EXAMPLE 3.6.4 Let f be a complex-valued continuous† function of the real variable x defined on the interval $\{x \mid a \le x \le b\}$. Prove that

$$\left| \int_a^b f(x)\,dx \right| \le (b-a)M, \text{ where } M = \max |f(x)|$$

We can consider f as a vector in the complex vector space of complex-valued continuous functions defined on the interval $\{x \mid a \le x \le b\}$. The reader should verify that this is a vector space. In this space we introduce the scalar product

$$(f \cdot g) = \int_a^b f(x)\overline{g(x)}\,dx$$

The reader should check the five properties. Using this scalar product and the Cauchy inequality for f and $g = 1$, we have

$$\left| \int_a^b f(x)\,dx \right| \le \left(\int_a^b |f(x)|^2\,dx \right)^{1/2} \left(\int_a^b dx \right)^{1/2}$$
$$\le [(b-a)M^2]^{1/2}(b-a)^{1/2} = (b-a)M$$

It is very common to refer to vectors in a vector space as *points*. For example, in R^3 if we have a vector (x,y,z), we could consider the three numbers as the coordinates of a point in three-dimensional euclidean space. Thinking, in general, of vectors as points in a vector space V with a scalar product, we can introduce the concept of distance between two points. Let $\mathbf{u}$ and $\mathbf{v}$ be in V; then we define the distance between $\mathbf{u}$ and $\mathbf{v}$ as $\|\mathbf{u} - \mathbf{v}\|$. This distance function has the following four desirable properties:

(i) $\|\mathbf{u} - \mathbf{v}\| = \|\mathbf{v} - \mathbf{u}\|$.
(ii) $\|\mathbf{u} - \mathbf{v}\| \ge 0$.
(iii) $\|\mathbf{u} - \mathbf{v}\| = 0$ if and only if $\mathbf{u} = \mathbf{v}$.
(iv) $\|\mathbf{u} - \mathbf{v}\| \le \|\mathbf{u} - \mathbf{w}\| + \|\mathbf{w} - \mathbf{v}\|$ (triangle inequality).

These properties follow easily from Theorem 3.6.3. For example, for (i), $\|\mathbf{u} - \mathbf{v}\| = \|(-1)(\mathbf{v} - \mathbf{u})\| = |-1|\,\|\mathbf{v} - \mathbf{u}\| = \|\mathbf{v} - \mathbf{u}\|$. For (iv), we have

$$\|\mathbf{u} - \mathbf{v}\| = \|(\mathbf{u} - \mathbf{w}) + (\mathbf{w} - \mathbf{v})\| \le \|\mathbf{u} - \mathbf{w}\| + \|\mathbf{w} - \mathbf{v}\|$$

Whenever a vector space has a distance between pairs of points defined satisfying properties (i) to (iv), we say it is a *metric space*. We have therefore shown that

† Continuous here means that both real and imaginary parts are continuous functions of x. If $f(x) = u(x) + iv(x)$, where u and v are real, then
$\int_a^b f(x)\,dx = \int_a^b u(x)\,dx + i\int_a^b v(x)\,dx$

every vector space with a scalar product is a metric space. There are, however, distance functions which are not derivable from a scalar product (see Exercise 3.6.11). There are even vector spaces which do not have a distance function, but such discussions are beyond the scope of this book.

EXERCISES 3.6

1 Let $\mathbf{u} = (1,-2,3,0)$ and $\mathbf{v} = (-2,4,5,-1)$. Compute $(\mathbf{u} \cdot \mathbf{v})$, $(\mathbf{v} \cdot \mathbf{u})$, $(2\mathbf{u} \cdot \mathbf{v})$, and $(\mathbf{u} \cdot 4\mathbf{u} + 3\mathbf{v})$.

2 Consider the space of continuous real-valued functions defined on the interval $\{x \mid 0 \le x \le 2\pi\}$. Let $f(x) = \sin x$ and $g(x) = \cos x$. Compute $(f \cdot g)$, $\|f\|$, and $\|g\|$.

3 Show that the space of Example 3.6.4 is a complex vector space. Show that the product of this example is a scalar product. Let $f(x) = e^{ix}$ and $g(x) = e^{2ix}$. Compute $(f \cdot g)$, where $a = 0$ and $b = 2\pi$.

4 Let V be a complex vector space with scalar product $(\mathbf{u} \cdot \mathbf{v})$. Let $\|\mathbf{u}\| = (\mathbf{u} \cdot \mathbf{u})^{1/2}$ be the norm. Show that $\|\mathbf{u} - \mathbf{v}\| \ge \big|\|\mathbf{u}\| - \|\mathbf{v}\|\big|$. *Hint*: Apply the triangle inequality to $\mathbf{u} = (\mathbf{u} - \mathbf{v}) + \mathbf{v}$ and $\mathbf{v} = (\mathbf{v} - \mathbf{u}) + \mathbf{u}$.

5 Let V be a real vector space with scalar product $(\mathbf{u} \cdot \mathbf{v})$. Let $\|\mathbf{u}\| = (\mathbf{u} \cdot \mathbf{u})^{1/2}$ be the norm. Prove the pythagorean theorem: $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ if and only if $(\mathbf{u} \cdot \mathbf{v}) = 0$. Why is this called the pythagorean theorem?

6 Let V be a complex vector space with scalar product $(\mathbf{u} \cdot \mathbf{v})$. Let $\|\mathbf{u}\| = (\mathbf{u} \cdot \mathbf{u})^{1/2}$ be the norm. Prove the parallelogram rule: $\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$. Why is this called the parallelogram rule?

7 Show that Cauchy's inequality is an equality if and only if the two vectors are proportional. *Hint*: Consider the proof for the case when the discriminant is zero.

8 Show that the triangle inequality (Theorem 3.6.3) is an equality if and only if the two vectors are proportional and the constant of proportionality is a non-negative real number.

9 Let f be a continuous real-valued function defined on the interval $\{x \mid a \le x \le b\}$. Prove that

$$\int_a^b |f(x)|^2\,dx \le \left(\int_a^b |f(x)|\,dx \right)^{1/2} \left(\int_a^b |f(x)|^3\,dx \right)^{1/2}$$

10 Let V be the vector space of n-tuples of real numbers. If $\mathbf{u} = (u_1, u_2, \ldots, u_n)$, let $\|\mathbf{u}\|^* = |u_1| + |u_2| + \cdots + |u_n|$. Show that $\|\mathbf{u} - \mathbf{v}\|^*$ satisfies the four properties of a distance function.

