ordered set of three numbers a vector in three-dimensional space. By analogy, we may call an ordered set of five numbers a "vector in five-dimensional space" or an ordered set of n numbers a "vector in n-dimensional space." (See Section 11 for a more careful discussion of the meaning of a vector.) A great deal of the geometrical terminology which is familiar in two and three dimensions can be extended to problems in n dimensions (that is, n variables) by using the algebra which parallels the geometry. For example, the distance from the origin to the point (x, y, z) is $\sqrt{x^2 + y^2 + z^2}$. By analogy, in a problem in the five variables x, y, z, u, v, we define the "distance" from the "origin" $(0, 0, 0, 0, 0)$ to the "point" (x, y, z, u, v) as $\sqrt{x^2 + y^2 + z^2 + u^2 + v^2}$. By using the algebra which goes with the geometry, we can easily extend such ideas as the length of a vector, the dot product of two vectors (and therefore the angle between the vectors and the idea of orthogonality), etc. (Problems 4 to 9). We saw above that an orthogonal transformation in two or three dimensions corresponds to a rotation of axes. Thus we might say, in a problem in n variables, that a linear transformation (that is, a linear change of variables) satisfying "sum of squares of new variables = sum of squares of old variables" [compare (3.2)] corresponds to a "rotation in n-dimensional space."

PROBLEMS, SECTION 3

1. Prove (3.4) in three dimensions; that is, show that if M is the matrix of a linear transformation from x, y, z to x', y', z' for which $x^2 + y^2 + z^2 = x'^2 + y'^2 + z'^2$, then M satisfies (3.4).

2. Prove the converse of Problem 1, that is, that if $M^T = M^{-1}$, then the length of a vector is not changed by the transformation M. *Hint:* The matrix r and its transpose r^T are:

$$r = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \qquad r^T = (x \quad y \quad z).$$

Show that $r^T r = x^2 + y^2 + z^2$; then use $r' = Mr$ and (1.2).

3. From (3.4), show that if M is orthogonal, then det $M = +1$ or -1. (When det $M = 1$, the transformation is called a *proper* rotation; when det $M = -1$, one or all three axes have been reflected, in addition to rotation.) *Hint:* Find det (MM^T); how is a determinant affected by interchanging rows and columns?

Do problems 4 to 9 by extending familiar definitions in two and three dimensions to the required number of dimensions.

4. Find the "distance" between the "points"

(a) $(4, -1, 2, 7)$ and $(2, 3, 1, 9)$;
(b) $(-1, 5, -3, 2, 4)$ and $(2, 6, 2, 7, 6)$.

5. Find the "length" of the "vectors"

(a) $(2, 0, 4, 6, 5)$, (b) $(-5, 1, 5, 3, -2)$.

6. Find the "cosine of the angle" between the two "vectors" in Problem 5. *Hint:* Generalize the dot product.

7. Show that the following "vectors" are orthogonal:

$$(1, -5, 7, 2, 3) \quad \text{and} \quad (2, 1, -2, 7, 1).$$

(*Hint:* Consider the "dot product.")

8. In three-dimensional space, the vectors **i**, **j**, **k** have components $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$. In, say, five-dimensional space, we can define unit basis vectors with components $(1, 0, 0, 0, 0)$, $(0, 1, 0, 0, 0)$, $(0, 0, 1, 0, 0)$, $\cdots$. Show that these five basis vectors in five dimensions are all mutually orthogonal. Generalize to n dimensions. Show that any vector in n dimensions can be written in terms of the n basis vectors.

9. Show that any $n + 1$ vectors in n dimensions are linearly dependent. *Hint:* See Chapter 3, Section 8.

4. EIGENVALUES AND EIGENVECTORS; DIAGONALIZING MATRICES

We can give the following physical interpretation to Figure 2.1 and equations (2.1) or (2.4). Suppose the (x, y) plane is covered by an elastic membrane which can be stretched, shrunk, or rotated (with the origin fixed). Then any point (x, y) of the membrane becomes some point (X, Y) after the deformation, and we can say that the matrix M describes the deformation. Notice that we use interchangeably "the point (x, y)" and "the vector **r**" and similarly for **R** [see (2.3)]. Let us now ask whether there are any vectors which are not changed in direction by the deformation, that is, vectors such that $\mathbf{R} = \mu\mathbf{r}$ where $\mu =$ const. Such vectors are called *eigenvectors* (or *characteristic vectors*) of the transformation, and the values of μ are called the *eigenvalues* (or characteristic values) of the matrix M of the transformation.

Eigenvalues To illustrate finding eigenvalues we use equations (2.2) which in matrix form become

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}. \tag{4.1}$$

The eigenvector condition, $\mathbf{R} = \mu\mathbf{r}$, is, in matrix notation,

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \mu\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \mu x \\ \mu y \end{pmatrix},$$

or written out in equation form:

$$\begin{aligned} 5x - 2y &= \mu x, \\ -2x + 2y &= \mu y, \end{aligned} \quad \text{or} \quad \begin{aligned} (5 - \mu)x - 2y &= 0, \\ -2x + (2 - \mu)y &= 0. \end{aligned} \tag{4.2}$$

If we tried to solve such a set of homogeneous equations by determinants, we would get $x = 0$, $y = 0$ (because the constants on the right-hand side are zero) unless the determinant of the coefficients were equal to zero [see Chapter 3, equation (8.11)]. In the latter case the equations would be dependent and we would get an infinite set of solutions. The condition then for there to be solutions of (4.2) other than $x = y = 0$ is that

$$\begin{vmatrix} 5-\mu & -2 \\ -2 & 2-\mu \end{vmatrix} = 0. \tag{4.3}$$

This is called the *characteristic equation* of the matrix M.

