GILBERT DE B. ROBINSON

VECTOR GEOMETRY

Vector
Geomentry

Gilbert de B. Robinson

TO JOAN, NANCY, AND JOHN

PREFACE

Even though some books need none, it has become conventional to write a preface. Many people have to be thanked for thier assistance in preparing the manuscript or in reading the proof—but such prefaces need not be read! Another kind of preface, however, which is sometimes not written, should be read, since it explains the sort of background which is assumed and for whom the book is intended.

BACKGROUND. No specific assumptions are made, but a student should have had a preliminary course in synthetic and also in analytical plane geometry. Permutations and combinations will come to the fore in Chapters 2 and 4, and a general feeling for algebraic processes is important throught.

A course such as the one presented here is preliminary in Toronto to several more detailed and systematic courses in algebra and in geometry for those students who specialize in mathematics. Students who specialize in physics or in chemistry, however, may not meet these ideas again until they are brought face to face with their applications, and in such a context the practical aspects of the problem are all-important. Although a knowledge of the calculus is desirable, as the Appendix makes clear, it is not essential for understanding the ideas described here.

AIMS OF THE COURSE. This then was the problem—to give an introductory course in modern *algebra* and *geomety*—and I have proceeded on the assumption that neither is complete without the other, that they are truly two sides of the same coin.

In seeking to coordinate Euclidean, projective, and non-Euclidean geometry in an elementary way with matrices, determinants, and linear transformations, the notion of a *vector* has been exploited to the full. There is nothing new in this book, but an attempt has been made to present ideas at a level suitable to first-year students and in a manner to arouse their interest. For these associations of ideas are the stuff from which modern mathematics and many of its applications are made.

The course has been given for three successive years, and my thanks are due to three successive classes of mathematics, physics, and chemistry students who have helped me to coordinate my ideas concerning the appropriate material and the order of its presentation. Neither of these factors need be fixed and additions or alterations can easily be made, but the underlying pattern of a linear transformation and its geometrical interpretation in different contexts remains the thread which connects the different topics. The brief introduction of a quadratic transformation in Chapter 8 only serves to emphasize the pattern!

A WORD TO STUDENTS. I have tried to keep the presentation as informal as possible in an attempt to arouse and maintain interest. Some of your established ideas may be challenged in Chapter 8 but this is all part of the process! The exercises have been constructed to illustrate the subject in hand and sometimes to carry the ideas a little further, but emphasis by mere repetition has been avoided. This matter of exercises is important. You should work at them contemplatively and expect to be frustrated sometimes, for this is the only way to make the ideas your own.

G. DE B. ROBINSON
University of Toronto

REFERENCES

The number of books on algebra and geometry is increasing every day, but the following list provides a reasonably diversified selection to which the reader can turn for further material.

WITH EMPHASIS ON ALGEBRA

1. Birkhoff, G., and S. MacLane, *Survey of Modern Algebra* (New York: The Macmillan Co., 1953).
2. Jaeger, A., *Introduction to Analytic Geometry and Linear Algebra* (New York: Holt, Rinehart and Winston, Inc., 1960).
3. McCoy, N. H., *Introduction to Modern Algebra* (Boston: Allyn and Bacon, Inc., 1960).
4. Murdoch, D. C., *Linear Algebra for Undergraduates* (New York: John Wiley & Sons, Inc., 1957).
5. Schreier, O., and E. Sperner, *Introduction to Modern Algebra and Matrix Theory* (New York: Chelsea Publishing Co., 1959).
6. Todhunter, I., and J. G. Leatham, *Spherical Trigonometry* (London: The Macmillan Co., 1919).
7. Thrall, R. M., and L. Tornheim, *Vector Spaces and Matrices* (New York: John Wiley & Sons, Inc., 1957).
8. Turnbull, H. W., and A. C. Aitken, *An Introduction to the Theory of Canonical Matrices* (Glasgow: Blackie and Son, Ltd., 1932).

WITH EMPHASIS ON FOUNDATIONS

1. Bonola, R., *Non-Euclidean Geometry* (New York: Dover Publications, Inc., 1955).
2. Klein, F., *Vorlesungen uber Nicht-Euklidische Geometrie* (Berlin: Springer-Verlag, 1928).
3. ____, *Elementary Mathematics from an Advanced Standpoint*, Vol. 2 (New York: The Macmillan Co., 1939).
4. Robinson, G. de B., *Foundations of Geometry* (Toronto: University of Toronto Press, 1959).

WITH EMPHASIS ON GEOMETRY

1. Askwith, E. H., *Pure Geometry* (Cambridge: Cambridge University Press, 1937).

2. Bell, R. J. T., *Coordinate Geometry of Three Dimensions* (London: The Macmillan Co., 1956).
3. Coolidge, J. L., *A Treatise on Algebraic Plane Curves* (Oxford: The Clarendon Press, 1931).
4. Coxeter, H. S. M., *The Real Projective Plane* (New York: McGraw-Hill Book Co., 1949).
5. ____, *Introduction to Geometry* (New York: John Wiley & Sons, Inc., 1961).
6. Schreier, O., and E. Sperner, *Projective Geometry* (New York: Chelsea Publishing Co., 1961).
7. Sommerville, D. M. Y., *Analytical Geometry of Three Dimensions* (Cambridge: Cambridge University Press, 1934).
8. ____, *An Introduction to the Geometry of n Dimensions* (New York: Dover Publications, Inc., 1958).

CONTENTS

VECTOR GEOMETRY

1

LINES AND PLANES

1.1 COORDINATE GEOMETRY

The study of *geometry* is essentially the study of relations which are suggested by the world in which we live. Of course our environment suggests many relations, physical, chemical and psychological, but those which concern us here have to do with *relative positions in space* and with *distances*. We shall begin with Euclidean geometry, which is based on *Pythagoras' theorem:*

The square on the hypotenuse of a right-angled triangle is equal to the sum of the squares on the other two sides.

The statement of this fundamental result implies a knowledge of *length* and *area* as well as the notion of a *right angle*. If we know what we mean by length and may assume its invariance under what we call "motion," we can construct a right angle using a ruler and compass. We define the area of a rectangle as the product of its length and breadth. To be rigorous in these things is not desirable at this stage, but later on we shall consider a proper set of axioms for geometry.

While the Greeks did not explicitly introduce *coordinates*, it is hard to believe that they did not envisage their usefulness. The utilization of coordinates was the great contribution of Descartes in 1637, and to us now it is a most natural procedure. Take an arbitrary point O in space, the corner of the room, for instance, and three mutually perpendicular *coordinate axes*. These lines could be the three lines of intersection of the "walls" and the "floor" at O; the planes so defined we call the *coordinate planes*. In order to describe the position of a point X, we measure its perpendicular distances from each of these three planes, denoting the distances $x1, x2, x3$ as in Figure 1.1 It is important to distinguish *direction* in making these measurements. Any point within the "room" has all its *coordinates* $(x1, x2, x3)$ positive; measurements on the opposite side of any coordinate plane would be negative. Thus the following eight combinations of sign describe the eight octants of space about O:

$$+ + +, \quad + + -, \quad + - +, \quad + - -$$
$$- + +, \quad - + -, \quad - - +, \quad - - -$$
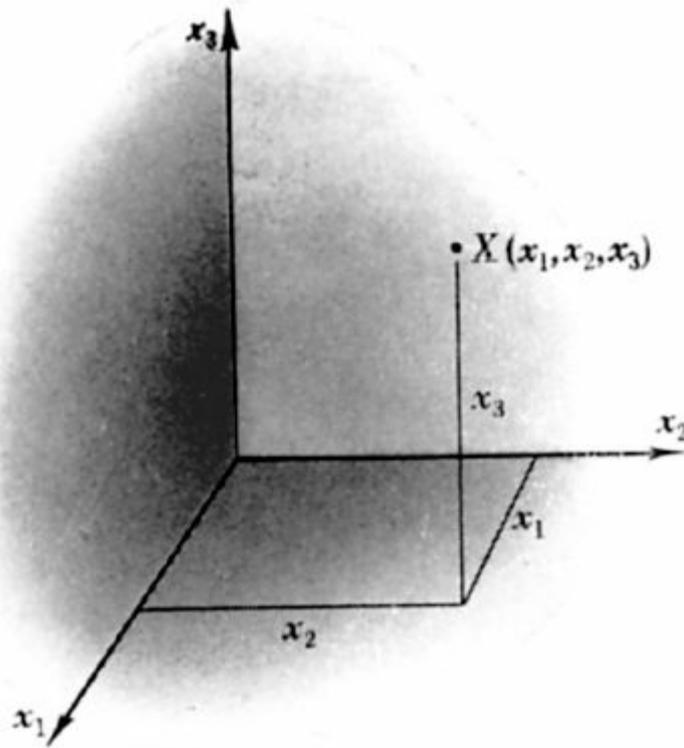
FIG. 1.1

We may describe the points on the "floor" by saying that $x3 = 0$; this is the *equation* of this coordinate plane. Limiting our attention to such points, we have plane geometry. If we call the number of mutually perpendicular coordinate axes the *dimension* of a space, then a plane has two *dimensions* and the position of each point is given by two coordinates, while *space* as we have been describing it has three dimensions.

1.2 EQUATIONS OF A LINE

If we assume that a *line* is determined uniquely by any two of its points, it is natural to seek characterizing properties dependent on these two points only. To this end we refer to Figure 1.2, assuming X to have any position on the line ZY, and complete the rectangular parallelepiped as indicated. If the coordinates of the points in question are

$$X(x_1, x_2, x_3), \quad Y(y_1, y_2, y_3), \quad Z(z_1, z_2, z_3)$$

and if XA, AB, AC are parallel to the coordinate axes with XD parallel to ZP, then from similar triangles,

$$ZX:ZA:ZB:ZC:ZD = ZY:ZP:ZQ:ZR:ZS$$

It follows from this proportionality that if we set $ZX = \tau ZY$, then

$$ZB = \tau ZQ, \quad ZC = \tau ZR, \quad ZD = \tau ZS$$

so that, in terms of coordinates,

1.21 $$\qquad\qquad x_i - z_i = \tau(y_i - z_i) \qquad\qquad i = 1,2,3$$

These equations may be rewritten thus:

1.22
$$\begin{aligned}
x_1 &= z_1 + \tau(y_1 - z_1) = \tau y_1 + (1 - \tau)z_1 \\
x_2 &= z_2 + \tau(y_2 - z_2) = \tau y_2 + (1 - \tau)z_2 \\
x_3 &= z_3 + \tau(y_3 - z_3) = \tau y_3 + (1 - \tau)z_3
\end{aligned}$$

in which form they define the coordinates of X as linear functions of the *parameter* τ. Clearly, if $\tau = 0$ then $X = Z$, and if $\tau = 1$ then $X = Y$.