11 Show that $\|\mathbf{u}\|^*$ of Exercise 10 cannot be derived from a scalar product. *Hint*: See Exercise 6.

12 Prove that a scalar product can be defined for any finite-dimensional vector space. *Hint*: If the dimension is $n \geq 1$, there is a basis $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$. Then $(\mathbf{v} \cdot \mathbf{w}) = \sum_{i=1}^{n} v_i \bar{w}_i$ is a scalar product, where v_i and w_i are coordinates with respect to the basis.

3.7 ORTHONORMAL BASES

In finite-dimensional vector spaces with a scalar product, we can select bases with special properties. These are called *orthonormal bases*, and they have many desirable properties, which we shall bring out in this section.

Definition 3.7.1 Let V be a vector space with a scalar product $(\mathbf{u} \cdot \mathbf{v})$. Two nonzero vectors are orthogonal if $(\mathbf{u} \cdot \mathbf{v}) = 0$. A vector $\mathbf{u}$ is normalized if $\|\mathbf{u}\| = 1$. A set of vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ is orthonormal if $(\mathbf{u}_i \cdot \mathbf{u}_j) = \delta_{ij}$, $i = 1, 2, \ldots, k; j = 1, 2, \ldots, k$.

Theorem 3.7.1 A set of orthonormal vectors is independent.

PROOF Let $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ be an orthonormal set. Consider $c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k = \mathbf{0}$. Let $1 \leq j \leq k$. Then

$$0 = (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k \cdot \mathbf{u}_j) = c_j$$

Theorem 3.7.2 Every finite-dimensional vector space which is not the zero space has an orthonormal basis.

PROOF If V has dimension $n > 0$, then it has a basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, none of which is zero. We shall now discuss a process for constructing an orthonormal basis from a given basis. We start with $\mathbf{v}_1$. Let $\mathbf{u}_1 = \mathbf{v}_1/\|\mathbf{v}_1\|$. Then $\|\mathbf{u}_1\| = 1$. Next let

$$\mathbf{u}_2 = \frac{\mathbf{v}_2 - c_1\mathbf{u}_1}{\|\mathbf{v}_2 - c_1\mathbf{u}_1\|}$$

where $c_1 = (\mathbf{v}_2 \cdot \mathbf{u}_1)$. Then

$$(\mathbf{u}_2 \cdot \mathbf{u}_1) = \frac{(\mathbf{v}_2 \cdot \mathbf{u}_1) - c_1(\mathbf{u}_1 \cdot \mathbf{u}_1)}{\|\mathbf{v}_2 - c_1\mathbf{u}_1\|} = 0$$

and $\|\mathbf{u}_2\| = 1$. We must check that $\|\mathbf{v}_2 - c_1\mathbf{u}_1\| \neq 0$. If not, $\mathbf{v}_2$ would be a multiple of $\mathbf{v}_1$ and the v's would not be independent. We now have $\mathbf{u}_1$ and $\mathbf{u}_2$ orthonormal. Next we let

$$\mathbf{u}_3 = \frac{\mathbf{v}_3 - c_2\mathbf{u}_1 - c_3\mathbf{u}_2}{\|\mathbf{v}_3 - c_2\mathbf{u}_1 - c_3\mathbf{u}_2\|}$$

EXERCISES 3.7

1 Test the following set of vectors in R^3 for independence and construct from it an orthonormal basis: (1,0,1), (1,−1,1), (0,1,1).

2 Test the following set of vectors in R^4 for independence and construct from it an orthonormal basis: (1,0,1,0), (1,−1,0,1), (0,1,−1,1), (1,−1,1,−1).

3 Consider the space of real-valued polynomials of degree 2 or less defined on the interval $\{x \mid -1 \le x \le 1\}$. Using the scalar product $(p \cdot q) = \int_{-1}^{1} p(x)q(x)\,dx$, construct an orthonormal basis from the independent polynomials 1, x, x^2.

4 Consider the n-dimensional real vector space V. Let $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ and $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be two orthonormal bases for V such that $\mathbf{v}_i = \sum_{k=1}^{n} a_{ki}\mathbf{u}_k$, $i = 1, 2, \ldots, n$. Prove that the matrix A with elements a_{ij} is orthogonal. Express the u's in terms of the v's.

5 Consider the n-dimensional complex vector space V. Let $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ and $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be two orthonormal bases for V such that $\mathbf{v}_i = \sum_{k=1}^{n} a_{ki}\mathbf{u}_k$, $i = 1, 2, \ldots, n$. Prove that the matrix A with elements a_{ij} is unitary. Express the u's in terms of the v's.

6 Given two arbitrary bases $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ and $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ in a vector space V such that $\mathbf{v}_i = \sum_{k=1}^{n} a_{ki}\mathbf{u}_k$. Prove that the matrix A with elements a_{ij} is nonsingular. Express the u's in terms of the v's.

7 Consider the plane given implicitly by the equation $x + y + z = 0$ in euclidean three-dimensional space R^3. Construct an orthonormal basis as follows: select an orthonormal basis for the subspace consisting of those points in the given plane and then find a third unit vector orthogonal to the given plane.

8 Consider the subspace of R^4 spanned by the two vectors $\mathbf{u}_1 = (1,0,1,0)$ and $\mathbf{u}_2 = (1,-1,1,-1)$. Construct an orthonormal basis $\mathbf{v}_1$, $\mathbf{v}_2$ for this subspace. Now construct an orthonormal basis for R^4 containing $\mathbf{v}_1$ and $\mathbf{v}_2$.

9 Given any subspace U of dimension $m \ge 1$ in an n-dimensional vector space V $(m < n)$, prove that V has an orthonormal basis consisting of m vectors in U and $n - m$ vectors orthogonal to all vectors in U.

10 Given a vector $\mathbf{v}$ in an n-dimensional vector space V and given a subspace U of dimension m $(1 \le m < n)$. Prove that $\mathbf{v}$ can be expressed uniquely as $\mathbf{v} = \mathbf{u} + \mathbf{w}$, where $\mathbf{u}$ is in U and $\mathbf{w}$ is orthogonal to U (orthogonal to all vectors in U). $\mathbf{u}$ is called the *projection* of $\mathbf{v}$ on U.