> To obtain the characteristic equation of a matrix M, we subtract μ from the elements on the main diagonal of M, and then set the determinant of the resulting matrix equal to zero.

We solve (4.3) for μ to find the characteristic values of M:

$$\begin{aligned} &(5-\mu)(2-\mu) - 4 = \mu^2 - 7\mu + 6 = 0, \\ &\mu = 1 \quad \text{or} \quad \mu = 6. \end{aligned} \tag{4.4}$$

Eigenvectors Substituting the μ values from (4.4) into (4.2), we get:

$$\begin{aligned} 2x - y &= 0 \qquad \text{from either of the equations (4.2) when } \mu = 1; \\ x + 2y &= 0 \qquad \text{from either of the equations (4.2) when } \mu = 6. \end{aligned} \tag{4.5}$$

We were looking for vectors $\mathbf{r} = \mathbf{i}x + \mathbf{j}y$ such that the transformation (2.2) would give an $\mathbf{R}$ parallel to $\mathbf{r}$. What we have found is that *any* vector $\mathbf{r}$ with x and y components satisfying either of the equations (4.5) has this property. Since equations (4.5) are equations of straight lines through the origin, such vectors lie along these lines (Figure 4.1). Then equations (4.5) show that any vector $\mathbf{r}$ from the origin to a point on $x + 2y = 0$ is changed by the transformation (2.2) to a vector in the same direction but six times as long, and any vector from the origin to a point on $2x - y = 0$ is unchanged by the transformation (2.2). These vectors (along $x + 2y = 0$ and $2x - y = 0$) are the eigenvectors of the transformation. Along these two directions (and only these), the deformation of the elastic membrane was a pure stretch with no shear (rotation).

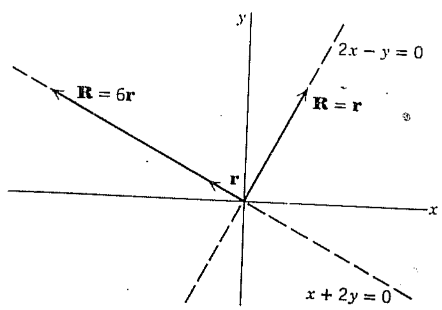

FIGURE 4.1

Diagonalizing a Matrix We next write (4.2) once with $\mu = 1$, and again with $\mu = 6$, using subscripts 1 and 2 to identify the corresponding eigenvectors:

$$\begin{aligned} 5x_1 - 2y_1 &= x_1, \qquad & 5x_2 - 2y_2 &= 6x_2, \\ -2x_1 + 2y_1 &= y_1, \qquad & -2x_2 + 2y_2 &= 6y_2. \end{aligned} \tag{4.6}$$

These four equations can be written as one matrix equation, as you can easily verify by multiplying out both sides (Problem 1):

$$\begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}. \tag{4.7}$$

All we really can say about (x_1, y_1) is that $2x_1 - y_1 = 0$; however, it is convenient to pick numerical values of x_1 and y_1 to make $\mathbf{r}_1 = (x_1, y_1)$ a unit vector, and similarly for $\mathbf{r}_2 = (x_2, y_2)$. Then we have

$$x_1 = \frac{1}{\sqrt{5}}, \qquad y_1 = \frac{2}{\sqrt{5}}, \qquad x_2 = \frac{-2}{\sqrt{5}}, \qquad y_2 = \frac{1}{\sqrt{5}}, \tag{4.8}$$

and (4.7) becomes

$$\begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}. \tag{4.9}$$

Representing these matrices by letters we can write

$$MC = CD, \qquad \text{where}$$

$$M = \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}, \qquad C = \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}, \qquad D = \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}. \tag{4.10}$$

If, as here, the determinant of C is not zero, then C has an inverse C^{-1}; let us multiply (4.10) by C^{-1} and remember that $C^{-1}C$ is the unit matrix; then $C^{-1}MC = C^{-1}CD = D$.

> $$C^{-1}MC = D. \tag{4.11}$$
>
> The matrix D has elements different from zero only down the main diagonal; it is called a *diagonal matrix*. The matrix D is called *similar* to M, and when we obtain D given M, we say that we have *diagonalized M by a similarity transformation.*

We shall see shortly that this amounts physically to a simplification of the problem by a better choice of variables. For example, in the problem of the membrane, we shall find it simpler to describe the deformation if we use axes along the eigenvectors. We shall see later several more examples of the use of the diagonalization process.

Observe that it is easy to find D; we need only solve the characteristic equation of M. Then D is a matrix with these characteristic values down the main diagonal and zeros elsewhere. We can also find C (with more work), but for many purposes only D is needed.