If we set

$$l_1 = y_1 - z_1, \quad l_2 = y_2 - z_2, \quad l_3 = y_3 - z_3$$

then $l1, l2, l3$ are called the *direction numbers* of the line l. If X and X' are any two distinct points on l, then

$$x_i - x_i' = (\tau - \tau')(y_i - z_i) \qquad\qquad i = 1,2,3$$

so that *numbers proportional to $l1, l2, l3$ are determined by any two distinct points on l.* Two lines whose direction numbers are proportional are said to be *parallel*. We can summarize these results by writing

$$(x_1 - z_1):(x_2 - z_2):(x_3 - z_3) = (y_1 - z_1):(y_2 - z_2):(y_3 - z_3) = l_1:l_2:l_3$$

It follows that we may write the *equations* of l in the symmetric form

1.231 $$\quad \frac{x_1 - z_1}{y_1 - z_1} = \frac{x_2 - z_2}{y_2 - z_2} = \frac{x_3 - z_3}{y_3 - z_3}$$

or

1.232 $$\quad \frac{x_1 - z_1}{l_1} = \frac{x_2 - z_2}{l_2} = \frac{x_3 - z_3}{l_3}$$

but it should be emphasized that these are valid *only if all the denominators are different from zero*, i.e., provided the line l is *not* parallel to one of the coordinate planes. As will appear in the sequel, it is the parametric equations 1.22 which are most significant. Moreover, they generalize easily and provide the important link between classical geometry and modern algebra.
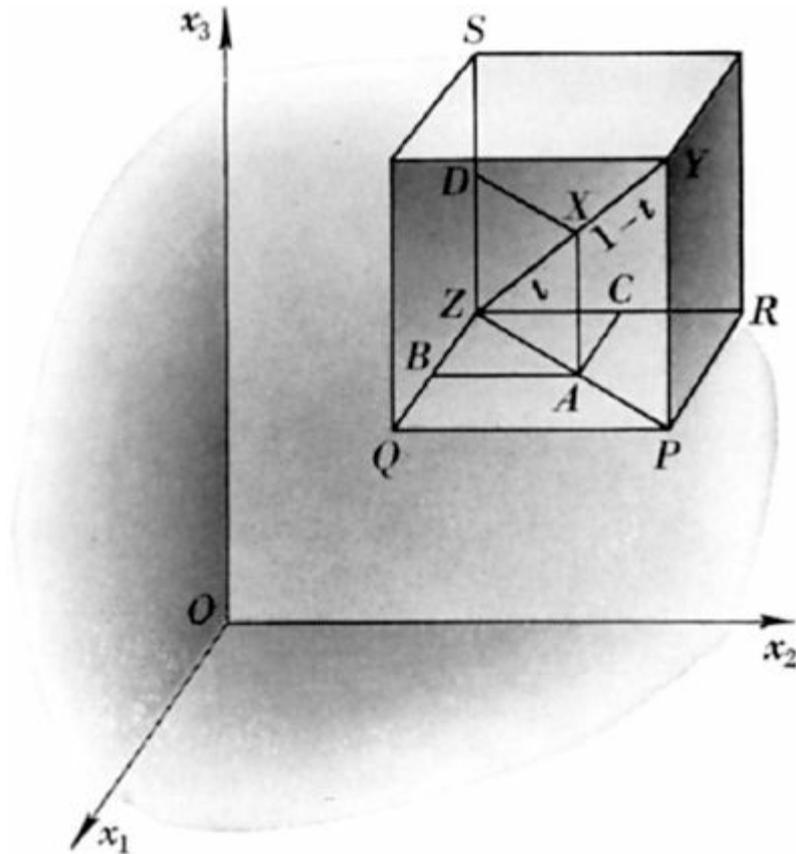
FIG. 1.2

Let us now assume that ZY makes angles $\theta 1$, $\theta 2$, $\theta 3$ with ZQ, ZR, ZS, i.e., with $Ox1$, $Ox2$, $Ox3$. One must be careful here to insist on the direction being from Z to Y; otherwise the angles θi might be confused with $\pi - \theta i$. With such a convention,

1.24

$$\cos \theta_1 = \frac{ZQ}{ZY} = \frac{ZB}{ZX} = \lambda_1$$

$$\cos \theta_2 = \frac{ZR}{ZY} = \frac{ZC}{ZX} = \lambda_2$$

$$\cos \theta_3 = \frac{ZS}{ZY} = \frac{ZD}{ZX} = \lambda_3$$

and $\lambda 1$, $\lambda 2$, $\lambda 3$ are called the *direction cosines* of the line l. By Pythagoras' theorem, $ZP2 = ZQ2 + ZR2$, so that

1.25

$$\left(\frac{ZQ}{ZY}\right)^2 + \left(\frac{ZR}{ZY}\right)^2 + \left(\frac{ZS}{ZY}\right)^2 = \lambda_1^2 + \lambda_2^2 + \lambda_3^2 = 1$$

Thus, given $l1$, $l2$, $l3$, we have

$$\textbf{1.26} \qquad \lambda_1 = \frac{l_1}{\sqrt{l_1^2 + l_2^2 + l_3^2}}, \qquad \lambda_2 = \frac{l_2}{\sqrt{l_1^2 + l_2^2 + l_3^2}}, \qquad \lambda_3 = \frac{l_3}{\sqrt{l_1^2 + l_2^2 + l_3^2}}$$

and parallel lines make equal angles with the coordinate axes.

Clearly, $\lambda 1$, $\lambda 2$, $\lambda 3$ may be substituted for $l1$, $l2$, $l3$ in 1.232, and we may write the first set of equations of 1.22 in the form

$$\textbf{1.27} \qquad \begin{aligned} x_1 &= z_1 + l\lambda_1 \\ x_2 &= z_2 + l\lambda_2 \\ x_3 &= z_3 + l\lambda_3 \end{aligned}$$

EXERCISES

1. Find the equations, in parametric and symmetric form, of the line joining the two points $Y(1, -2, -1)$ and $Z(2, -1, 0)$.

Solution. The parametric equations of the line in question are, by 1.22,

$$\begin{aligned} x_1 &= \tau + 2(1 - \tau) = 2 - \tau \\ x_2 &= -2\tau - 1(1 - \tau) = -1 - \tau \\ x_3 &= -\tau \end{aligned}$$

and in the symmetric form 1.231,

$$\frac{x_1 - 2}{1 - 2} = \frac{x_2 + 1}{-2 + 1} = \frac{x_3}{-1}$$

2. What are the direction cosines of the line in Exercise 1? Write the equations of the line in the form 1.27.

3. Find parametric equations for the line through the point $(1, 0, 0)$ parallel to the line joining the origin to the point $(0, 1, 2)$. Could these equations be written in the form 1.232?

4. Find the equations of the edges of the cube whose vertices are the eight points $(\pm1, \pm1, \pm1)$, as in Figure 5 of Chapter 4.

5. Find the direction cosines of the edges of the regular tetrahedron with vertices

$$A(1, -1, -1), \quad B(-1, 1, -1), \quad C(-1, -1, 1), \quad D(1, 1, 1)$$

1.3 VECTOR ADDITION

The notion of a vector in three dimensions, or 3-space, can be introduced in two ways:

(i) A *vector* is a directed line segment of fixed length.

(ii) A *vector X* is an ordered* triple of three numbers $(x1, x2, x3)$, called the *components* of *X*.

It is important to have both definitions clearly in mind. If we write a small arrow above the symbols to indicate *direction*, then this is determined for $V = \overrightarrow{ZY}$ in Figure 2 by the components

$$\overrightarrow{(y_1 - z_1, \quad y_2 - z_2, \quad y_3 - z_3)} = (v_1, v_2, v_3)$$

also, the length of *ZY* or the *magnitude* of *V* is given by

$$1.31 \qquad |V| = \sqrt{(y_1 - z_1)^2 + (y_2 - z_2)^2 + (y_3 - z_3)^2} = \sqrt{v_1^2 + v_2^2 + v_3^2}$$

The *position* of a vector is immaterial, so we may assume it to have one end "tied" to the origin. Sometimes a vector is called "free" if it can take up any position, but this distinction is not made in either (i) or (ii). In this sense a vector is more general than any particular directed segment, and could be described as an *equivalence class*† of directed segments.

That the two definitions (i) and (ii) are equivalent follows from the theorem:

.32 *Two vectors are equal if and only if their components are equal.*
Proof. Since the components $(v1, v2, v3)$ determine the magnitude and direction of a vector *V*, the condition is certainly sufficient. Conversely, if $|U| = |V|$ then

$$u_1^2 + u_2^2 + u_3^2 = v_1^2 + v_2^2 + v_3^2$$

and if *U* and *V* have the same direction, we must have

$$u_1 = kv_1, \quad u_2 = kv_2, \quad u_3 = kv_3$$

so that $k2 = 1$. It follows that $k = 1$ and the two vectors must coincide.

Following this line of thought, we denote the vector with components $(ku1, ku2, ku3)$ by kU so that

$$|kU| = |k||U|$$

k may be any real number, and $|k|$ is k taken positive. In particular, k may be zero, in which case kU is the *zero vector 0* with components $(0, 0, 0)$. Evidently the magnitude of *0* is zero and its direction is undefined.

We define the *sum*

$$W = U + V$$

of two vectors U and V to be the diagonal of the parallelogram formed by U and V. Alternatively, we may define W by means of the formulas

1.33 $w_i = u_i + v_i$ $i = 1,2,3$

It will be sufficient to consider these definitions in the plane where a vector is defined by two components only. We take the vectors $U(u1, u2)$ and $V(v1, v2)$ and complete the parallelogram, as in Figure 1.3; it follows immediately that the components of W satisfy the relation 1.33. But there is more to be learned from the figure. For example, we arrive at the same result whether we go one way around the parallelogram or the other way around, so that

$$U + V = V + U$$
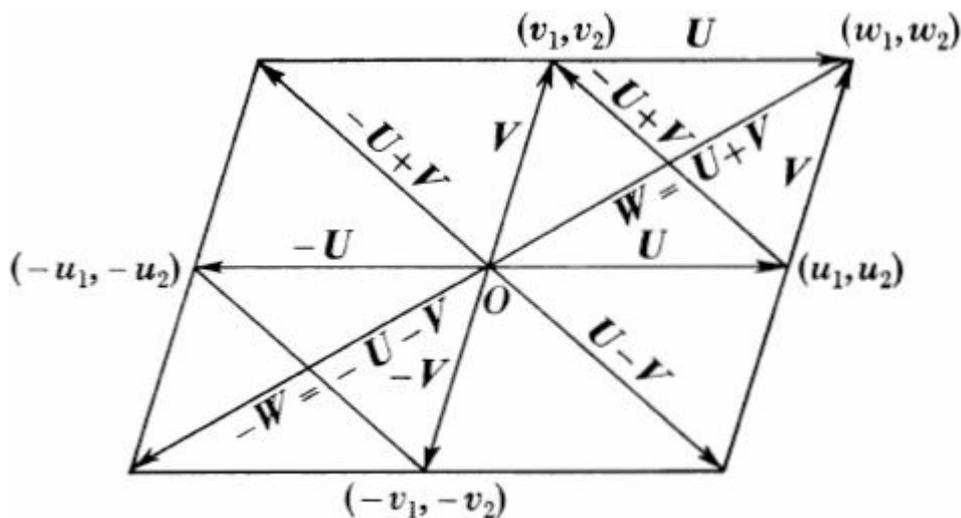
FIG. 1.3

and vector addition is *commutative*. This is also a consequence of the commutativity of addition as applied to the components in 1.33. Finally, by reversing the direction of U we obtain the vector $-U$ so that

$$-U + U = 0 = U - U$$

where 0 is the zero vector. The other diagonal of the parallelogram is the vector $-U + V$, as indicated.