11 Find the projection of (1,2,3) on the plane given implicitly by $x + y + z = 0$ (see Exercise 7).

12 Find the projection of (1,2,3,4) on the subspace of R^4 spanned by $\mathbf{u}_1 = (1,0,1,0)$ and $\mathbf{u}_2 = (1,-1,1,-1)$ (see Exercise 8).

13 Show that the space of $m \times n$ real matrices is isomorphic to R^{nm}. Exhibit a one-to-one correspondence.

14 Show that the space of $m \times n$ complex matrices is isomorphic to C^{nm}. Exhibit a one-to-one correspondence.

15 Show that the space of complex-valued polynomials in the complex variable z of degree n or less is isomorphic to C^{n+1}.

16 Prove that two finite-dimensional vector spaces which are isomorphic have the same dimension.

*3.8 INFINITE-DIMENSIONAL VECTOR SPACES

We have already established the existence of infinite-dimensional vector spaces; for example, the space of real-valued continuous functions defined on the interval $\{x \mid 0 \le x \le 1\}$. However, we have not had much to say about such spaces for a couple of good reasons. One is that our primary concern in this book is with finite-dimensional vector spaces. The other is that the theory of infinite-dimensional spaces is quite a bit more complicated than that for finite-dimensional spaces. This theory is properly a part of the branch of mathematics called *functional analysis*. However, it is possible to give a very brief introduction to the subject, which we propose to do in this section.

One of the easiest ways to obtain an infinite-dimensional vector space is to extend from R^n, the space of n-tuples of real numbers, to the space of infinite sequences of real numbers (infinite-tuples). Let $\mathbf{u} = (u_1, u_2, u_3, \ldots)$ and $\mathbf{v} = (v_1, v_2, v_3, \ldots)$ be infinite sequences of real numbers. We shall say that $\mathbf{u} = \mathbf{v}$ if $u_i = v_i$, for all positive integers i. We define the sum $\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2, u_3 + v_3, \ldots)$ and multiplication by a real scalar a as $a\mathbf{u} = (au_1, au_2, au_3, \ldots)$. The zero vector we can define as $\mathbf{0} = (0, 0, 0, \ldots)$ and the negative by $-\mathbf{u} = (-u_1, -u_2, -u_3, \ldots)$. It is easy to verify that we have a real vector space. However, since we shall want to have a scalar product in this space, we shall restrict the sequences somewhat. We shall want to define the scalar product

$$(\mathbf{u} \cdot \mathbf{v}) = u_1v_1 + u_2v_2 + u_3v_3 + \cdots = \sum_{i=1}^{\infty} u_iv_i$$

and hence the norm as

$$\|\mathbf{u}\| = \left(\sum_{i=1}^{\infty} u_i^2\right)^{1/2}$$

Since we are now dealing with infinite sequences, in order to ensure convergence we shall restrict our sequences to those such that $\sum_{i=1}^{\infty} u_i^2 < \infty$. Since we have put a restriction on the sequences which we have in the space, we shall have to recheck the axioms. The only ones which can cause trouble are A1 and

PROOF Clearly $f \circ g$ is defined on all of U. Also, if U and V are defined on the same scalars and so are V and W, then the same is true of U and W. Finally,

$$\begin{aligned}[f \circ g](a\mathbf{u}_1 + b\mathbf{u}_2) &= g[f(a\mathbf{u}_1 + b\mathbf{u}_2)] \\ &= g[af(\mathbf{u}_1) + bf(\mathbf{u}_2)] \\ &= ag[f(\mathbf{u}_1)] + bg[f(\mathbf{u}_2)] \\ &= a[f \circ g](\mathbf{u}_1) + b[f \circ g](\mathbf{u}_2)\end{aligned}$$

EXERCISES 4.2

1 Let U be R^2 and V be R^2. Let $f(\mathbf{u})$ be the reflection of $\mathbf{u}$ in the x axis; that is, if $\mathbf{u} = (x,y)$, then $f(\mathbf{u}) = (x,-y)$. Show that f is a linear transformation. Find the null space and range of f.

2 Let U be R^2 and V be R^2. Let $f(\mathbf{u})$ be the orthogonal complement of $\mathbf{u}$ with respect to the line $x = y$; that is, if $\mathbf{z}$ is a unit vector along the given line, then $(\mathbf{u} \cdot \mathbf{z})\mathbf{z}$ is the projection on the line and $\mathbf{u} - (\mathbf{u} \cdot \mathbf{z})\mathbf{z}$ is the orthogonal complement. Show that f is a linear transformation. Find the null space and range of f.

3 Let U be R^3 and V be R^3. Let $f(\mathbf{u})$ be the reflection of $\mathbf{u}$ in the xy plane; that is, if $\mathbf{u} = (x,y,z)$, then $f(\mathbf{u}) = (x,y,-z)$. Show that f is a linear transformation. Find the null space and range of f.

4 Let U be R^3 and V be R^3. Let $f(\mathbf{u})$ be the reflection of $\mathbf{u}$ in the z axis; that is, if $\mathbf{u} = (x,y,z)$, then $f(\mathbf{u}) = (-x,-y,z)$. Show that f is a linear transformation. Find the null space and range of f.

5 Let U be R^3 and V be R^3. Let $f(\mathbf{u})$ be the orthogonal complement of $\mathbf{u}$ with respect to the plane represented implicitly by $x + y + z = 0$. Show that f is a linear transformation. Find the null space and range of f.

6 Let U be C^n and V be C^n. Let $f(\mathbf{u}) = c\mathbf{u}$, where c is a complex number. Show that f is a linear transformation. Find the null space and range of f.

7 Let $U = R^4$ and $V = R^3$. Let (x_1,x_2,x_3,x_4) be coordinates of $\mathbf{u}$ relative to the standard basis in U. Let (y_1,y_2,y_3) be the coordinates of $f(\mathbf{u})$ relative to the standard basis in V, and

$$\begin{aligned}y_1 &= x_1 - x_2 + 2x_3 - x_4 \\ y_2 &= -x_1 + 2x_2 - 3x_3 + x_4 \\ y_3 &= x_1 - 3x_2 + 4x_3 - x_4\end{aligned}$$

Show that f is a linear transformation. Find the null space and range of f.

8 Let $U = C^n$ and $V = C^1$. Let $f(\mathbf{u}) = u_1$. Show that f is a linear transformation. Find the null space and range of f.

9 Let $U = C^n$ and $V = C^1$. Let $f(\mathbf{u}) = u_1 + u_2 + \cdots + u_n$. Show that f is a linear transformation. Find the null space and range of f.

