

Examples 3 and 4 suggest that an operator whose characteristic polynomial splits is diagonalizable if and only if the dimension of each eigenspace is equal to the multiplicity of the corresponding eigenvalue. This is indeed true, as we now show. We begin with the following lemma, which is a slight variation of Theorem 5.5.

**Lemma.** *Let  $T$  be a linear operator, and let  $\lambda_1, \lambda_2, \dots, \lambda_k$  be distinct eigenvalues of  $T$ . For each  $i = 1, 2, \dots, k$ , let  $v_i \in E_{\lambda_i}$ , the eigenspace corresponding to  $\lambda_i$ . If*

$$v_1 + v_2 + \cdots + v_k = 0,$$

*then  $v_i = 0$  for all  $i$ .*

*Proof.* Suppose otherwise. By renumbering if necessary, suppose that, for  $1 \leq m \leq k$ , we have  $v_i \neq 0$  for  $1 \leq i \leq m$ , and  $v_i = 0$  for  $i > m$ . Then, for each  $i \leq m$ ,  $v_i$  is an eigenvector of  $T$  corresponding to  $\lambda_i$  and

$$v_1 + v_2 + \cdots + v_m = 0.$$

But this contradicts Theorem 5.5, which states that these  $v_i$ 's are linearly independent. We conclude, therefore, that  $v_i = 0$  for all  $i$ . ■

**Theorem 5.8.** *Let  $T$  be a linear operator on a vector space  $V$ , and let  $\lambda_1, \lambda_2, \dots, \lambda_k$  be distinct eigenvalues of  $T$ . For each  $i = 1, 2, \dots, k$ , let  $S_i$  be a finite linearly independent subset of the eigenspace  $E_{\lambda_i}$ . Then  $S = S_1 \cup S_2 \cup \cdots \cup S_k$  is a linearly independent subset of  $V$ .*

*Proof.* Suppose that for each  $i$

$$S_i = \{v_{i1}, v_{i2}, \dots, v_{in_i}\}.$$

Then  $S = \{v_{ij} : 1 \leq j \leq n_i, \text{ and } 1 \leq i \leq k\}$ . Consider any scalars  $\{a_{ij}\}$  such that

$$\sum_{i=1}^k \sum_{j=1}^{n_i} a_{ij} v_{ij} = 0.$$

For each  $i$ , let

$$w_i = \sum_{j=1}^{n_i} a_{ij} v_{ij}.$$

Then  $w_i \in E_{\lambda_i}$  for each  $i$ , and  $w_1 + \cdots + w_k = 0$ . Therefore, by the lemma,  $w_i = 0$  for all  $i$ . But each  $S_i$  is linearly independent, and hence  $a_{ij} = 0$  for all  $j$ . We conclude that  $S$  is linearly independent. ■

Theorem 5.8 tells us how to construct a linearly independent subset of eigenvectors, namely, by collecting bases for the individual eigenspaces. The next theorem tells us when the resulting set is a basis for the entire space.

**Theorem 5.9.** *Let  $T$  be a linear operator on a finite-dimensional vector space  $V$  such that the characteristic polynomial of  $T$  splits. Let  $\lambda_1, \lambda_2, \dots, \lambda_k$  be the distinct eigenvalues of  $T$ . Then*

- (a)  *$T$  is diagonalizable if and only if the multiplicity of  $\lambda_i$  is equal to  $\dim(E_{\lambda_i})$  for all  $i$ .*
- (b) *If  $T$  is diagonalizable and  $\beta_i$  is an ordered basis for  $E_{\lambda_i}$  for each  $i$ , then  $\beta = \beta_1 \cup \beta_2 \cup \dots \cup \beta_k$  is an ordered basis<sup>2</sup> for  $V$  consisting of eigenvectors of  $T$ .*

*Proof.* For each  $i$ , let  $m_i$  denote the multiplicity of  $\lambda_i$ ,  $d_i = \dim(E_{\lambda_i})$ , and  $n = \dim(V)$ .