Consider now the similarity between the formulas 1.33 defining vector addition and the parametric equations of a line in 1.27. If we denote by Z the vector $\overrightarrow{OZ}$ with components (z1, z2, z3) and by Λ the vector with components (λ1, λ2, λ3), then *the relations 1.27 are the scalar equations equivalent to the vector equation*

1.34 $$X = Z + t\Lambda$$

It follows that the notion of a *vector* is of central significance in Euclidean geometry. As the title of this book suggests, our purpose is to develop these ideas in several different contexts. Some of these contexts are officially "algebraic" while others are "geometric," but with this thread to guide us, we shall see their interrelations and why it is that mathematics is a living subject, changing and progressing with the introduction of new ideas.
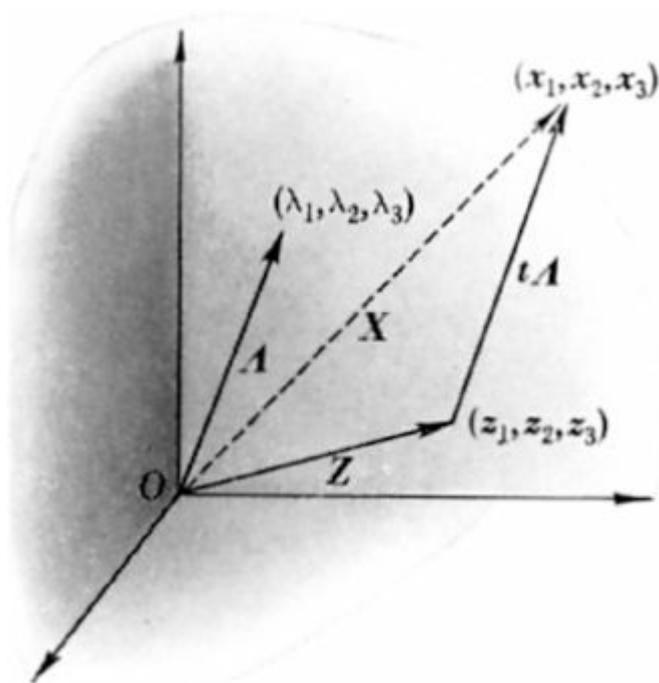
FIG. 1.4

EXERCISES

1. Show that the following vectors are equal: $\overrightarrow{OA}, \overrightarrow{AP}, \overrightarrow{QR}$ where O is the origin, A is the point $(2, -3, 1)$, P is the point $(4, -6, 2)$, Q is the point $(-7, 3, 1)$, and R is the point $(-5, 0, 2)$.

2. Find the components of the vectors $\overrightarrow{AB}, \overrightarrow{BC}, \overrightarrow{CA}$ where A is the point $(1, 2, 3)$, B is the point $(-2, 3, 1)$, and C is the point $(3, -2, -4)$, and show that

$$\overrightarrow{AB} + \overrightarrow{BC} + \overrightarrow{CA} = 0$$

3. Determine the length and the direction cosines of the vector $\overrightarrow{AB}$ in Exercise 2. What would be the components of a parallel vector of unit length?

4. If U, V, W are three arbitrary vectors, show that

$$(U + V) + W = U + (V + W)$$

(the associative law of addition).

5. Prove that the medians of any triangle ABC are concurrent.

Solution. If D is the midpoint of *BC*, then $\overrightarrow{OD} = \frac{1}{2}(\overrightarrow{OB} + \overrightarrow{OC})$. Since the centroid *G* divides *AD* in the ratio 2:1,

$$\overrightarrow{OG} = \tfrac{1}{3}(\overrightarrow{OA} + 2\overrightarrow{OD}) = \tfrac{1}{3}(\overrightarrow{OA} + \overrightarrow{OB} + \overrightarrow{OC})$$

Since this result is symmetric in *A*, *B*, *C*, the medians must be concurrent in *G*.

1.4 THE INNER PRODUCT

In the preceding section we defined the multiplication of a vector U by a scalar k. Such multiplication is called *scalar multiplication* and it is obviously commutative,

1.41 $$kU = Uk$$

There is another kind of multiplication of vectors which is of great importance. To define it we use the generalized Pythagorean theorem to yield

1.42 $$|\overrightarrow{YZ}|^2 = |\overrightarrow{XY}|^2 + |\overrightarrow{XZ}|^2$$
$$- 2|\overrightarrow{XY}||\overrightarrow{XZ}| \cos \theta$$

so that

1.43 $$\cos \theta = \frac{|\overrightarrow{XY}|^2 + |\overrightarrow{XZ}|^2 - |\overrightarrow{YZ}|^2}{2|\overrightarrow{XY}||\overrightarrow{XZ}|}$$
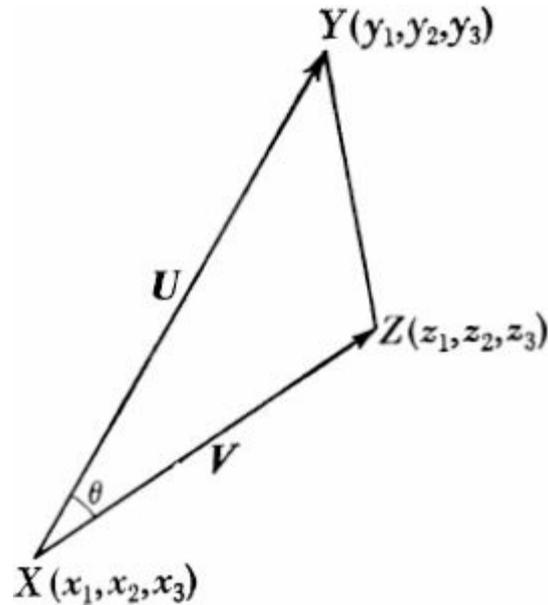
FIG. 1.5

Substituting from 1.31 and simplifying, we have

$$1.44 \qquad \cos \theta = \frac{(y_1 - x_1)(z_1 - x_1) + (y_2 - x_2)(z_2 - x_2) + (y_3 - x_3)(z_3 - x_3)}{|\overrightarrow{XY}||\overrightarrow{XZ}|}$$

$$= \cos \varphi_1 \cos \psi_1 + \cos \varphi_2 \cos \psi_2 + \cos \varphi_3 \cos \psi_3$$

where $\cos \varphi 1$, $\cos \varphi 2$, $\cos \varphi 3$ are the direction cosines of XY, and $\cos \psi 1$, $\cos \psi 2$, $\cos \psi 3$ are those of XZ. Since it is important to have a convenient expression for the sum of products appearing in 1.44, we define the *inner* or *scalar product* of the vectors U, V to be

$$1.45 \qquad\qquad U \cdot V = u_1 v_1 + u_2 v_2 + u_3 v_3 = |U||V| \cos \theta$$

where U has components $(u1, u2, u3)$ and V as components $(v1, v2, v3)$.

All these formulas are valid also in the plane, but in this case a vector U has only two components $(u1, u2)$, and the angles $\varphi 1$, $\varphi 2$ between U and the coordinate axes are complementary. Thus $\cos \varphi 2 = \sin \varphi 1$, so that

$$\cos \theta = \cos \varphi_1 \cos \psi_1 + \sin \varphi_1 \sin \psi_1 = \cos (\varphi_1 - \psi_1)$$

and it is convenient to write the equation of a line (note that there is now only *one* equation),

$$\frac{x_1 - z_1}{\cos \theta_1} = \frac{x_2 - z_2}{\cos \theta_2}$$

in the form

$$x_2 - z_2 = (\tan \theta_1)(x_1 - z_1)$$

Rather than try to visualize a *space* of more than three dimensions, one should think of a vector V as having n components $(v_1, v_2, \ldots, v_n)$. The sum of two vectors, $W = U + V$, is defined by the equations

$$w_i = u_i + v_i \qquad\qquad i = 1, 2, \ldots n$$

and the vector kV has components $(kv_1, kv_2, \ldots kv_n)$ for any real number k.

Through use of Pythagoras' theorem, the distance between two points $Z(z_1, z_2, \ldots z_n)$ and $Y(y_1, y_2, \ldots y_n)$ is given by the relation

1.46 $$|ZY|^2 = (y_1 - z_1)^2 + (y_2 - z_2)^2 + \ldots + (y_n - z_n)^2$$

All that we have said generalizes so that

$$\cos \theta = \cos \varphi_1 \cos \psi_1 + \cos \varphi_2 \cos \psi_2 + \ldots + \cos \varphi_n \cos \psi_n$$

and

1.47 $$U \cdot V = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n = |U||V| \cos \theta$$

In Particular,

$$U \cdot U = |U|^2 = u_1^2 + u_2^2 + \ldots + u_n^2$$

and such a space is still called *Euclidean*, of n dimensions.

Finally, inner multiplication is *commutative*, and since

$$(U + V) \cdot W = U \cdot W + V \cdot W$$

it is also *distributive*. In such a relation it is not necessary to refer to the dimensionality of the space in which the vectors lie. We consider this "abstract" approach to vectors in the following section.

EXERCISES

1. Prove that if the vectors U and V are perpendicular, then $U \cdot V = 0$, and conversely.

2. Prove that each face of the regular tetrahedron in Exercise 5 of is an equilateral triangle, (a) by finding the lengths of the edges and (b) by finding the angles between the edges.

3. Show that if W is perpendicular to U and also to V, then W is perpendicular

to any vector $aU + bV$. How would such a vector $aU + bV$ be related to U and V? Draw a figure to illustrate the following solution.

Solution. If $W \cdot U = 0$ and $W \cdot V = 0$, then

$$W \cdot (aU + bV) = aW \cdot U + bW \cdot V = 0$$

The vector $aU + bV$ would be obtained by first constructing aU collinear with U and bV collinear with V and then finding the diagonal of the parallelogram formed by aU and bV.

4. If we denote the vector $\overrightarrow{ZY}$ in Figure 5 by W, then $W = U - V$. Derive the relation 1.42 by calculating the inner product $W \cdot W$.