10 Let U be the space of continuous real-valued functions defined on the interval $\{x \mid 0 \le x \le 1\}$. Let $T[f] = f(x_0)$, $0 \le x_0 \le 1$, where x_0 is fixed. Show that T is a linear transformation. Find the null space and range of T.

11 Find the null space and range of the linear transformations of Examples 4.2.8 and 4.2.9.

12 Prove that for any linear transformation f, $f(\mathbf{0}) = \mathbf{0}$.

13 Show that condition (iii) of Definition 4.2.2 can be replaced by the two conditions $f(a\mathbf{u}) = af(\mathbf{u})$ and $f(\mathbf{u}_1 + \mathbf{u}_2) = f(\mathbf{u}_1) + f(\mathbf{u}_2)$.

14 Find the most general linear transformation from R^1 to R^1.

15 Let f be a linear transformation from R^n to R^n represented by $f(\mathbf{u}) = A\mathbf{u}$, where A is $n \times n$, $\mathbf{u}$ is the column matrix of coordinates relative to the standard basis in U, and $f(\mathbf{u})$ is the column matrix of coordinates relative to the standard basis in V. Show that the following statements are all equivalent by citing the appropriate theorems:

(*a*) A is nonsingular.
(*b*) $|A| \neq 0$.
(*c*) The null space of f is the zero space.
(*d*) The dimension of the range of f is n.
(*e*) A is invertible.
(*f*) f has an inverse.
(*g*) The columns of A are independent.
(*h*) The rows of A are independent.
(*i*) The equations $AX = B$ have a unique solution.
(*j*) The equations $AX = 0$ have only the trivial solution.

16 Let f be a linear transformation from R^n to R^m represented by $f(\mathbf{u}) = A\mathbf{u}$, where A is $m \times n$. Prove that f is not invertible if $n > m$.

17 A linear transformation f is said to be *onto* if every vector in the range space is a value of $f(\mathbf{u})$ for at least one $\mathbf{u}$ in the domain. If the domain is finite-dimensional, show that f is onto if and only if the dimension of the domain is equal to the dimension of the null space plus the dimension of the range space.

18 Which of the linear transformations in Exercises 1 to 10 have inverses? Find the inverses where they exist.

19 If the domain and range of a linear transformation are the same and $f(\mathbf{u}) = \mathbf{u}$, then f is called the *identity transformation*. Show that the composition of an invertible linear transformation with its inverse (in either order) is the identity.

20 Find the compositions of the linear transformations of Example 4.2.5 and Exercise 1 in both orders. Is the operation of composition commutative? Is the operation of composition associative?

Theorem 4.3.9 Let f be a linear transformation from the n-dimensional space U to the n-dimensional space V, represented by the $n \times n$ matrix A with respect to the bases $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ in U and $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ in V. Then f has an inverse f^{-1} if and only if A is invertible. The matrix representation of f^{-1} is A^{-1} with respect to the given bases in U and V.

PROOF Let X be the coordinates of $\mathbf{u}$ with respect to $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$. Then $Y = AX$ are the coordinates of $f(\mathbf{u})$ with respect to the basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Furthermore, f is invertible if and only if the dimension of the null space is zero. Therefore, f is invertible if and only if $AX = 0$ has only the trivial solution, and $AX = 0$ has only the trivial solution if and only if A is invertible. Now suppose A^{-1} exists. Then $A^{-1}Y = A^{-1}(AX) = (A^{-1}A)X = X$. Therefore, $X = A^{-1}Y$ expresses the coordinates of $\mathbf{u}$ in terms of the coordinates of $\mathbf{v} = f(\mathbf{u})$. Therefore, A^{-1} is the matrix representation of f^{-1} relative to the given bases in V and U.

EXAMPLE 4.3.4 Let $U = R^3$ and $V = R^3$. Let $f(\mathbf{u})$ be the vector obtained from $\mathbf{u}$ by first rotating $\mathbf{u}$ about the z axis through an angle of 90° in the counterclockwise direction and then through an angle of 90° in the counterclockwise direction about the x axis. We shall find a matrix representation of f with respect to the standard bases in U and V.

$$f(\mathbf{e}_1) = (0,0,1)$$
$$f(\mathbf{e}_2) = (-1,0,0)$$
$$f(\mathbf{e}_3) = (0,-1,0)$$

Therefore, f has the representation

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Now the matrix of the transformation is orthogonal and is therefore invertible. The inverse of the transformation has the representation

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$$

In the next section, we study the question of how the representation of a linear transformation changes when we change the bases in the domain and range spaces.

EXERCISES 4.3

1 Find the matrix representation of the linear transformation of Example 4.2.1 with respect to the standard bases. Find the representation with respect to arbitrary bases.

2 Find the matrix representation of the linear transformation of Example 4.2.2 with respect to arbitrary bases.

3 Let $U = R^3$ and $V = R^2$. Let $\mathbf{u}$ be any vector in R^3, and let $f(\mathbf{u})$ be the projection of $\mathbf{u}$ on the xy plane. Find the matrix representation of f with respect to the standard bases.

4 Let U be the space of real-valued polynomials of degree n or less in the real variable x. Let f be the operation of integration over the interval $\{x \mid 0 \le x \le 1\}$. Find the matrix representation of f with respect to the basis $1, x, x^2, \ldots, x^n$ in U.

5 Find the matrix representation of the linear transformation of Exercise 4.2.1 with respect to the standard bases. Find the representation with respect to the basis (1,1), (1, −1) in both domain and range spaces.

6 Find the matrix representation of the linear transformation of Exercise 4.2.2 with respect to the standard bases. Find the representation with respect to the basis (1,1), (1, −1) in both domain and range spaces.

7 Find the matrix representation of the linear transformation of Exercise 4.2.4 with respect to the standard bases. Find the representation of the inverse with respect to the standard bases.

8 Find the matrix representation of the linear transformation of Exercise 4.2.7 with respect to the standard bases. Find a basis for the null space of the transformation and a basis for the domain consisting of this basis and other vectors orthogonal to the null space. Find a representation of the linear transformation with respect to this new basis and the standard basis in the range space.

9 Find matrix representations of the linear transformations in Exercises 4.2.8 and 4.2.9 with respect to the standard bases in domain and range spaces.

10 Let $U = R^3$ and $V = R^3$. Let A be the representation of a linear transformation f with respect to the standard bases. Which transformations are invertible? Find the inverse if its exists.