Note that the order of the eigenvalues down the main diagonal of D is arbitrary; for example, we could write (4.6) as

$$\begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} x_2 & x_1 \\ y_2 & y_1 \end{pmatrix} = \begin{pmatrix} x_2 & x_1 \\ y_2 & y_1 \end{pmatrix}\begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix} \tag{4.12}$$

instead of as (4.7). Then (4.11) still holds, with a different C, of course, and with

$$D = \begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix}$$

instead of as in (4.10) (Problem 1).

Meaning of C and D To see more clearly the meaning of (4.11) let us find what the matrices C and D mean physically. We consider two sets of axes (x, y) and (x', y') with (x', y') rotated through θ from (x, y) (Figure 4.2). The (x, y) and (x', y') coordinates of *one* point (or components of one vector $\mathbf{r} = \mathbf{r}'$) relative to the two systems are related by (3.1). Solving (3.1) for x and y, we have

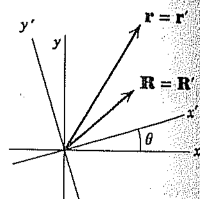

FIGURE 4.2

$$\begin{aligned} x &= x' \cos\theta - y' \sin\theta, \\ y &= x' \sin\theta + y' \cos\theta, \end{aligned} \tag{4.13}$$

or in matrix notation

$$r = Cr' \quad \text{where} \quad C = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{4.14}$$

This equation is true for *any* single vector with components given in the two systems. Suppose we have another vector $\mathbf{R} = \mathbf{R}'$ (Figure 4.2) with components X, Y and X', Y'; these components are related by

$$R = CR'. \tag{4.15}$$

Now let M be a matrix which describes a deformation of the plane in the (x, y) system. Then the equation

$$R = Mr \tag{4.16}$$

says that the vector $\mathbf{r}$ becomes the vector $\mathbf{R}$ after the deformation, both vectors given relative to the (x, y) axes. Let us ask how we can describe the deformation in the (x', y') system, that is, what matrix carries $\mathbf{r}'$ into $\mathbf{R}'$? We substitute (4.14) and (4.15) into (4.16) and find $CR' = MCr'$ or

$$R' = C^{-1}MCr'. \tag{4.17}$$

Thus the answer to our question is that

> $D = C^{-1}MC$ is the matrix which describes in the (x', y') system the same deformation that M describes in the (x, y) system.

Next we want to show that if the matrix C is chosen to make $D = C^{-1}MC$ a diagonal matrix, then the new axes (x', y') are along the directions of the eigenvectors of M. Recall from (4.10) that the columns of C are the components of the unit eigenvectors. If the eigenvectors are perpendicular, as they are in our example (see Problem 2) then new axes (x', y') along the eigenvector directions are a set of perpendicular axes rotated from axes (x, y) by some angle θ (Figure 4.3). The unit eigenvectors $\mathbf{r}_1$ and $\mathbf{r}_2$ (*unit* means $|\mathbf{r}_1| = 1$, $|\mathbf{r}_2| = 1$) are shown in Figure 4.3; from the figure we find

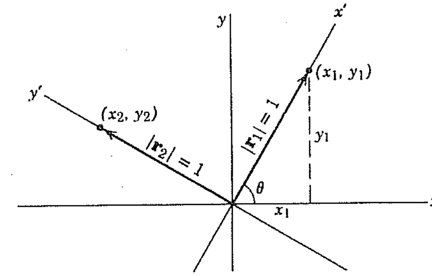

FIGURE 4.3

$$\begin{aligned} x_1 &= |\mathbf{r}_1| \cos\theta = \cos\theta, & x_2 &= -|\mathbf{r}_2| \sin\theta = -\sin\theta, \\ y_1 &= |\mathbf{r}_1| \sin\theta = \sin\theta, & y_2 &= |\mathbf{r}_2| \cos\theta = \cos\theta; \end{aligned}$$
$$C = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{4.18}$$

> Thus, the matrix C which diagonalizes M is the rotation matrix C in (4.14) when the (x', y') axes are along the directions of the eigenvectors of M.

Relative to these new axes, the diagonal matrix D describes the deformation. For our example we have

$$R' = Dr' \quad \text{or} \quad \begin{pmatrix} X' \\ Y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix} \quad \text{or} \tag{4.19}$$
$$X' = x', \quad Y' = 6y'.$$

In words, (4.19) says that [in the (x', y') system] each point (x', y') has its x' coordinate unchanged by the deformation and its y' coordinate multiplied by 6, that is, the deformation is simply a stretch in the y' direction. This is a simpler description of the deformation and clearer physically than the description given by (4.1).

You can see now why the order of eigenvalues down the main diagonal in D is arbitrary and why (4.12) is just as satisfactory as (4.7). The new axes (x', y') are along the eigenvectors, but it is unimportant which eigenvector we call x' and which we call y'. In doing a problem we simply select a D with the eigenvalues of M in some (arbitrary) order down the main diagonal. Our choice of D then determines which eigenvector direction is called the x' axis and which is called y'.