First, suppose that  $T$  is diagonalizable. Let  $\beta$  be a basis for  $V$  consisting of eigenvectors of  $T$ . For each  $i$ , let  $\beta_i = \beta \cap E_{\lambda_i}$ , the set of vectors in  $\beta$  that are eigenvectors corresponding to  $\lambda_i$ , and let  $n_i$  denote the number of vectors in  $\beta_i$ . Then  $n_i \leq d_i$  for each  $i$  because  $\beta_i$  is a linearly independent subset of a subspace of dimension  $d_i$ , and  $d_i \leq m_i$  by Theorem 5.7. The  $n_i$ 's sum to  $n$  because  $\beta$  contains  $n$  vectors. The  $m_i$ 's also sum to  $n$  because the degree of the characteristic polynomial of  $T$  is equal to the sum of the multiplicities of the eigenvalues. Thus

$$n = \sum_{i=1}^k n_i \leq \sum_{i=1}^k d_i \leq \sum_{i=1}^k m_i = n.$$

It follows that

$$\sum_{i=1}^k (m_i - d_i) = 0.$$

Since  $(m_i - d_i) \geq 0$  for all  $i$ , we conclude that  $m_i = d_i$  for all  $i$ .

Conversely, suppose that  $m_i = d_i$  for all  $i$ . We simultaneously show that  $T$  is diagonalizable and prove (b). For each  $i$ , let  $\beta_i$  be an ordered basis for  $E_{\lambda_i}$ , and let  $\beta = \beta_1 \cup \beta_2 \cup \dots \cup \beta_k$ . By Theorem 5.8,  $\beta$  is linearly independent. Furthermore, since  $d_i = m_i$  for all  $i$ ,  $\beta$  contains

$$\sum_{i=1}^k d_i = \sum_{i=1}^k m_i = n$$

---

<sup>2</sup>We regard  $\beta_1 \cup \beta_2 \cup \dots \cup \beta_k$  as an ordered basis in the natural way—the vectors in  $\beta_1$  are listed first (in the same order as in  $\beta_1$ ), then the vectors in  $\beta_2$  (in the same order as in  $\beta_2$ ), etc.

vectors. Therefore  $\beta$  is an ordered basis for  $V$  consisting of eigenvectors of  $T$ , and we conclude that  $T$  is diagonalizable. ■

This theorem completes our study of the diagonalization problem. We summarize our results.

### Test for Diagonalization

Let  $T$  be a linear operator on an  $n$ -dimensional vector space  $V$ . Then  $T$  is diagonalizable if and only if both of the following conditions hold.

1. The characteristic polynomial of  $T$  splits.
2. For each eigenvalue  $\lambda$  of  $T$ , the multiplicity of  $\lambda$  equals  $n - \text{rank}(T - \lambda I)$ .

These same conditions can be used to test if a square matrix  $A$  is diagonalizable because diagonalizability of  $A$  is equivalent to diagonalizability of the operator  $L_A$ .

If  $T$  is a diagonalizable operator and  $\beta_1, \beta_2, \dots, \beta_k$  are ordered bases for the eigenspaces of  $T$ , then the union  $\beta = \beta_1 \cup \beta_2 \cup \dots \cup \beta_k$  is an ordered basis for  $V$  consisting of eigenvectors of  $T$ , and hence  $[T]_\beta$  is a diagonal matrix.

When testing  $T$  for diagonalizability, it is usually easiest to choose a convenient basis  $\alpha$  for  $V$  and work with  $B = [T]_\alpha$ . If the characteristic polynomial of  $B$  splits, then use condition 2 above to check if the multiplicity of each *repeated* eigenvalue of  $B$  equals  $n - \text{rank}(B - \lambda I)$ . (By Theorem 5.7, condition 2 is automatically satisfied for eigenvalues of multiplicity 1.) If so, then  $B$ , and hence  $T$ , is diagonalizable.

If  $T$  is diagonalizable and a basis  $\beta$  for  $V$  consisting of eigenvectors of  $T$  is desired, then we first find a basis for each eigenspace of  $B$ . The union of these bases is a basis  $\gamma$  for  $F^n$  consisting of eigenvectors of  $B$ . Each vector in  $\gamma$  is the coordinate vector relative to  $\alpha$  of an eigenvector of  $T$ . The set consisting of these  $n$  eigenvectors of  $T$  is the desired basis  $\beta$ .