(a) $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 2 & 0 & 3 \end{pmatrix}$ (b) $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ -1 & 5 & -1 \end{pmatrix}$

11 Show that the representation of a linear transformation from a finite-dimensional domain U to a finite-dimensional range space V with respect to given bases in U and V is unique. *Hint:* If $Y = AX$ and $Y = BX$, then $0 = (A - B)X$ for *all* vectors X.

12 Let f be a linear transformation from U with basis $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ to V with basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Let g be a linear transformation from V to W with basis $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$. If f has the representation A and g has the representation B with

respect to the given bases, where A and B are nonsingular, show that $(f \circ g)^{-1}$ has the representation $A^{-1}B^{-1}$.

4.4 CHANGE OF BASES

In this section, we shall again consider linear transformation with a finite-dimensional domain and range space. If we pick a basis in the domain and a basis in the range space, we shall have a unique matrix representation of the transformation. If we change the bases in the domain and the range spaces, we shall, in general, change the representation. Our purpose is to find an easy way to find the new representation. Our approach will be the following. We shall first show that a change of basis in an n-dimensional vector space can be interpreted as an invertible linear transformation from C^n to C^n or R^n to R^n, depending on whether the space is complex or real. Then we shall show that the change of representation of a linear transformation can be obtained by composing three linear transformations.

Theorem 4.4.1 Let U be an n-dimensional vector space with a basis $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$. Let $\mathbf{u}'_1, \mathbf{u}'_2, \ldots, \mathbf{u}'_n$ be another basis for U, such that

$$\begin{aligned}
\mathbf{u}_1 &= p_{11}\mathbf{u}'_1 + p_{21}\mathbf{u}'_2 + \cdots + p_{n1}\mathbf{u}'_n \\
\mathbf{u}_2 &= p_{12}\mathbf{u}'_1 + p_{22}\mathbf{u}'_2 + \cdots + p_{n2}\mathbf{u}'_n \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\mathbf{u}_n &= p_{1n}\mathbf{u}'_1 + p_{2n}\mathbf{u}'_2 + \cdots + p_{nn}\mathbf{u}'_n
\end{aligned}$$

If X is the column matrix of coordinates of $\mathbf{u}$ with respect to $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ and X' is the column matrix of coordinates of $\mathbf{u}$ with respect to $\mathbf{u}'_1, \mathbf{u}'_2, \ldots, \mathbf{u}'_n$, then $X' = PX$, where P is the $n \times n$ matrix with elements p_{ij}. Also, P is invertible and $X = P^{-1}X'$.

PROOF Let $\mathbf{u} = x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \cdots + x_n\mathbf{u}_n$. Then

$$\begin{aligned}
\mathbf{u} &= x_1(p_{11}\mathbf{u}'_1 + p_{21}\mathbf{u}'_2 + \cdots + p_{n1}\mathbf{u}'_n) \\
&\quad + x_2(p_{12}\mathbf{u}'_1 + p_{22}\mathbf{u}'_2 + \cdots + p_{n2}\mathbf{u}'_n) \\
&\quad + \cdots + x_n(p_{1n}\mathbf{u}'_1 + p_{2n}\mathbf{u}'_2 + \cdots + p_{nn}\mathbf{u}'_n) \\
&= (p_{11}x_1 + p_{12}x_2 + \cdots + p_{1n}x_n)\mathbf{u}'_1 \\
&\quad + (p_{21}x_1 + p_{22}x_2 + \cdots + p_{2n}x_n)\mathbf{u}'_2 \\
&\quad + \cdots + (p_{n1}x_1 + p_{n2}x_2 + \cdots + p_{nn}x_n)\mathbf{u}'_n \\
&= x'_1\mathbf{u}'_1 + x'_2\mathbf{u}'_2 + \cdots + x'_n\mathbf{u}'_n
\end{aligned}$$

Therefore,

$$\begin{aligned}
x'_1 &= p_{11}x_1 + p_{12}x_2 + \cdots + p_{1n}x_n \\
x'_2 &= p_{21}x_1 + p_{22}x_2 + \cdots + p_{2n}x_n \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
x'_n &= p_{n1}x_1 + p_{n2}x_2 + \cdots + p_{nn}x_n
\end{aligned}$$

Definition 4.4.1 Let A and B be $n \times n$ matrices. If there exists an invertible $n \times n$ matrix P such that $B = PAP^{-1}$, then we say that B is similar to A and B is obtained from A by a similarity transformation.

Theorem 4.4.4† Let A, B, and C be $n \times n$ matrices. Then (i) A is similar to A, for all A, (ii) if A is similar to B, then B is similar to A, and (iii) if A is similar to B, and B is similar to C, then A is similar to C.

PROOF (i) The $n \times n$ identity matrix is invertible, and $A = IAI^{-1}$.

(ii) If A is similar to B, there is an invertible matrix P such that $A = PBP^{-1}$. But then $B = P^{-1}AP = P^{-1}A(P^{-1})^{-1}$.

(iii) If A is similar to B, there is an invertible matrix P such that $A = PBP^{-1}$. If B is similar to C, there is an invertible matrix Q such that $B = QCQ^{-1}$. Then

$$A = PBP^{-1} = P(QCQ^{-1})P^{-1} = (PQ)C(Q^{-1}P^{-1}) = SCS^{-1}$$

where $S = PQ$.

Some of the other important properties of similarity transformations are given by the next theorem.

Theorem 4.4.5

(i) If A is similar to B, then $|A| = |B|$.

(ii) If A_1 is similar to B_1 and A_2 is similar to B_2 under the same similarity transformation, then $A_1 + A_2$ is similar to $B_1 + B_2$.

(iii) If A is similar to B, then A^k is similar to B^k under the same similarity transformation for any positive integer k.

(iv) If A is similar to B, then $p(A)$ is similar to $p(B)$ under the same similarity transformation, where p is a polynomial.‡

(v) If A is similar to B and A is nonsingular, then B is nonsingular and A^{-1} is similar to B^{-1}.

PROOF (i) There exists a nonsingular matrix P such that $A = PBP^{-1}$. Hence $|A| = |P|\,|B|\,|P^{-1}| = |B|\,|P|\,|P^{-1}| = |B|$, since $|P|\,|P^{-1}| = |PP^{-1}| = |I| = 1$.