5. Prove the two following inequalities:

$$|U + V| \leqslant |U| + |V| \qquad \text{(triangle inequality)}$$
$$U \cdot V \leqslant |U||V| \qquad \text{(Schwarz's inequality)}$$

1.5 LINEAR DEPENDENCE

As the simplest example of this important concept, let us consider a space of n dimensions and points

$$E_1 = (1,0,0, \ldots, 0), \quad E_2 = (0,1,0, \ldots, 0), \ldots, \quad E_n = (0,0,0, \ldots, 1)$$

one on each coordinate axis. If we denote the vector $\overrightarrow{OE}i$ by Ei, then it is an easy extension of the ideas of the preceding section to write any vector $X(x1, x2, \ldots xn)$ in the form

.51
$$X = x_1 E_1 + x_2 E_2 + \ldots + x_n E_n$$

The vector X is said to be *linearly dependent* on the *basis* vectors Ei ($i = 1, 2, \ldots n$).

More generally, we shall say that vectors U, V, $\ldots W$ are *linearly dependent* if there exists a set of constants a, b, $\ldots c$, not all zero, such that

.52
$$aU + bV + \ldots + cW = 0$$

If no such constants exist then the vectors U, V, $\ldots W$ are said to be *linearly independent*.

As we have mentioned before, a vector equation 1.51 or 1.52 is equivalent to, or is a short-hand way of writing, a set of n scalar equations. For example, the scalar equations corresponding to 1.52 are

$$x_1 = x_1 \cdot 1 + x_2 \cdot 0 + \ldots + x_n \cdot 0$$
$$x_2 = x_1 \cdot 0 + x_2 \cdot 1 + \ldots + x_n \cdot 0$$
$$\vdots$$
$$x_n = x_1 \cdot 0 + x_2 \cdot 0 + \ldots + x_n \cdot 1$$

while those corresponding to 1.52 are

$$au_i + bv_i + \ldots + cw_i = 0 \qquad\qquad i = 1, 2, \ldots n$$

We shall develop the notion of a *basis* in subsequent chapters, but the vectors E_i are particularly important; not only are they pairwise *orthogonal*, i.e., perpendicular, but they are also *normal*, i.e., of unit length. We express both these facts by writing

1.53
$$\boldsymbol{E}_i \cdot \boldsymbol{E}_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

In 3-space it is sometimes convenient to use the notation $x1 = x$, $x2 = y$, $x3 = z$, in which case we write $E1 = i$, $E2 = j$, $E3 = k$. The advantage of the suffix notation, however, is that it extends to any number of dimensions.

It is interesting to see that we could have approached our subject from a purely abstract point of view, defining an *abstract vector space* υ as a set of vectors A, B, C, … with the property of addition such that:

(i) If A and B are vectors in υ so also is $A + B$

(ii) $A + B = B + A$ (commutative law of addition)

(iii) $(A + B) + C = A + (B + C)$ (associative law of addition)

(iv) There exists a vector in υ called the *zero* vector $\boldsymbol{0}$ such that $A + \boldsymbol{0} = A = \boldsymbol{0} + A$

(v) With every vector A in υ is associated a vector $-A$ such that

$$A + (-A) = \boldsymbol{0} = (-A) + A$$

(vi) With every real number k and vector A in υ is associated a vector $kA = Ak$ such that $k1A + k2A = (k1 + k2)A$ and $k1(k2A) = k1k2A$. We assume that $1A = A$ for all A.

.54 Definition The number of linearly independent vectors in υ is called the *dimension* of υ.

We may introduce the notion of an inner product $A \cdot B$ by assuming this operation to satisfy the further axioms

(vii) $A \cdot B = B \cdot A$ (commutative law of inner multiplication)

(viii) $A \cdot (B + C) = A \cdot B + A \cdot C$ (distributive law)

(ix) $(kA \cdot B) = (A \cdot kB) = k(A \cdot B)$

(x) For any vector A in υ, $A \cdot A$ is a real positive number or zero.

(xi) $A \cdot A = 0$ implies that $A = 0$.

Thus we may set $|A|2 = A \cdot A$ and call $|A| \geqslant 0$ the *magnitude* of A. Similarly, for any two vectors A, B we may define

$$\cos \theta = \frac{A \cdot B}{|A||B|}$$

thus avoiding the use of components at all. However, the geometrical definitions given in Section 3 provide the most familiar realization of a vector space υ and the only one with which we shall be concerned.

EXERCISES

1. If A and B have coordinates $(-1,2,0)$ and $(2,1,-1)$ respectively, express the vector $\overrightarrow{AB}$ in terms of the basis vectors i, j, k.

2. Find the lengths of the two diagonals of the parallelogram formed by the vectors $\overrightarrow{OA}$ and $\overrightarrow{OB}$ in Exercise 1.

3. Prove that the two vectors $U = i - 8j + 2k$ and $V = 6i + 2j + 5k$ are orthogonal, and find $|U|$ and $|V|$.

4. Find X perpendicular to $3i - j + 2k$ and $2i + 5j + 7k$, and such that $|X| = 2$.

5. In the regular tetrahedron with vertices $A(1,-1,-1)$, $B(-1,1,-1)$, $C(-1,-1,1)$, $D(1,1,1)$, prove that the vectors $\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AD}$ are linearly independent.

6. Express the vector $\overrightarrow{BC}$ in Exercise 5 as a linear combination of the vectors $\overrightarrow{AB}$ and $\overrightarrow{AC}$. Express each of these vectors in terms of the basis vectors i, j, k, and show that the same relation holds.

Solution. The components of the vectors in question are

$$\overrightarrow{AB}(-2,2,0), \quad \overrightarrow{BC}(0,-2,2), \quad \overrightarrow{AC}(-2,0,2)$$

so that $\overrightarrow{BC} = \overrightarrow{AC} - \overrightarrow{AB}$. In terms of i, j, k,

$$\overrightarrow{AB} = -2i + 2j, \quad \overrightarrow{BC} = -2j + 2k, \quad \overrightarrow{AC} = -2i + 2k$$

and clearly the same relation holds good.

7. If the centroid G of the tetrahedron in Exercise 5 is defined by the equation

$$\overrightarrow{OG} = \tfrac{1}{4}(\overrightarrow{OA} + \overrightarrow{OB} + \overrightarrow{OC} + \overrightarrow{OD})$$

find the coordinates of G.

8. Prove that the lines joining the midpoints of opposite edges of any tetrahedron concur in the centroid of the tetrahedron.

9. Prove that the lines joining the vertices to the centroids of opposite faces concur in the centroid of the tetrahedron.

1.6 EQUATIONS OF A PLANE

Let us begin by writing the parametric equations 1.21 or 1.22 of a line in the vector form

1.61 $$(X - Z) - \tau(Y - Z) = 0$$

or

1.62 $$X = \tau Y + (1 - \tau)Z$$

.63 *The vectors* $\mathbf{X} - \mathbf{Z}$ *and* $\mathbf{Y} - \mathbf{Z}$ *are linearly dependent if and only if the points X, Y, Z are collinear.*

   If now X, Y, Z are *not* collinear, they will define a plane π; and if $\mathbf{U}$ is any point in π we may complete the parallelogram as in Figure 1.6 and write
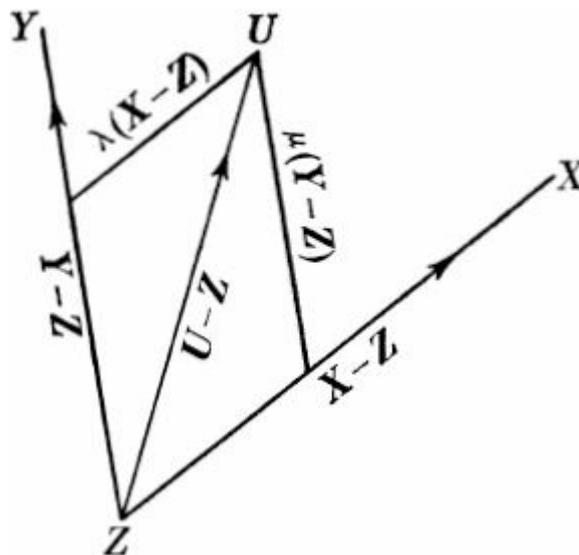
FIG. 1.6

1.64 $$(\mathbf{U} - \mathbf{Z}) - \lambda(\mathbf{X} - \mathbf{Z}) - \mu(\mathbf{Y} - \mathbf{Z}) = 0$$

or

1.65
$$U = \lambda X + \mu Y + (1 - \lambda - \mu) Z$$

Essentially, *we have established a coordinate system in* π *with origin Z and axes ZX and ZY.* The equation of *ZX* is μ = 0 while that of *ZY* is λ = 0, and the *coordinates* of *U* are (λ, μ). Note that these axes *ZX* and *ZY* need not be orthogonal; all that we do require is that λ, μ be determined by lines *parallel* to *ZX* and *ZY.* Since all these steps are reversible, we conclude that

.66 *The vectors* **U** − **Z**, **X** − **Z**, *and* **Y** − **Z** *are linearly dependent if and only if the points U, X, Y, Z are coplanar.*

However, we can approach the problem from quite a different point of view. Let us assume that π passes through the point $Z(z1, z2, z3)$ and is perpendicular to a given vector **U**. Then if *X* is any point of π, the vector **V** = $\overrightarrow{ZX}$ is perpendicular to **U** so that **U**·**V** = 0, or

1.67
$$u_1(x_1 - z_1) + u_2(x_2 - z_2) + u_3(x_3 - z_3) = 0$$

and this is the *equation* of π. Note that the normal vector **U** is not unique since 1.67 may be multiplied through by any constant $k \neq 0$.