Furthermore, if  $A$  is an  $n \times n$  diagonalizable matrix, we can use the corollary to Theorem 2.23 (p. 115) to find an invertible  $n \times n$  matrix  $Q$  and a diagonal  $n \times n$  matrix  $D$  such that  $Q^{-1}AQ = D$ . The matrix  $Q$  has as its columns the vectors in a basis of eigenvectors of  $A$ , and  $D$  has as its  $j$ th diagonal entry the eigenvalue of  $A$  corresponding to the  $j$ th column of  $Q$ .

We now consider some examples illustrating the preceding ideas.

### Example 5

We test the matrix

$$A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix} \in M_{3 \times 3}(R)$$

for diagonalizability.

The characteristic polynomial of  $A$  is  $\det(A - tI) = -(t-4)(t-3)^2$ , which splits, and so condition 1 of the test for diagonalization is satisfied. Also  $A$  has eigenvalues  $\lambda_1 = 4$  and  $\lambda_2 = 3$  with multiplicities 1 and 2, respectively. Since  $\lambda_1$  has multiplicity 1, condition 2 is satisfied for  $\lambda_1$ . Thus we need only test condition 2 for  $\lambda_2$ . Because

$$A - \lambda_2 I = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

has rank 2, we see that  $3 - \text{rank}(A - \lambda_2 I) = 1$ , which is not the multiplicity of  $\lambda_2$ . Thus condition 2 fails for  $\lambda_2$ , and  $A$  is therefore not diagonalizable.  $\blacklozenge$

### Example 6

Let  $T$  be the linear operator on  $P_2(R)$  defined by

$$T(f(x)) = f(1) + f'(0)x + (f'(0) + f''(0))x^2.$$

We first test  $T$  for diagonalizability. Let  $\alpha$  denote the standard ordered basis for  $P_2(R)$  and  $B = [T]_\alpha$ . Then

$$B = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}.$$

The characteristic polynomial of  $B$ , and hence of  $T$ , is  $-(t-1)^2(t-2)$ , which splits. Hence condition 1 of the test for diagonalization is satisfied. Also  $B$  has the eigenvalues  $\lambda_1 = 1$  and  $\lambda_2 = 2$  with multiplicities 2 and 1, respectively. Condition 2 is satisfied for  $\lambda_2$  because it has multiplicity 1. So we need only verify condition 2 for  $\lambda_1 = 1$ . For this case,

$$3 - \text{rank}(B - \lambda_1 I) = 3 - \text{rank} \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} = 3 - 1 = 2,$$

which is equal to the multiplicity of  $\lambda_1$ . Therefore  $T$  is diagonalizable.

We now find an ordered basis  $\gamma$  for  $R^3$  of eigenvectors of  $B$ . We consider each eigenvalue separately.

The eigenspace corresponding to  $\lambda_1 = 1$  is

$$E_{\lambda_1} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in R^3 : \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0 \right\},$$

which is the solution space for the system

$$x_2 + x_3 = 0,$$

and has

$$\gamma_1 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\}$$

as a basis.

The eigenspace corresponding to  $\lambda_2 = 2$  is

$$E_{\lambda_2} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : \begin{pmatrix} -1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{0} \right\},$$

which is the solution space for the system

$$\begin{aligned} -x_1 + x_2 + x_3 &= 0 \\ x_2 &= 0, \end{aligned}$$

and has

$$\gamma_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

as a basis.

Let

$$\gamma = \gamma_1 \cup \gamma_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Then  $\gamma$  is an ordered basis for  $\mathbb{R}^3$  consisting of eigenvectors of  $B$ .

Finally, observe that the vectors in  $\gamma$  are the coordinate vectors relative to  $\alpha$  of the vectors in the set

$$\beta = \{1, -x + x^2, 1 + x^2\},$$

which is an ordered basis for  $P_2(R)$  consisting of eigenvectors of  $T$ . Thus

$$[T]_{\beta} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}. \quad \blacklozenge$$

Our next example is an application of diagonalization that is of interest in Section 5.3.