(ii) There exists a nonsingular matrix P such that $A_1 = PB_1P^{-1}$ and $A_2 = PB_2P^{-1}$. Then

$$A_1 + A_2 = PB_1P^{-1} + PB_2P^{-1} = P(B_1 + B_2)P^{-1}$$

† This theorem shows that similarity is an equivalence relation.

‡ If A and B are real, p is to have real coefficients; while if A and B are complex, p is to have complex coefficients.

(iii) There exists a nonsingular matrix P such that $A = PBP^{-1}$. Then $A^2 = (PBP^{-1})(PBP^{-1}) = PB(P^{-1}P)BP^{-1} = P(BB)P^{-1} = PB^2P^{-1}$ The rest follows by induction.

(iv) There is a nonsingular matrix P such that $A = PBP^{-1}$. Let c be a scalar. Then $cA = c(PBP^{-1}) = P(cB)P^{-1}$. Therefore, similarity is preserved under multiplication by a scalar. Now let $p(A) = a_0I + a_1A + a_2A^2 + \cdots + a_kA^k$. Then by (ii) and (iii) of this theorem,

$$\begin{aligned} Pp(B)P^{-1} &= a_0PIP^{-1} + a_1PBP^{-1} + a_2PB^2P^{-1} + \cdots + a_kPB^kP^{-1} \\ &= a_0I + a_1A + a_2A^2 + \cdots + a_kA^k = p(A) \end{aligned}$$

(v) There is a nonsingular matrix P such that $A = PBP^{-1}$. By (i), $|B| = |A| \neq 0$. Therefore, B is nonsingular. Also $A^{-1} = (PBP^{-1})^{-1} = PB^{-1}P^{-1}$, showing that A^{-1} is similar to B^{-1}.

EXERCISES 4.4

1 Let $U = V = R^2$, and let $f(\mathbf{u})$ be the reflection of $\mathbf{u}$ in the line $x = y$. Find the matrix representation of f:

(*a*) Relative to the standard basis in both U and V.

(*b*) Relative to the standard basis in U and the basis (1,1), (1,−1) in V.

(*c*) Relative to the standard basis in V and the basis (1,1), (1,−1) in U.

(*d*) Relative to the basis (1,1), (1,−1) in both U and V.

2 Let $U = V = R^3$, and let $f(\mathbf{u})$ be the reflection of $\mathbf{u}$ in the plane given implicitly by $x + y + z = 0$. Find the matrix representation of f:

(*a*) Relative to the standard basis in both U and V.

(*b*) Relative to the standard basis in U and the basis (1,0,−1), (1,−2,1), (1,1,1) in V.

(*c*) Relative to the standard basis in V and the basis (1,0,−1), (1,−2,1), (1,1,1) in U.

(*d*) Relative to the basis (1,0,−1), (1,−2,1), (1,1,1) in both U and V.

3 Show that the matrix representing the change of basis from one orthonormal set to another in a complex vector space is a unitary matrix.

4 Consider the linear transformation of Example 4.3.4. Find a vector which is transformed into itself. Use this vector and two other vectors orthogonal to it and to each other as a basis. Find the representation with respect to the new basis in both domain and range space.

5 Show that if A is similar to B and A is nonsingular, then A^k is similar to B^k for all integers k.

6 Suppose A is similar to a diagonal matrix D with diagonal elements $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $|\lambda_i| < 1$ for $i = 1, 2, \ldots, n$. Let $p_k(A) = I + A + A^2 + \cdots + A^k = P(I + D + D^2 + \cdots + D^k)P^{-1}$, since $A = PDP^{-1}$. Consider $\lim_{k\to\infty} p_k(A)$. Show that this limit exists. If we denote the series $I + A + A^2 + \cdots$ by B, prove that $B = (I - A)^{-1}$. *Hint*: Consider $\lim_{k \to \infty} (I - A)p_k(A)$ and $\lim_{k \to \infty} p_k(A)(I - A)$.

7 Suppose A is similar to a diagonal matrix D with diagonal elements $\lambda_1, \lambda_2, \ldots, \lambda_n$. Let

$$p_k(A) = I + \frac{A}{1!} + \frac{A^2}{2!} + \cdots + \frac{A^k}{k!}$$
$$= P\left(I + \frac{D}{1!} + \frac{D^2}{2!} + \cdots + \frac{D^k}{k!}\right)P^{-1}$$

since $A = PDP^{-1}$. Show that $\lim_{k \to \infty} p_k(A)$ exists and is equal to

$$P\begin{pmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & e^{\lambda_n} \end{pmatrix}P^{-1}$$

4.5 CHARACTERISTIC VALUES AND CHARACTERISTIC VECTORS

In this section, we consider only linear transformations for which the domain is a subspace of the range space. Suppose the range space of f is a complex vector space, and suppose there is a complex number λ and a nonzero vector $\mathbf{u}$ such that $f(\mathbf{u}) = \lambda\mathbf{u}$. Then we say that λ is a *characteristic value* (eigenvalue) of f and $\mathbf{u}$ is a *characteristic vector* (eigenvector) of f corresponding to λ.

Definition 4.5.1 Let f be a linear transformation from the complex (real) vector space U to V, where U is contained in V. Let λ be a complex (real) number and $\mathbf{u}$ be a nonzero vector in U such that $f(\mathbf{u}) = \lambda\mathbf{u}$. Then λ is a characteristic value of f, and $\mathbf{u}$ is a characteristic vector of f corresponding to λ.

EXAMPLE 4.5.1 Let f be the identity transformation from the complex vector space U to U. Then $f(\mathbf{u}) = \mathbf{u}$, and clearly $\lambda = 1$ is a characteristic value of f with corresponding characteristic vector $\mathbf{u} \neq \mathbf{0}$. Therefore, every nonzero vector is a characteristic vector of f.

EXAMPLE 4.5.2 Let f be the zero transformation from the complex vector space U to U. Then $f(\mathbf{u}) = \mathbf{0} = 0\mathbf{u}$, and clearly 0 is a characteristic value of f

Let us find the basis (characteristic vectors) relative to which this is the representation. If $\lambda = \lambda_1 = 6$, we must solve

$$\begin{pmatrix} 3 & -3 & 0 \\ -3 & 6 & -3 \\ 0 & -3 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

or $x - y = 0$, $y - z = 0$. In other words, $x = y = z$ and a characteristic vector is $\mathbf{v}_1 = \mathbf{u}_1 + \mathbf{u}_2 + \mathbf{u}_3$. If $\lambda = \lambda_2 = 9$, we must solve

$$\begin{pmatrix} 0 & -3 & 0 \\ -3 & 3 & -3 \\ 0 & -3 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

or $y = 0$, $x + z = 0$, and a corresponding characteristic vector is $\mathbf{v}_2 = \mathbf{u}_1 - \mathbf{u}_3$. If $\lambda = \lambda_3 = 15$, we must solve

$$\begin{pmatrix} -6 & -3 & 0 \\ -3 & -3 & -3 \\ 0 & -3 & -6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

or $2x + y = 0$, $x + y + z = 0$. A corresponding characteristic vector is $\mathbf{v}_3 = \mathbf{u}_1 - 2\mathbf{u}_2 + \mathbf{u}_3$.