We conclude this discussion of planes in space by computing the perpendicular distance p from a point $X(x1, x2, x3)$ of general position to the plane π with equation 1.67. If this distance is measured along a normal $\overrightarrow{XN}$ making an angle θ with $\overrightarrow{XZ}$, then

$$p = |\overrightarrow{XZ}| \cos \theta$$

$$= \frac{|\overrightarrow{XZ}||U \cdot (X - Z)|}{|\overrightarrow{XZ}||U|} \qquad \text{by 1.45}$$

$$= \frac{|u_1(x_1 - z_1) + u_2(x_2 - z_2) + u_3(x_3 - z_3)|}{\sqrt{u_1^2 + u_2^2 + u_3^2}}$$

If the equation 1.67 had been simplified and written in the form

$$u_1 x_1 + u_2 x_2 + u_3 x_3 + u_4 = 0$$

then we would have

1.68 $$p = \frac{|u_1 x_1 + u_2 x_2 + u_3 x_3 + u_4|}{\sqrt{u_1^2 + u_2^2 + u_3^2}}$$

The important question: *What locus is represented by the general linear equation with real coefficients?* can be answered by first finding a point $Z(z1, z2, z3)$ whose coordinates satisfy the equation, and then rearranging it in the form 1.67. We conclude from this and Section 1.2 that:
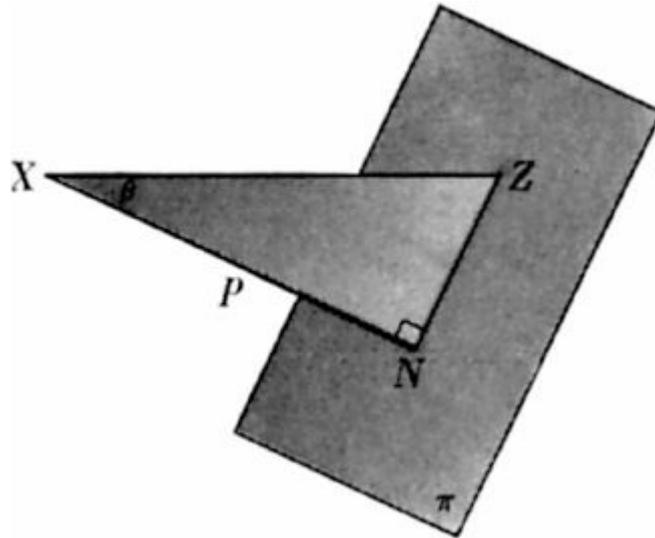
FIG. 1.7

.69 *Every linear equation in $x1, x2, x3$ represents a plane; two linear equations represent the line of intersection of the two planes, unless the two planes are parallel.*

The determination of the parametric or symmetric form of the equations of a line l, when l is defined by two linear equations, is illustrated by the following example.

Example. Consider the intersection l of the two planes

$$x_1 - x_2 + 2x_3 = 1$$
$$x_1 + x_2 - x_3 = 3$$

In order to find a point on l we first look for the intersection of l with, say, the plane $x3 = 0$, which yields the point $Z(2, 1, 0)$. The direction numbers $l1, l2, l3$ of l must satisfy the two linear equations

$$l_1 - l_2 + 2l_3 = 0$$
$$l_1 + l_2 - l_3 = 0$$

which express the fact that l is orthogonal to the normal direction of each plane containing it. Thus

$$l_1 : l_2 : l_3 = -1 : 3 : 2$$

so that the parametric equations of l may be written

$$x_1 = 2 - t, \quad x_2 = 1 + 3t, \quad x_3 = 2t$$

and the symmetric equations

$$\frac{x_1 - 2}{-1} = \frac{x_2 - 1}{3} = \frac{x_3}{2}$$

The direction cosines of l are $-1/\sqrt{14}, \; 3/\sqrt{14}, \; 2/\sqrt{14}$.

EXERCISES

1. Find the equations of the faces of the tetrahedron with vertices

$$A(1,1,1), \quad B(1,-1,-1), \quad C(-1,1,-1), \quad D(-1,-1,1)$$

and their angles of intersection.

Solution. The equation of the face ABC may be taken to be

$$u_1(x_1 - 1) + u_2(x_2 - 1) + u_3(x_3 - 1) = 0$$

Substituting the coordinates of B and C,

$$
\begin{aligned}
-2u_2 - 2u_3 &= 0 \\
-2u_1 \qquad\; - 2u_3 &= 0
\end{aligned}
$$

so that the equation of ABC is

$$(x_1 - 1) + (x_2 - 1) - (x_3 - 1) = 0$$

Similarly, the equation of the face ABD is

$$(x_1 - 1) - (x_2 - 1) + (x_3 - 1) = 0, \text{ etc.}$$

The angle θ between the faces will be the angle between their normals, properly directed, so that

$$\cos \theta = \frac{-1.1 + 1.1 + 1.1}{3} = \frac{1}{3}$$

2. Find the equations of the faces of the octahedron with vertices

$$
\begin{aligned}
A(1,0,0), &\quad B(0,1,0), &\quad C(0,0,1) \\
A'(-1,0,0), &\quad B'(0,-1,0), &\quad C'(0,0,-1)
\end{aligned}
$$

and the angles between faces which (a) intersect in an edge, (b) intersect in a vertex. Which faces are parallel to one another?

3. (a) Give the components of vectors parallel to each of the edges of the octahedron in Exercise 2.

(b) Express each of these vectors in terms of the basis vectors E_1, E_2, E_3.

4. Show that the plane $u_1 x_1 + u_2 x_2 + u_3 x_3 + u_4 = 0$

(i) meets Ox_i at a distance $-u_4/u_i (i = 1, 2, 3)$ from 0

(ii) is parallel to Ox_i if $u_i = 0$

(iii) is perpendicular to Ox_i if $u_j = u_k = 0 (j \neq k \neq i)$

Generalize these statements to an Euclidean space of n dimensions.

* 'Ordered' in the sense that the order of the components x_1, x_2, x_3 is important, so that e.g. $(x_1, x_2, x_3) \neq (x_2, x_1, x_3)$.

† Relations which are *reflexive*, *symmetric*, and *transitive* are known as *equivalence relations*, and the sets to which they apply, as *equivalence classes*. For a discussion of these ideas see Birkhoff and MacLane, *Survey of Modern Algebra*.

2

DETERMINANTS AND LINEAR EQUATIONS

2.1 THE PROBLEM DEFINED

By introducing coordinates, Descartes aimed to make it possible to solve geometrical problems "analytically." Thus, as we saw in Chapter 1, the study of lines and planes in space is translated into the study of simultaneous linear equations. There are two important aspects of this problem which become confused when the number of variables is small, namely, (a) the finding of *actual solutions* of a given system of simultaneous equations, and (b) the investigation of the *properties* of such solutions in general without explicitly determining of them. Though we shall introduce the abbreviation known as a *determinant* for a "multilinear" polynomial expression and use determinants to solve systems of linear equations, the reader should be warned that the real importance of determinants is theoretical rather than practical.

Let us begin with the simple case

2.111
$$a_{11}x_1 + a_{12}x_2 = a_{10}$$

2.112
$$a_{21}x_1 + a_{22}x_2 = a_{20}$$

If we multiply 2.111 by $a22$ and 2.112 by $-a12$ and add, we have

2.113
$$(a_{11}a_{22} - a_{21}a_{12})x_1 = (a_{10}a_{22} - a_{20}a_{12})$$

Similarly, if we multiply 2.111 by $-a21$ and 2.112 by $a11$ and add, we have

2.114
$$(a_{11}a_{22} - a_{21}a_{12})x_2 = (a_{11}a_{20} - a_{21}a_{10})$$

For convenience in writing the solution we set

2.12
$$a_{11}a_{22} - a_{21}a_{12} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \Delta$$

called a *determinant* of order 2, so that

$$\Delta x_1 = \begin{vmatrix} a_{10} & a_{12} \\ a_{20} & a_{22} \end{vmatrix} = \Delta_1, \qquad \Delta x_2 = \begin{vmatrix} a_{11} & a_{10} \\ a_{21} & a_{20} \end{vmatrix} = \Delta_2$$

Observe that the determinant Δ_1 (Δ_2) is formed by replacing the first (second) column of Δ by the vector (a_{10}, a_{20}). Geometrically, we have found the point of intersection of two coplanar lines when $\Delta \neq 0$. If $\Delta = 0$, the two lines are parallel and so have no point of intersection unless they coincide.

Example. The two equations

$$x_1 - x_2 = 1$$
$$x_1 + x_2 = 3$$

have as solution

$$x_1 = \frac{\begin{vmatrix} 1 & -1 \\ 3 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix}} = \frac{4}{2} = 2, \qquad x_2 = \frac{\begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix}}{\begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix}} = \frac{2}{2} = 1$$

The two equations

$$x_1 - \quad x_2 = 1$$
$$2x_1 - 2x_2 = 1$$

represent parallel lines since $\Delta = 0$; they have no solution and are said to be *inconsistent*.

Let us now seek the solution of the three linear equations

2.131 $\qquad\qquad\qquad a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = a_{10}$

2.132 $\qquad\qquad\qquad a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = a_{20}$

2.133 $\qquad\qquad\qquad a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = a_{30}$

To proceed systematically, we begin by reducing the problem to the case $n = 2$ by eliminating x_3, thus:

2.134 $\quad (a_{11}a_{23} - a_{21}a_{13})x_1 + (a_{12}a_{23} - a_{22}a_{13})x_2 = a_{10}a_{23} - a_{20}a_{13}$

2.135 $\quad (a_{11}a_{33} - a_{31}a_{13})x_1 + (a_{12}a_{33} - a_{32}a_{13})x_2 = a_{10}a_{33} - a_{30}a_{13}$

2.136 $\quad (a_{21}a_{33} - a_{31}a_{23})x_1 + (a_{22}a_{33} - a_{32}a_{23})x_2 = a_{20}a_{33} - a_{30}a_{23}$

If we multiply the equation 2.134 by a_{33}, 2.135 by a_{23}, and subtract, we obtain 2.136 multiplied by a_{13}, but for the sake of symmetry we retain all three equations. If now we eliminate x_2 by multiplying 2.134 by $-a_{32}$, 2.135 by a_{22},

2.136 by $-a12$, and add, we arrive at the equation

$$\begin{aligned}
2.14 \qquad & [a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{21}(a_{12}a_{33} - a_{32}a_{13}) + a_{31}(a_{12}a_{23} - a_{22}a_{13})]x_1 \\
& = a_{10}(a_{22}a_{33} - a_{32}a_{23}) - a_{20}(a_{12}a_{33} - a_{32}a_{13}) + a_{30}(a_{12}a_{23} - a_{22}a_{13})
\end{aligned}$$

By a similar procedure we could obtain equations for $x2$ and $x3$.

Again, let us set

$$a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{21}(a_{12}a_{33} - a_{32}a_{13}) + a_{31}(a_{12}a_{23} - a_{22}a_{13})$$

$$2.15 \qquad = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \Delta$$

so that

$$\Delta x_1 = \Delta_1, \quad \Delta x_2 = \Delta_2, \quad \Delta x_3 = \Delta_3$$

As before, Δi is obtained by replacing the ith column of Δ by the vector ($a10$, $a20$, $a30$).

