$$

Let $j$ be the largest integer $i$ such that $\alpha_i \neq 0$. Then

$$0 = \alpha_1 + \cdots + \alpha_j, \qquad \alpha_j \neq 0.$$

Thus $\alpha_j = -\alpha_1 - \cdots - \alpha_{j-1}$ is a non-zero vector in $W_j \cap (W_1 + \cdots + W_{j-1})$.

Now that we know (a) and (b) are the same, let us see why (a) is equivalent to (c). Assume (a). Let $\mathcal{B}_i$ be a basis for $W_i$, $1 \leq i \leq k$, and let $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_k)$. Any linear relation between the vectors in $\mathcal{B}$ will have the form

$$\beta_1 + \cdots + \beta_k = 0$$

where $\beta_i$ is some linear combination of the vectors in $\mathcal{B}_i$. Since $W_1, \ldots, W_k$ are independent, each $\beta_i$ is 0. Since each $\mathcal{B}_i$ is independent, the relation we have between the vectors in $\mathcal{B}$ is the trivial relation.

We relegate the proof that (c) implies (a) to the exercises (Exercise 2). ∎

If any (and hence all) of the conditions of the last lemma hold, we say that the sum $W = W_1 + \cdots + W_k$ is **direct** or that $W$ is the **direct sum** of $W_1, \ldots, W_k$ and we write

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

In the literature, the reader may find this direct sum referred to as an independent sum or the interior direct sum of $W_1, \ldots, W_k$.

EXAMPLE 11. Let $V$ be a finite-dimensional vector space over the field $F$ and let $\{\alpha_1, \ldots, \alpha_n\}$ be any basis for $V$. If $W_i$ is the one-dimensional subspace spanned by $\alpha_i$, then $V = W_1 \oplus \cdots \oplus W_n$.

EXAMPLE 12. Let $n$ be a positive integer and $F$ a subfield of the complex numbers, and let $V$ be the space of all $n \times n$ matrices over $F$. Let $W_1$ be the subspace of all **symmetric** matrices, i.e., matrices $A$ such that $A^t = A$. Let $W_2$ be the subspace of all **skew-symmetric** matrices, i.e., matrices $A$ such that $A^t = -A$. Then $V = W_1 \oplus W_2$. If $A$ is any matrix in $V$, the unique expression for $A$ as a sum of matrices, one in $W_1$ and the other in $W_2$, is