In Example 4.5.7, we had two characteristic values and only two independent characteristic vectors in R^3. Therefore, it will not always be possible to find a diagonal representation. On the other hand, Example 4.5.6 illustrates that there may be n independent characteristic vectors even when there are not n distinct characteristic values. Hence, Theorem 4.5.3 gives a sufficient but not necessary condition for a diagonal representation. In the next section, we take up a couple of special cases where it will always be possible to obtain a diagonal representation.

EXERCISES 4.5

1 Find the characteristic values and characteristic vectors of the following matrices:

(*a*) $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ (*b*) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ (*c*) $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ (*d*) $\begin{pmatrix} 1 & 3 \\ -2 & 1 \end{pmatrix}$

2 Find the characteristic values and characteristic vectors of the following matrices:

$$(a)\ \begin{pmatrix}3&1&0\\1&3&0\\0&0&2\end{pmatrix}\qquad (b)\ \begin{pmatrix}2&1&2\\0&-1&3\\0&0&3\end{pmatrix}\qquad (c)\ \begin{pmatrix}5&1&1\\-3&1&-3\\-2&-2&-2\end{pmatrix}$$

$$(d)\ \begin{pmatrix}1&1&1\\0&1&1\\0&0&2\end{pmatrix}\qquad (e)\ \begin{pmatrix}-8&5&4\\5&3&1\\4&1&0\end{pmatrix}$$

3 Let $U = V = R^2$, and let $f(\mathbf{u})$ be the reflection of $\mathbf{u}$ in the line $y = -x$. Find all the characteristic values and characteristic vectors of f. Find a representation of f which is diagonal.

4 Let $U = V = R^3$, and let $f(\mathbf{u})$ be the projection of $\mathbf{u}$ on the plane given implicitly by $x + 2y - z = 0$. Find all the characteristic values and characteristic vectors of f. Find a representation of f which is diagonal.

5 Let $U = V = R^3$, and let f be a linear transformation with the representation matrix relative to the standard basis

$$A = \begin{pmatrix}4&-20&-10\\-2&10&4\\6&-30&-13\end{pmatrix}$$

Find a basis with respect to which the representation is diagonal.

6 If λ is a characteristic value of a square matrix A, show that λ^n is a characteristic value of A^n, where n is a positive integer.

7 Show that a square matrix A is invertible if and only if $\lambda = 0$ is not a characteristic value of A.

8 Show that if λ is a characteristic value of an invertible matrix A, then λ^{-1} is a characteristic value of A^{-1}.

9 If λ is a characteristic value of a square matrix A, show that $\lambda^3 - 3\lambda^2 + \lambda - 2$ is a characteristic value of $A^3 - 3A^2 + A - 2I$.

10 If $p(\lambda) = 0$ is the characteristic equation of the $n \times n$ matrix A and A has n independent characteristic vectors $X_1, X_2, \ldots, X_n$, prove that $p(A) = 0$. *Hint*: Show that $p(A)X_i = 0$ for $i = 1, 2, \ldots, n$.

11 Show that if $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ are a set of characteristic vectors of a linear transformation f corresponding to the same characteristic value λ, then they span a subspace S such that for any $\mathbf{u}$ in S, $f(\mathbf{u}) = \lambda\mathbf{u}$. *Note*: Such subspaces are called *invariant subspaces*.

12 Let f be a linear transformation from U to U, where U is n-dimensional. Show that f has a diagonal representation if and only if the sum of the dimensions of its invariant subspaces is n.

13 Suppose we want to find a matrix C such that $C^2 = A$. (C might be called a square root of A.) Suppose A is similar to a diagonal matrix B with diagonal elements $\lambda_1, \lambda_2, \ldots, \lambda_n$, with $\lambda_i \geq 0$. Then $B = PAP^{-1}$. Let D be a diagonal matrix with diagonal elements $\pm\sqrt{\lambda_1}, \pm\sqrt{\lambda_2}, \ldots, \pm\sqrt{\lambda_n}$. Then $B = D^2$. Let $C = P^{-1}DP$. Show that $C^2 = A$. Use this method to find square roots of

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

14 Solve the following system of equations:

$$\frac{dx}{dt} = 2x + y$$

$$\frac{dy}{dt} = x + 2y$$

Hint: Let $\mathbf{u} = \begin{pmatrix}x\\y\end{pmatrix}$, and write the equations as $d\mathbf{u}/dt = A\mathbf{u}$. Find P such that $\mathbf{v} = P\mathbf{u}$. Hence, $P^{-1}\,d\mathbf{v}/dt = (AP^{-1})\mathbf{v}$ and $d\mathbf{v}/dt = (PAP^{-1})\mathbf{v}$. If

$$PAP^{-1} = \begin{pmatrix}\lambda_1&0\\0&\lambda_2\end{pmatrix}$$

then the equations are separated.

15 Solve the following system of equations:

$$\frac{dx}{dt} = 3x + z$$

$$\frac{dy}{dt} = 3y + z$$

$$\frac{dz}{dt} = x + y + 2z$$

16 Consider the differential equation

$$\frac{d^2x}{dt^2} - 3\frac{dx}{dt} - 4x = 0$$

Look for solutions of the form $x = e^{\lambda t}$. Show that λ must be a root of the equation $\lambda^2 - 3\lambda - 4 = 0$. Show that the given equation is equivalent to the system $dx/dt = y$, $dy/dt = 4x + 3y$. Compare with Exercise 14.

4.6 SYMMETRIC AND HERMITIAN MATRICES

We saw, in the last section, that an $n \times n$ matrix is similar to a diagonal matrix if and only if it has n independent characteristic vectors. We also saw square matrices which are not similar to diagonal matrices. In this section, we shall study two types of matrices, real symmetric and complex hermitian, which are always similar to diagonal matrices. We shall begin with real symmetric matrices.