It is important to observe that *the groups of terms multiplying $a11$, $a21$, $a31$ in 2.15 are just the second-order determinants obtained by crossing out the row and column containing $a11$, $a21$, $a31$ in Δ multiplied by ± 1*. In fact, we could multiply 2.131 by

$$(a_{22}a_{33} - a_{32}a_{23}) = A_{11}$$

2.132 by

$$-(a_{12}a_{33} - a_{32}a_{13}) = A_{21}$$

2.133 by

$$(a_{12}a_{23} - a_{22}a_{13}) = A_{31}$$

and add to obtain 2.14 directly, since

$$2.161 \qquad a_{11}A_{11} + a_{21}A_{21} + a_{31}A_{31} = \Delta$$

$$2.162 \qquad a_{1i}A_{11} + a_{2i}A_{21} + a_{3i}A_{31} = 0 \qquad i = 2,3$$

$$2.163 \qquad a_{10}A_{1j} + a_{20}A_{2j} + a_{30}A_{3j} = \Delta_j \qquad j = 1,2,3$$

The equation 2.161 leads to an *inductive* definition of a determinant of order n which runs as follows:

.17 Definition If $A = (a_{ij})$ is an $n \times n$ array or *matrix*, the *minor* M_{ij} of an element a_{ij} in $|A|$ is the determinant of that $(n-1) \times (n-1)$ matrix obtained from A by striking out the ith row and jth column, and the *cofactor* A_{ij} is defined by the equation

$$2.171 \qquad\qquad A_{ij} = (-1)^{i+j} M_{ij} \qquad\qquad i,j = 1, 2, \ldots n$$

We define the determinant $|A|$ of A by the relation

$$2.172 \qquad\qquad |A| = a_{11} A_{11} + a_{21} A_{21} + \ldots + a_{n1} A_{n1}$$

Since $|A|$ has been defined in 2.12 for $n = 2$ and in 2.15 for $n = 3$, the inductive definition is complete.

We summarize the method developed above, expressing the solution of a system of n linear equations in n unknowns in

.18 Cramer's Rule *The solution of n linear equations,*

$$2.181 \qquad\qquad
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= a_{10} \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= a_{20} \\
&\;\;\vdots \\
a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= a_{n0}
\end{aligned}$$

is given by the formulas

$$2.182 \qquad\qquad \Delta x_1 = \Delta_1, \quad \Delta x_2 = \Delta_2, \ldots, \quad \Delta x_n = \Delta_n$$

where $\Delta = |a_{ij}| \neq 0$, and Δ_i is obtained by replacing the ith column of Δ by the vector

$$(a_{10}, a_{20}, \ldots, a_{n0})$$

As in the case $n = 2$, we may interpret these results geometrically, but we postpone this until we have developed the general properties of determinants suggested above, proving Cramer's rule in Section 2.3.

EXERCISES

1. Evaluate the following determinants by multiplying out all the terms in their "expansions" according to 2.15 above:

$$\begin{vmatrix} 1 & 0 & 1 \\ 0 & 3 & -5 \\ -2 & 1 & 4 \end{vmatrix}, \quad \begin{vmatrix} 5 & -3 & 1 \\ -4 & 1 & -1 \\ 3 & 5 & -1 \end{vmatrix}, \quad \begin{vmatrix} 2 & -3 & 1 \\ -6 & 1 & 5 \\ 3 & 2 & -5 \end{vmatrix}$$

2. Solve the simultaneous equations

$$x_1 + x_3 = 0, \quad 3x_2 - 5x_3 = 1, \quad 2x_1 - x_2 - 4x_3 = 2$$

(a) by elimination and (b) by determinants.

3. Have the equations

$$2x_1 - 3x_2 + x_3 = 0, \quad 6x_1 - x_2 - 5x_3 = 1$$
$$3x_1 + 2x_2 - 5x_3 = 3$$

a solution? If not, show explicitly that they are inconsistent.

2.2 DETERMINANTS

There is another definition of a determinant which is important:

.21 Definition If $(aij) = A$ is an $n \times n$ "array" or *matrix*, then the determinant of A is the expression

2.211
$$|A| = \underset{n!}{\Sigma} \, \text{sgn} \, \pi \cdot a_{\pi(1),1} \cdot a_{\pi(2),2} \cdots \cdots \, a_{\pi(n),n}$$

where $\pi(1), \pi(2), \ldots, \pi(n)$ is a permutation of 1, 2, …, n, sgn $\pi = \pm 1$ according as π is *even* or *odd*, and the summation is over all n! such permutations.

Before we attempt to reconcile these two definitions we must make clear the notion of the *evenness* or *oddness* of a permutation. To this end consider all 3! arrangements of the symbols 1,2,3:

2.22   123,   132,   213,   231,   312,   321

By comparison with the initial arrangement 123, we may construct for each arrangement the *permutation* π which accomplishes the change, by starting with any given symbol a and setting next on the right the symbol b into which a is transformed by π, then setting next to $b = \pi(a)$ the symbol $c = \pi(b)$, and so on. In due course, we return to a, completing the *cycle*. If this exhausts the symbols, we have the desired permutation π. If not, we begin again with another symbol not already considered and so construct all the cycles of π. If we do this for the arrangements of 2.22 we obtain

2.23   (1)(2)(3),   (1)(23),   (12)(3),   (123),   (132),   (13)(2)

or

2.24 I,   (23),   (12),   (123),   (132),   (13)

Cycles of length one are usually omitted and the identity permutation is represented by I.

It is often convenient to write a permutation π in *two-rowed* form, where $\pi(a)$ is placed immediately beneath a. Thus, the permutations 2.24 could also be written

$$\begin{pmatrix} 123 \\ 123 \end{pmatrix}, \quad \begin{pmatrix} 123 \\ 132 \end{pmatrix}, \quad \begin{pmatrix} 123 \\ 213 \end{pmatrix}, \quad \begin{pmatrix} 123 \\ 231 \end{pmatrix}, \quad \begin{pmatrix} 123 \\ 312 \end{pmatrix}, \quad \begin{pmatrix} 123 \\ 321 \end{pmatrix}$$

It is important that the order of writing the various columns of the two-rowed form does not matter, so that permutations can be combined as in the following paragraph.

Observe that the permutations (123) and (132) can be written as a sequence of *transpositions*, i.e., cycles of length 2, in the following manner:

$$(123) = \begin{pmatrix} 123 \\ 132 \\ 231 \end{pmatrix} = (12)(23), \qquad (132) = \begin{pmatrix} 123 \\ 132 \\ 312 \end{pmatrix} = (13)(32)$$

since it is customary to operate or *multiply* permutations from *right* to *left*. It can easily be verified that

$$(123) = (12)(23) = (23)(13) = (23)(13)tt$$

for any transposition t, since $tt = I$. Thus the number of transpositions in terms of which (123) may be expressed seems to preserve its parity, though the transpositions themselves may differ; certainly the number of such transpositions is not unique.

.25 Definition A permutation is said to be *even* or *odd* according as the number of transpositions required to express it is even or odd.

If this definition is to be significant we must prove the following

.26 Theorem *The number of transpositions in terms of which a given permutation π on n symbols may be expressed is always even or always odd.*
Proof. Consider the general case and suppose that the permutation π operates on the *subscripts* of $x1, x2, \ldots, xn$. To prove the theorem we construct the function

$$P = \prod_{i<j} (x_i - x_j)$$

and consider the effect of π on P. If $\pi = (ij)$:

2.261

$$
\begin{aligned}
(x_i - x_j) &\longrightarrow -(x_i - x_j) & & \\
(x_k - x_i) &\longleftrightarrow (x_k - x_j) & &\text{for} \quad k < i < j \\
\left.\begin{aligned}(x_i - x_k) &\longrightarrow -(x_k - x_j)\\ (x_k - x_j) &\longrightarrow -(x_i - x_k)\end{aligned}\right\} & & &\text{for} \quad i < k < j \\
(x_i - x_k) &\longleftrightarrow (x_j - x_k) & &\text{for} \quad i < j < k
\end{aligned}
$$

Since these are the only factors of P affected by $\pi = (ij)$, we conclude that $\pi(P) = -P$ *for any transposition* π.

On the other hand, $\pi(P)$ is well defined for any permutation π, so $\pi(P)$ is either P or $-P$. We conclude that π must always be expressible as a product of an *even* or *odd* number of transpositions, as required.

With these explanations, our definition 2.21 of a determinant is complete, and it agrees with that of the preceding section for $n = 2,3$. (Note that I, (123), (132) are *even* while (12), (13), (23) are *odd* permutations, yielding the signs of 2.15 as written.) In order to identify the two definitions we observe that:

(i) No term in the expansion of $|A|$ in 2.21 has two factors a_{ij} with the same first or second suffixes, so that $|A|$ is *linear* in the elements of any row or column.

(ii) Since the $\pi(1)$, $\pi(2)$, … $\pi(n)$ are just 1,2, … n rearranged, we could equally well suppose the *first* suffixes to be arranged in natural order 1,2, … n and the second suffixes permuted by the inverse permutation $\pi{-1}$ of π; for example, if

$$\pi = (123) = \begin{pmatrix} 123 \\ 231 \end{pmatrix}$$

then turned upside down,

$$\begin{pmatrix} 123 \\ 312 \end{pmatrix} = (132) = \pi^{-1}$$

Clearly $\pi\pi{-1} = I = \pi{-1}\pi$, and $\pi{-1}$ is even or odd when π is even or odd.

From (ii) we conclude that

2.27
$$|A| = |A^t|$$

where A_t is the *transpose* of the matrix A, obtained by writing the ith row of A as the ith column of A_t.

The property (i) of $|A|$ is more subtle. Certainly we can collect together those terms in 2.211 which include $a11$ as a factor; the number of these is $(n-1)!$ and they are just those terms which make up $M11 = A11$ as defined in the preceding section. Similarly, we may collect those terms in 2.211 which include $a21$ as a factor; they will be distinct from those which include $a11$ by (i) and will make up $-M21$. To prove this last statement it is sufficient to observe that, to obtain the π's corresponding to terms containing $a21$ as a factor, we need only operate on those π's corresponding to terms containing $a11$ as a factor by the transposition (12), and this accounts for the minus sign. By 2.171,

$$-M_{21} = A_{21}$$

Again, we collect the terms containing $a31$ as a factor; they are distinct by (i) and make up $M31 = A31$, and so on. We conclude that

2.28
$$|A| = a_{11}A_{11} + a_{21}A_{21} + \ldots + a_{n1}A_{n1}$$

as before.

The argument of the preceding paragraph is quite general, and by introducing an extra transposition we can shift the first column into the second column position or, by a further transposition, into the third column position, and so on, leaving the minors of the elements in the column unchanged. Otherwise described, one can think of moving aij, into the position of $a11$, leaving its minor unchanged. Since this same process can be applied to both the rows of A and those of At, i.e., the columns of A, the number of transpositions is

$$(i - 1) + (j - 1) = i + j - 2$$

and since $(-1)^{i+j} = (-1)^{i+j-2}$, we have explained the significance of 2.171. Moreover, we have generalized 2.28 so that we have the following important result:

2.29
$$|A| = \sum_{i=1}^{n} a_{ik}A_{ik} = \sum_{j=1}^{n} a_{kj}A_{kj} \qquad 1 \leqslant k \leqslant n$$

2.3 EVALUATION OF A DETERMINANT

While the definition 2.21 of $|A|$ is theoretically significant, $n!$ therein increases so rapidly that to evaluate an $n \times n$ determinant by calculating each term in its expansion becomes prohibitive. Nor is the calculation of the $(n-1) \times (n-1)$ minors more feasible, so we must develop a third and more practical method of evaluating $|A|$. To this end we prove a sequence of theorems of disarming simplicity.