### Example 7

Let

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

We show that  $A$  is diagonalizable and find a  $2 \times 2$  matrix  $Q$  such that  $Q^{-1}AQ$  is a diagonal matrix. We then show how to use this result to compute  $A^n$  for any positive integer  $n$ .

First observe that the characteristic polynomial of  $A$  is  $(t-1)(t-2)$ , and hence  $A$  has two distinct eigenvalues,  $\lambda_1 = 1$  and  $\lambda_2 = 2$ . By applying the corollary to Theorem 5.5 to the operator  $L_A$ , we see that  $A$  is diagonalizable. Moreover,

$$\gamma_1 = \left\{ \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\} \quad \text{and} \quad \gamma_2 = \left\{ \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$$

are bases for the eigenspaces  $E_{\lambda_1}$  and  $E_{\lambda_2}$ , respectively. Therefore

$$\gamma = \gamma_1 \cup \gamma_2 = \left\{ \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$$

is an ordered basis for  $\mathbb{R}^2$  consisting of eigenvectors of  $\mathbb{R}^2$ . Let

$$Q = \begin{pmatrix} -2 & -1 \\ 1 & 1 \end{pmatrix},$$

the matrix whose columns are the vectors in  $\gamma$ . Then, by the corollary to Theorem 2.23 (p. 115),

$$D = Q^{-1}AQ = [L_A]_{\beta} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

To find  $A^n$  for any positive integer  $n$ , observe that  $A = QDQ^{-1}$ . Therefore

$$\begin{aligned} A^n &= (QDQ^{-1})^n \\ &= (QDQ^{-1})(QDQ^{-1}) \cdots (QDQ^{-1}) \\ &= QD^nQ^{-1} \\ &= Q \begin{pmatrix} 1^n & 0 \\ 0 & 2^n \end{pmatrix} Q^{-1} \\ &= \begin{pmatrix} -2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2^n \end{pmatrix} \begin{pmatrix} -1 & -1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 2-2^n & 2-2^{n+1} \\ -1+2^n & -1+2^{n+1} \end{pmatrix}. \quad \blacklozenge \end{aligned}$$

We now consider an application that uses diagonalization to solve a system of differential equations.

### Systems of Differential Equations

Consider the system of differential equations

$$\begin{aligned}x'_1 &= 3x_1 + x_2 + x_3 \\x'_2 &= 2x_1 + 4x_2 + 2x_3 \\x'_3 &= -x_1 - x_2 + x_3,\end{aligned}$$

where, for each  $i$ ,  $x_i = x_i(t)$  is a differentiable real-valued function of the real variable  $t$ . Clearly, this system has a solution, namely, the solution in which each  $x_i(t)$  is the zero function. We determine all of the solutions to this system.

Let  $x: R \rightarrow \mathbb{R}^3$  be the function defined by

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix}.$$

The **derivative** of  $x$ , denoted  $x'$ , is defined by

$$x'(t) = \begin{pmatrix} x'_1(t) \\ x'_2(t) \\ x'_3(t) \end{pmatrix}.$$

Let

$$A = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & -1 & 1 \end{pmatrix}$$

be the coefficient matrix of the given system, so that we can rewrite the system as the matrix equation  $x' = Ax$ .

It can be verified that for

$$Q = \begin{pmatrix} -1 & 0 & -1 \\ 0 & -1 & -2 \\ 1 & 1 & 1 \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix},$$

we have  $Q^{-1}AQ = D$ . Substitute  $A = QDQ^{-1}$  into  $x' = Ax$  to obtain  $x' = QDQ^{-1}x$  or, equivalently,  $Q^{-1}x' = DQ^{-1}x$ . The function  $y: R \rightarrow \mathbb{R}^3$  defined by  $y(t) = Q^{-1}x(t)$  can be shown to be differentiable, and  $y' = Q^{-1}x'$  (see Exercise 16). Hence the original system can be written as  $y' = Dy$ .