At $t = 0$, we have

$$X_0 = X(0) = P^{-1}\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = c_1X_1 + c_2X_2 + \cdots + c_nX_n$$

where $X_1, X_2, \ldots, X_n$ are characteristic vectors of A. In this case, the characteristic vectors are independent, and therefore they form a basis for R^n. No matter what initial vector X_0 is prescribed, we can always find constants $c_1, c_2, \ldots, c_n$ such that $X_0 = c_1X_1 + c_2X_2 + \cdots + c_nX_n$.

EXAMPLE 4.6.4 Find a solution of the following system of differential equations

$$\frac{dx_1}{dt} = 2x_1 + 3x_2 + 3x_3$$

$$\frac{dx_2}{dt} = 3x_1 - x_2$$

$$\frac{dx_3}{dt} = 3x_1 - x_3$$

satisfying the initial conditions $x_1(0) = 1$, $x_2(0) = -2$, $x_3(0) = 0$. The system can be written as $X' = AX$, where

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

is real and symmetric. The characteristic equation is

$$\begin{vmatrix} 2-\lambda & 3 & 3 \\ 3 & -1-\lambda & 0 \\ 3 & 0 & -1-\lambda \end{vmatrix} = (\lambda + 1)(\lambda + 4)(-\lambda + 5) = 0$$

The characteristic values are $\lambda_1 = -1$, $\lambda_2 = -4$, $\lambda_3 = 5$. The corresponding characteristic vectors are

$$X_1 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \qquad X_2 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \qquad X_3 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

Therefore, a solution is

$$x_1(t) = -c_2e^{-4t} + 2c_3e^{5t}$$
$$x_2(t) = c_1e^{-t} + c_2e^{-4t} + c_3e^{5t}$$
$$x_3(t) = -c_1e^{-t} + c_2e^{-4t} + c_3e^{5t}$$

In order for $X(0) = (1,-2,0)$, we must have

$$\begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 2 \\ 1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

Solving, we have $c_1 = -1$, $c_2 = -1$, $c_3 = 0$. Our solution satisfying the initial conditions is

$$x_1(t) = e^{-4t}$$
$$x_2(t) = -e^{-t} - e^{-4t}$$
$$x_3(t) = e^{-t} - e^{-4t}$$

EXERCISES 4.6

1 Find similarity transformations which reduce each of the following matrices to diagonal form:

(a) $\begin{pmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{pmatrix}$ (b) $\begin{pmatrix} 4 & 1 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ (c) $\begin{pmatrix} 4 & -2 & 2 \\ -2 & 1 & -1 \\ 2 & -1 & 1 \end{pmatrix}$

2 Find solutions of the system of differential equations $X' = AX$, where A is each of the matrices of Exercise 1, subject to the initial conditions $X(0) = (1,2,3)$.

3 Find similarity transformations which reduce each of the following matrices to diagonal form:

(a) $\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & 2-2i \\ 2+2i & 0 \end{pmatrix}$ (c) $\begin{pmatrix} 2 & 0 & 2i \\ 0 & 2 & -2i \\ -2i & 2i & 4 \end{pmatrix}$

4 Find a solution of the system of differential equations $Z' = AZ$, where A is the matrix of Exercise 3(c), subject to the initial conditions $Z(0) = (i,0,i)$.

5 Identify the figure in the xy plane given by the equation $3x^2 + 2xy + 3y^2 = 1$. Find the axes of symmetry.

6 Identify the surface in R^3 given by the equation

$$9x^2 + 12y^2 + 9z^2 - 6xy - 6yz = 1$$

Hint: An equation of the form $\lambda_1x^2 + \lambda_2y^2 + \lambda_3z^2 = 1$ represents an ellipsoid if $\lambda_1, \lambda_2, \lambda_3$ are all positive.

7 A quadratic form $q(x_1, x_2, \ldots, x_n) = \tilde{X}AX$, where A is real and symmetric, is called positive-definite if $q > 0$ for all $X \neq 0$. Prove that $q = \tilde{X}AX$ is positive-definite if and only if all the characteristic values of A are positive.

8 Let $q = \tilde{X}AX$ be a positive-definite quadratic form. Show that $(X \cdot Y) = \tilde{X}AY$ is a scalar product for R^n.

9 A hermitian form $h(z_1, z_2, \ldots, z_n) = \tilde{\bar{Z}}AZ$, where A is hermitian, is called positive-definite if $h > 0$ for all $Z \neq 0$. Prove that $h = \tilde{\bar{Z}}AZ$ is positive-definite if and only if all the characteristic values of A are positive.

10 Let $\tilde{\bar{Z}}AZ$ be a positive-definite hermitian form. Show that $(Z_1 \cdot Z_2) = \tilde{\bar{Z}}_2AZ_1$ is a scalar product for C^n.

*4.7 JORDAN FORMS

Theorem 4.5.2 gives necessary and sufficient conditions for an $n \times n$ matrix to be similar to a diagonal matrix, namely that it should have n independent characteristic vectors. We have also seen square matrices which are similar to no diagonal matrix. In this section, we shall discuss the so-called *Jordan canonical form*, a form of matrix to which every square matrix is similar. The Jordan form is not quite as simple as a diagonal matrix but is nevertheless simple enough to make it very applicable, particularly in the solution of systems of differential equations.

Before we embark upon the discussion of Jordan forms, it will be convenient to introduce the concept of partitioning of matrices. Suppose we write

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

where A is an $m \times n$ matrix, B is an $m \times p$ matrix, C is a $q \times n$ matrix, and D is a $q \times p$ matrix. In other words, M is an $(m + q) \times (n + p)$ matrix partitioned into blocks of matrices so that each block in a given row has the same number of rows and each block in a given column has the same number of columns. The reader should convince himself that the following product rule is valid for partitioned matrices. Let

$$P = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{pmatrix} \qquad Q = \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1p} \\ B_{21} & B_{22} & \cdots & B_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ B_{n1} & B_{n2} & \cdots & B_{np} \end{pmatrix}$$

Then

$$PQ = \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1p} \\ C_{21} & C_{22} & \cdots & C_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ C_{m1} & C_{m2} & \cdots & C_{mp} \end{pmatrix}$$

where $C_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj}$ provided P and Q are partitioned in such a way that A_{ik} has the same number of columns as B_{kj} has rows for all i, j, and k.