.31 Interchanging any two columns (rows) of (A) changes $|A|$ into $-|A|$.
Proof. Since such an interchange corresponds to introducing an extra transposition into each π in 2.211, so changing the sign of each term in the expansion of $|A|$ or $|At| = |A|$, and the statement follows.

.32 If two columns (rows) of A are the same, $|A| = 0$.
Proof. Suppose the ith column of A is equal to the jth column. Introducing the extra transposition (ij) does not affect A, but $|A|$ changes sign. Hence $|A| = -|A| = 0$, and similarly for At.

.33 Multiplication of a column (row) of A by k changes $|A|$ into $k|A|$.
Proof. This follows immediately from 2.211, since every term in the expansion contains just one factor from each row and each column.

.34 Adding a constant multiple of a column (row) of A to another column (row) of A leaves $|A|$ unchanged.
Proof. Let us assume we are adding k times the jth column to the jth column of A, which yields:

$$\begin{vmatrix} a_{11} & .. & a_{1i} & .. & a_{1j} + ka_{1i} & .. & a_{1n} \\ a_{21} & .. & a_{2i} & .. & a_{2j} + ka_{2i} & .. & a_{2n} \\ . \\ . \\ . \\ a_{n1} & .. & a_{ni} & .. & a_{nj} + ka_{ni} & .. & a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & .. & a_{1i} & .. & a_{1j} & .. & a_{1n} \\ a_{21} & .. & a_{2i} & .. & a_{2j} & .. & a_{2n} \\ . \\ . \\ . \\ a_{n1} & .. & a_{ni} & .. & a_{nj} & .. & a_{nn} \end{vmatrix} + k \begin{vmatrix} a_{11} & .. & a_{1i} & .. & a_{1i} & .. & a_{1n} \\ a_{21} & .. & a_{2i} & .. & a_{2i} & .. & a_{2n} \\ . \\ . \\ . \\ a_{n1} & .. & a_{ni} & .. & a_{ni} & .. & a_{nn} \end{vmatrix} = |A| + k \cdot 0 = |A|$$

The splitting into a sum of two determinants follows from 2.21 since each term in the expansion contains just one factor from the jth column and every such element is of the form $arj + kari$. The first determinant on the right is just $|A|$, while the second is zero by 2.32. As usual, the result for rows follows by considering At.

We are now in a position to give the general result of which 2.162 is a special case:

.35 For any $n \times n$ determinant $|A|$,

$$\sum_{k=1}^{n} a_{jk} A_{ik} = \sum_{k=1}^{n} a_{kj} A_{ki} = 0 \qquad i \neq j$$

Proof. Consider a determinant obtained from $|A|$ by replacing the ith column by a replica of the jth column, and denote the result by $|A0|$. Expanding $|A0|$ according to 2.29 we have

$$|A_0| = \sum_{k=1}^{n} a_{jk} A_{ik} = 0$$

by 2.32. The second result of 2.35 follows by considering At.

The application of these results to the evaluation of a determinant is immediate. Our aim is to simplify the expansion by introducing as many strategically placed zeros as possible by successive applications of 2.33 and 2.34. If we can arrange that all but one element in each row and column is zero, then the evaluation will be reduced to a mere multiplication of nonzero elements, after rearrangement according to 2.31. We shall study the stages of this reduction in detail later on with reference to the *matrix A* rather than $|A|$, but we can apply it effectively here.

Example. In order to evaluate the determinant

$$\Delta = \begin{vmatrix} 1 & -1 & 1 \\ 1 & 0 & -1 \\ 0 & -2 & 1 \end{vmatrix}$$

a first step could be to subtract the first from the second row,

$$\Delta = \begin{vmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & -2 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & -2 & 1 \end{vmatrix}$$

and then, by adding the first column to the second column and subtracting it from the third, obtain zeros in the first row except in the upper left-hand corner. Again, adding twice the second row to the third,

$$\Delta = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & -3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{vmatrix} = -3$$

and finally, adding twice the second column to the third column we obtain the desired product of single terms. Of course, we could have proceeded differently; we could indeed have stopped after the second stage and calculated the second-order determinant

$$\begin{vmatrix} 1 & -2 \\ -2 & 1 \end{vmatrix} = -3$$

While the evaluation of a determinant is relatively easy for $n = 2,3,4$, it rapidly becomes difficult, particularly if the coefficients are complicated. Using 2.31–2.35 we can now complete the

.36 Proof of Cramer's Rule If we multiply the equation 2.181 in order by Ai1, Ai2, …, Ain *and add, we have*

$$\Delta x_i = \Delta_i \qquad\qquad i = 1, 2, \ldots n$$

from 2.29 and 2.35, and the solutions xi of the equations 2.181 are well determined so long as $\Delta \neq 0$.

Example. Consider the three linear equations

$$\begin{aligned} x_1 - x_2 + x_3 &= 6 \\ x_1 \qquad - x_3 &= 1 \\ - 2x_2 + x_3 &= 4 \end{aligned}$$

for which we have seen that $\Delta = -3$ in the preceding example. Applying Cramer's rule and evaluating the determinants we have

$$\Delta_1 = \begin{vmatrix} 6 & -1 & 1 \\ 1 & 0 & -1 \\ 4 & -2 & 1 \end{vmatrix} = -9, \qquad \Delta_2 = \begin{vmatrix} 1 & 6 & 1 \\ 1 & 1 & -1 \\ 0 & 4 & 1 \end{vmatrix} = 3$$

$$\Delta_3 = \begin{vmatrix} 1 & -1 & 6 \\ 1 & 0 & 1 \\ 0 & -2 & 4 \end{vmatrix} = -6$$

so that

$$x_1 = \frac{-9}{-3} = 3, \quad x_2 = \frac{3}{-3} = -1, \quad x_3 = \frac{-6}{-3} = 2$$

EXERCISES

1. Evaluate the three determinants in Exercise 1 of Section 2.1 by adding and subtracting (a) rows, (b) columns, using scalar multiplication where necessary to produce a zero element.

2. Evaluate the determinants

$$\begin{vmatrix} 1 & 1 & 2 & 1 \\ 0 & 2 & 3 & -4 \\ 3 & 2 & 1 & 0 \\ 1 & -1 & 1 & 1 \end{vmatrix}, \quad \begin{vmatrix} 1 & -1 & 3 & -1 \\ 2 & -2 & -1 & 3 \\ 1 & -1 & -4 & 4 \\ 5 & -5 & -6 & 10 \end{vmatrix}, \quad \begin{vmatrix} 1 & 1 & 1 & 5 \\ 0 & 0 & 1 & 2 \\ 3 & 4 & -3 & 2 \\ 1 & -2 & 4 & 5 \end{vmatrix}$$

3. Solve the following system of equations by Cramer's rule:

$$\begin{aligned} x \quad\quad + 3z + w &= 2 \\ x + 2y + 2z - w &= -2 \\ 2x + 3y + z + w &= 2 \\ x - 4y \quad\quad + w &= 2 \end{aligned}$$

How could you have deduced the result by inspection?

4. Prove that

$$\begin{vmatrix} a & b & c \\ a^2 & b^2 & c^2 \\ a^3 & b^3 & c^3 \end{vmatrix} = abc(a - b)(b - c)(c - a)$$

without expanding the determinant.

2.4 INTERSECTIONS OF THREE PLANES

We saw in Chapter 1 that a plane in a 3-dimensional Euclidean space is represented by a linear equation

2.41
$$\pi_1: \quad a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = a_{10}$$

which passes through the origin if and only if $a_{10} = 0$. Let us consider the intersection of π_1 with the plane

2.42
$$\pi_2: \quad a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = a_{20}$$

If the normal vectors $A_1 = (a_{11}, a_{12}, a_{13})$ and $A_2 = (a_{21}, a_{22}, a_{23})$ are *parallel*, then

$$a_{11} : a_{12} : a_{13} = a_{21} : a_{22} : a_{23}$$

and without loss of generality we can assume that $A_1 = A_2$, so that π_1 and π_2 are distinct if $a_{10} \neq a_{20}$. Two parallel planes have no common points.

If π_1 and π_2 are *not* parallel, they intersect in a line l. In order to find the equations of l we first locate a point $Z(z_1, z_2, z_3)$ on l so that

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = a_{10} = a_{11}z_1 + a_{12}z_2 + a_{13}z_3$$
$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = a_{20} = a_{21}z_1 + a_{22}z_2 + a_{23}z_3$$

(e.g., we may set $x3 = 0$ and solve 2.41 and 2.42). Having thus determined Z, we may write 2.41 and 2.42 in the form

2.411 $$a_{11}(x_1 - z_1) + a_{12}(x_2 - z_2) + a_{13}(x_3 - z_3) = 0$$

2.421 $$a_{21}(x_1 - z_1) + a_{22}(x_2 - z_2) + a_{23}(x_3 - z_3) = 0$$

from which we obtain the parametric equations of l:

2.422
$$x_1 = z_1 + \tau A_{31}$$
$$x_2 = z_2 + \tau A_{32}$$
$$x_3 = z_3 + \tau A_{33}$$

where

$$A_{31} = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}, \quad A_{32} = -\begin{vmatrix} a_{11} & a_{13} \\ a_{23} & a_{21} \end{vmatrix}, \quad A_{33} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

are the direction numbers of l. Since we have assumed that $\pi 1$ and $\pi 2$ are not parallel, it follows that no two of $A31$, $A32$, $A33$ are zero.

Example. The two planes

$$x_1 - x_2 + 2x_3 = 0$$
$$x_1 + x_2 - x_3 = 2$$

intersect in a line l through the point $Z(1,1,0)$, so we may write the two equations in the form

$$(x_1 - 1) - (x_2 - 1) + 2x_3 = 0$$
$$(x_1 - 1) + (x_2 - 1) - x_3 = 0$$

Thus we have

$$x_1 = 1 - \tau, \quad x_2 = 1 + 3\tau, \quad x_3 = 2\tau$$

or in symmetric form

$$\frac{x_1 - 1}{1} = \frac{x_2 - 1}{3} = \frac{x_3}{2}$$

Consider now the intersection of $\pi 1$, $\pi 2$ and a third plane

$$\pi_3: \quad a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = a_{30}$$

If we assume that $\Delta = |A|$ and $\Delta 1$, $\Delta 2$, $\Delta 3$ are as defined in Section 2.1, then by Cramer's rule

.44 The coordinates of the point of intersection of the three planes $\pi 1$, $\pi 2$, $\pi 3$ are given by
$\Delta x1 = \Delta 1$, $\Delta x2 = \Delta 2$, $\Delta x3 = \Delta 3$ *provided* $\Delta \neq 0$.