Since  $D$  is a diagonal matrix, the system  $y' = Dy$  is easy to solve. Setting

$$y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix},$$

we can rewrite  $y' = Dy$  as

$$\begin{pmatrix} y_1'(t) \\ y_2'(t) \\ y_3'(t) \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix} = \begin{pmatrix} 2y_1(t) \\ 2y_2(t) \\ 4y_3(t) \end{pmatrix}.$$

The three equations

$$\begin{aligned} y_1' &= 2y_1 \\ y_2' &= 2y_2 \\ y_3' &= 4y_3 \end{aligned}$$

are independent of each other, and thus can be solved individually. It is easily seen (as in Example 3 of Section 5.1) that the general solution to these equations is  $y_1(t) = c_1 e^{2t}$ ,  $y_2(t) = c_2 e^{2t}$ , and  $y_3(t) = c_3 e^{4t}$ , where  $c_1, c_2$ , and  $c_3$  are arbitrary constants. Finally,

$$\begin{aligned} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} &= x(t) = Qy(t) = \begin{pmatrix} -1 & 0 & -1 \\ 0 & -1 & -2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} c_1 e^{2t} \\ c_2 e^{2t} \\ c_3 e^{4t} \end{pmatrix} \\ &= \begin{pmatrix} -c_1 e^{2t} & - & c_3 e^{4t} \\ & -c_2 e^{2t} - 2c_3 e^{4t} & \\ c_1 e^{2t} + c_2 e^{2t} + & c_3 e^{4t} & \end{pmatrix} \end{aligned}$$

yields the general solution of the original system. Note that this solution can be written as

$$x(t) = e^{2t} \left[ c_1 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right] + e^{4t} \left[ c_3 \begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix} \right].$$

The expressions in brackets are arbitrary vectors in  $E_{\lambda_1}$  and  $E_{\lambda_2}$ , respectively, where  $\lambda_1 = 2$  and  $\lambda_2 = 4$ . Thus the general solution of the original system is  $x(t) = e^{2t} z_1 + e^{4t} z_2$ , where  $z_1 \in E_{\lambda_1}$  and  $z_2 \in E_{\lambda_2}$ . This result is generalized in Exercise 15.

### Direct Sums\*

Let  $T$  be a linear operator on a finite-dimensional vector space  $V$ . There is a way of decomposing  $V$  into simpler subspaces that offers insight into the

behavior of  $T$ . This approach is especially useful in Chapter 7, where we study nondiagonalizable linear operators. In the case of diagonalizable operators, the simpler subspaces are the eigenspaces of the operator.

**Definition.** Let  $W_1, W_2, \dots, W_k$  be subspaces of a vector space  $V$ . We define the **sum** of these subspaces to be the set

$$\{v_1 + v_2 + \cdots + v_k : v_i \in W_i \text{ for } 1 \leq i \leq k\},$$

which we denote by  $W_1 + W_2 + \cdots + W_k$  or  $\sum_{i=1}^k W_i$ .

It is a simple exercise to show that the sum of subspaces of a vector space is also a subspace.

### Example 8

Let  $V = \mathbb{R}^3$ , let  $W_1$  denote the  $xy$ -plane, and let  $W_2$  denote the  $yz$ -plane. Then  $\mathbb{R}^3 = W_1 + W_2$  because, for any vector  $(a, b, c) \in \mathbb{R}^3$ , we have

$$(a, b, c) = (a, 0, 0) + (0, b, c),$$

where  $(a, 0, 0) \in W_1$  and  $(0, b, c) \in W_2$ . ♦

Notice that in Example 8 the representation of  $(a, b, c)$  as a sum of vectors in  $W_1$  and  $W_2$  is not unique. For example,  $(a, b, c) = (a, b, 0) + (0, 0, c)$  is another representation. Because we are often interested in sums for which representations are unique, we introduce a condition that assures this outcome. The definition of *direct sum* that follows is a generalization of the definition given in the exercises of Section 1.3.

**Definition.** Let  $W_1, W_2, \dots, W_k$  be subspaces of a vector space  $V$ . We call  $V$  the **direct sum** of the subspaces  $W_1, W_2, \dots, W_k$  and write  $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$ , if

$$V = \sum_{i=1}^k W_i$$

and

$$W_j \cap \sum_{i \neq j} W_i = \{0\} \quad \text{for each } j \ (1 \leq j \leq k).$$

### Example 9

Let  $V = \mathbb{R}^4$ ,  $W_1 = \{(a, b, 0, 0) : a, b \in \mathbb{R}\}$ ,  $W_2 = \{(0, 0, c, 0) : c \in \mathbb{R}\}$ , and  $W_3 = \{(0, 0, 0, d) : d \in \mathbb{R}\}$ . For any  $(a, b, c, d) \in V$ ,

$$(a, b, c, d) = (a, b, 0, 0) + (0, 0, c, 0) + (0, 0, 0, d) \in W_1 + W_2 + W_3.$$

Thus

$$V = \sum_{i=1}^3 W_i.$$

To show that  $V$  is the direct sum of  $W_1$ ,  $W_2$ , and  $W_3$ , we must prove that  $W_1 \cap (W_2 + W_3) = W_2 \cap (W_1 + W_3) = W_3 \cap (W_1 + W_2) = \{0\}$ . But these equalities are obvious, and so  $V = W_1 \oplus W_2 \oplus W_3$ .  $\blacklozenge$

Our next result contains several conditions that are equivalent to the definition of a direct sum.

**Theorem 5.10.** *Let  $W_1, W_2, \dots, W_k$  be subspaces of a finite-dimensional vector space  $V$ . The following conditions are equivalent.*

- (a)  $V = W_1 \oplus W_2 \oplus \dots \oplus W_k$ .
- (b)  $V = \sum_{i=1}^k W_i$  and, for any vectors  $v_1, v_2, \dots, v_k$  such that  $v_i \in W_i$  ( $1 \leq i \leq k$ ), if  $v_1 + v_2 + \dots + v_k = 0$ , then  $v_i = 0$  for all  $i$ .
- (c) Each vector  $v \in V$  can be uniquely written as  $v = v_1 + v_2 + \dots + v_k$ , where  $v_i \in W_i$ .
- (d) If  $\gamma_i$  is an ordered basis for  $W_i$  ( $1 \leq i \leq k$ ), then  $\gamma_1 \cup \gamma_2 \cup \dots \cup \gamma_k$  is an ordered basis for  $V$ .
- (e) For each  $i = 1, 2, \dots, k$ , there exists an ordered basis  $\gamma_i$  for  $W_i$  such that  $\gamma_1 \cup \gamma_2 \cup \dots \cup \gamma_k$  is an ordered basis for  $V$ .

*Proof.* Assume (a). We prove (b). Clearly

$$V = \sum_{i=1}^k W_i.$$

Now suppose that  $v_1, v_2, \dots, v_k$  are vectors such that  $v_i \in W_i$  for all  $i$  and  $v_1 + v_2 + \dots + v_k = 0$ . Then for any  $j$

$$-v_j = \sum_{i \neq j} v_i \in \sum_{i \neq j} W_i.$$

But  $-v_j \in W_j$  and hence

$$-v_j \in W_j \cap \sum_{i \neq j} W_i = \{0\}.$$

So  $v_j = 0$ , proving (b).

Now assume (b). We prove (c). Let  $v \in V$ . By (b), there exist vectors  $v_1, v_2, \dots, v_k$  such that  $v_i \in W_i$  and  $v = v_1 + v_2 + \dots + v_k$ . We must show

that this representation is unique. Suppose also that  $v = w_1 + w_2 + \cdots + w_k$ , where  $w_i \in W_i$  for all  $i$ . Then

$$(v_1 - w_1) + (v_2 - w_2) + \cdots + (v_k - w_k) = 0.$$

But  $v_i - w_i \in W_i$  for all  $i$ , and therefore  $v_i - w_i = 0$  for all  $i$  by (b). Thus  $v_i = w_i$  for all  $i$ , proving the uniqueness of the representation.

Now assume (c). We prove (d). For each  $i$ , let  $\gamma_i$  be an ordered basis for  $W_i$ . Since

$$V = \sum_{i=1}^k W_i$$

by (c), it follows that  $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$  generates  $V$ . To show that this set is linearly independent, consider vectors  $v_{ij} \in \gamma_i$  ( $j = 1, 2, \dots, m_i$  and  $i = 1, 2, \dots, k$ ) and scalars  $a_{ij}$  such that

$$\sum_{i,j} a_{ij} v_{ij} = 0.$$

For each  $i$ , set

$$w_i = \sum_{j=1}^{m_i} a_{ij} v_{ij}.$$

Then for each  $i$ ,  $w_i \in \text{span}(\gamma_i) = W_i$  and

$$w_1 + w_2 + \cdots + w_k = \sum_{i,j} a_{ij} v_{ij} = 0.$$

Since  $0 \in W_i$  for each  $i$  and  $0 + 0 + \cdots + 0 = w_1 + w_2 + \cdots + w_k$ , (c) implies that  $w_i = 0$  for all  $i$ . Thus

$$0 = w_i = \sum_{j=1}^{m_i} a_{ij} v_{ij}$$

for each  $i$ . But each  $\gamma_i$  is linearly independent, and hence  $a_{ij} = 0$  for all  $i$  and  $j$ . Consequently  $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$  is linearly independent and therefore is a basis for  $V$ .

Clearly (e) follows immediately from (d).

Finally, we assume (e) and prove (a). For each  $i$ , let  $\gamma_i$  be an ordered basis for  $W_i$  such that  $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$  is an ordered basis for  $V$ . Then

$$V = \text{span}(\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k)$$

$$= \text{span}(\gamma_1) + \text{span}(\gamma_2) + \cdots + \text{span}(\gamma_k) = \sum_{i=1}^k W_i$$

by repeated applications of Exercise 14 of Section 1.4. Fix  $j$  ( $1 \leq j \leq k$ ), and suppose that, for some nonzero vector  $v \in V$ ,

$$v \in W_j \cap \sum_{i \neq j} W_i.$$

Then

$$v \in W_j = \text{span}(\gamma_j) \quad \text{and} \quad v \in \sum_{i \neq j} W_i = \text{span} \left( \bigcup_{i \neq j} \gamma_i \right).$$

Hence  $v$  is a nontrivial linear combination of both  $\gamma_j$  and  $\left( \bigcup_{i \neq j} \gamma_i \right)$ , so that  $v$  can be expressed as a linear combination of  $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$  in more than one way. But these representations contradict Theorem 1.8 (p. 43), and so we conclude that

$$W_j \cap \sum_{i \neq j} W_i = \{0\},$$

proving (a). ■

With the aid of Theorem 5.10, we are able to characterize diagonalizability in terms of direct sums.

**Theorem 5.11.** *A linear operator  $T$  on a finite-dimensional vector space  $V$  is diagonalizable if and only if  $V$  is the direct sum of the eigenspaces of  $T$ .*

*Proof.* Let  $\lambda_1, \lambda_2, \dots, \lambda_k$  be the distinct eigenvalues of  $T$ .

First suppose that  $T$  is diagonalizable, and for each  $i$  choose an ordered basis  $\gamma_i$  for the eigenspace  $E_{\lambda_i}$ . By Theorem 5.9,  $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$  is a basis for  $V$ , and hence  $V$  is a direct sum of the  $E_{\lambda_i}$ 's by Theorem 5.10.

Conversely, suppose that  $V$  is a direct sum of the eigenspaces of  $T$ . For each  $i$ , choose an ordered basis  $\gamma_i$  of  $E_{\lambda_i}$ . By Theorem 5.10, the union  $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$  is a basis for  $V$ . Since this basis consists of eigenvectors of  $T$ , we conclude that  $T$  is diagonalizable. ■

### Example 10

Let  $T$  be the linear operator on  $\mathbb{R}^4$  defined by

$$T(a, b, c, d) = (a, b, 2c, 3d).$$