    If we assume that $\Delta = 0$ and *not all of the cofactors A11, A21, A31 are zero,* then from 2.29 and 2.35

$$a_{11}A_{11} + a_{21}A_{21} + a_{31}A_{31} = 0$$
$$a_{12}A_{11} + a_{22}A_{21} + a_{32}A_{31} = 0$$
$$a_{13}A_{11} + a_{23}A_{21} + a_{33}A_{31} = 0$$

But these are just the scalar equations equivalent to the vector equation

$$A_1A_{11} + A_2A_{21} + A_3A_{31} = 0$$
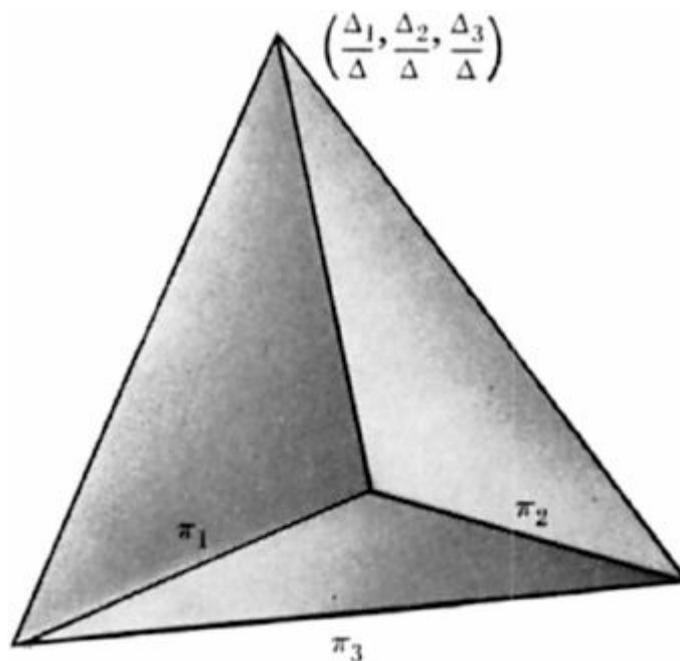
FIG. 2.1

where Ai is the ith row vector of A which defines the normal to the plane πi. We conclude that *these normal vectors are linearly dependent, and so coplanar* by 1.66. In other words, the planes $\pi 1$, $\pi 2$, $\pi 3$ are parallel to a fixed line l, perpendicular to the plane containing their normals. Put in this way we analyze the possibilities as follows.

.45 The three planes $\pi 1$, $\pi 2$, $\pi 3$ intersect in a line l. In this case the three

planes are linearly *dependent* and we may write

$$\pi_1 A_{11} + \pi_2 A_{21} + \pi_3 A_{31} = 0$$

This is equivalent to saying that in addition to the three equations written above we also have

$$a_{10} A_{11} + a_{20} A_{21} + a_{30} A_{31} = 0$$

so that the three rows of the nonsquare matrix

$$\tilde{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{10} \\ a_{21} & a_{22} & a_{23} & a_{20} \\ a_{31} & a_{32} & a_{33} & a_{30} \end{pmatrix}$$
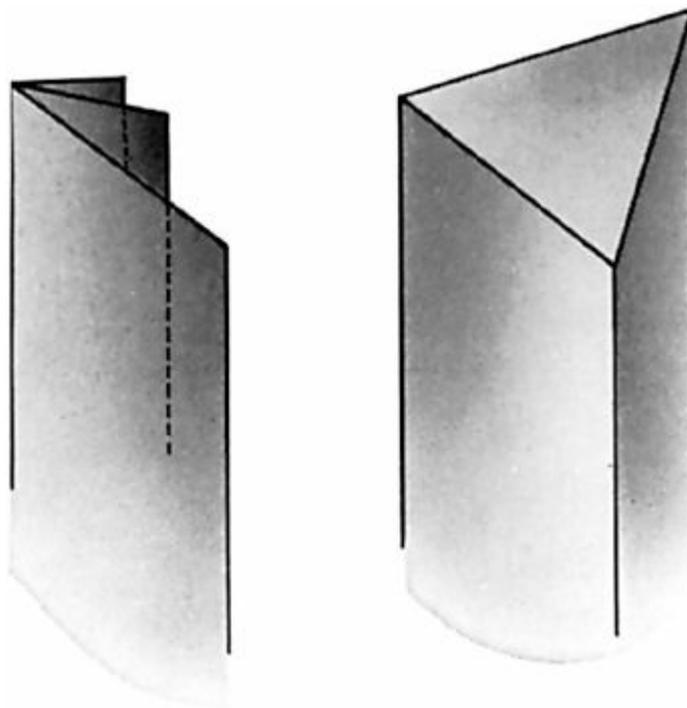
FIG. 2.2

are linearly dependent (note that we do *not* associate a determinant with a nonsquare matrix). The equations of l are given in 2.422, and the first two row vectors of $\tilde{A}$ are linearly *independent*. We say that the *row rank* of each of A and $\tilde{A}$ is 2.

.46 The three planes $\pi1$, $\pi2$, $\pi3$ intersect in three parallel lines. The direction numbers of these lines may be taken to be the co-factors in $|A|$, since

$$a_{11} A_{11} + a_{12} A_{12} + a_{13} A_{13} = 0$$
$$a_{21} A_{11} + a_{22} A_{12} + a_{23} A_{13} = 0$$

by 2.35. It follows that

$$A_{11}:A_{12}:A_{13} = A_{31}:A_{32}:A_{33}$$

and similarly

$$A_{21}:A_{22}:A_{23} = A_{31}:A_{32}:A_{33}$$

The row rank of A is still 2 but that of $\tilde{A}$ is now 3.

Example. The three planes (for which $\Delta = 0$),

$$\begin{aligned}
x_1 - x_2 + 2x_3 &= -1 \\
x_1 + x_2 - x_3 &= 1 \\
5x_1 + x_2 + x_3 &= k
\end{aligned}$$

intersect in a line with direction numbers $(-1,3,2)$ if $k = 1$, and otherwise in three parallel lines.

.47 The three planes are parallel. In this case we may assume that the row vectors of A coincide so that

$$A_1 = A_2 = A_3$$

but that $a10 \neq a20 \neq a30$. The row rank of A is now 1 while that of $\tilde{A}$ is 2. We could think of this as a limiting case of 2.45.

.48 Two of the three planes are parallel. We may suppose that $A1 = A2 \neq A3$ so that this is a limiting case of 2.46. The row ranks of A and $\tilde{A}$ are again 2 and 3. In order to distinguish this case from 2.46, we note that here

$$A_{31} = A_{32} = A_{33} = 0$$
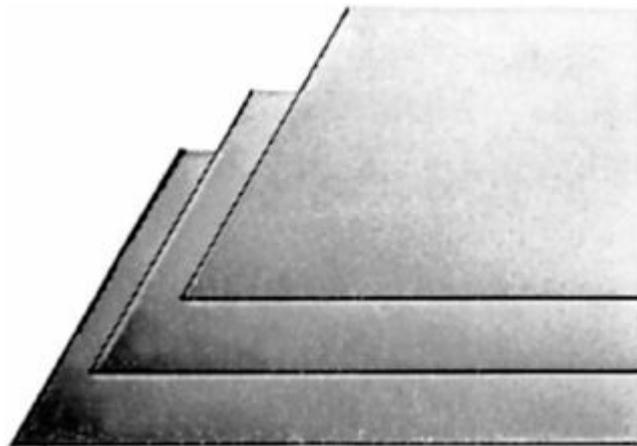
FIG. 2.3

whereas no such set A_{k1}, A_{k2}, A_{k3} vanishes in 2.46.

There remains the case in which $\Delta = 0$ with $A_{11} = A_{21} = A_{31} = 0$, so that the last two *column* vectors of A are linearly dependent. Since the direction numbers of the lines of intersection of π_1, π_2, π_3 are again proportional, we have a special case of 2.45 (2.46) in which the line (lines) is (are) parallel to a coordinate plane, or to one of the coordinate axes.

Example. In order to illustrate this special case of 2.45 where $\Delta = 0 = A_{11} = A_{21} = A_{31}$, we take the equations

$$\begin{aligned}
x_1 - x_2 + 2x_3 &= -1 \\
2x_1 + x_2 - 2x_3 &= 1 \\
3x_1 + 2x_2 - 4x_3 &= 2
\end{aligned}$$

in which the last three columns of $\tilde{A}$ are proportional, and the column rank of $\tilde{A}$ is 2. The direction numbers of the line of intersection are $(0,2,1)$. If we replaced the 2 on the right side of the last equation by, say, 0, the three planes would intersect in three lines parallel to the direction $(0,2,1)$, as in 2.46, and the column rank of $\tilde{A}$ would be 3.

EXERCISES

1. Are there any solutions of the equations

$$x_1 + x_2 - 3 = 0, \quad 2x_1 + x_2 - 4 = 0, \quad 3x_1 - 2x_2 - 1 = 0?$$

Plot the three lines on a piece of graph paper and explain the significance of your answer geometrically. Give the geometrical condition under which a common solution should exist. What would this amount to algebraically?

2. Find the parametric equations of the line of intersection of the two planes

$$\begin{aligned}
2x_1 - x_2 + x_3 &= 1 \\
x_1 + x_2 + x_3 &= 2
\end{aligned}$$

and prove that this line is parallel to the plane

$$x_1 - 2x_2 = 3$$

3. Prove that the three planes in Exercise 2 intersect in three parallel lines. Calculate the cofactors of the matrix A as in 2.46, and verify the required proportionality relations.

4. Find the equations of the lines of intersection of the three planes

$$x_1 - x_2 + x_3 = 1$$
$$x_1 - x_2 + x_3 = 2$$
$$x_1 - 2x_2 - x_3 = 1$$

in parametric form. Can these equations be written in symmetric form?

5. Examine the nature of the intersections of the sets of planes

(i) $\qquad x_1 - x_2 = 1, \quad x_2 - x_3 = 2, \quad 2x_1 - 2x_3 = 6$

(ii) $\qquad 2x_1 - x_2 + x_3 = 0, \quad x_2 - x_3 = 1, \quad x_1 + x_3 = 2$

and determine the coordinates of all common points.

2.5 HOMOGENEOUS EQUATIONS

There is still one possibility which we have not considered, namely, that the three planes $\pi 1$, $\pi 2$, $\pi 3$ *all* pass through the origin, in which case

2.51 $\qquad\qquad\qquad a_{10} = a_{20} = a_{30} = 0$

and the equations 2.41, 2.42, 2.43 are said to be *homogeneous*.

If these three homogeneous equations have a solution $x1 = a1, x2 = a2, x3 = a3$, then $x1 = ka1, x2 = ka2, x3 = ka3$ is also a solution for all real values of k; the three planes intersect in a line as in 2.45, and $\Delta = 0$. Conversely, our analysis of the preceding section shows that if $\Delta = 0$ and 2.51 holds, the distinction between 2.45 and 2.46 disappears whereas 2.47 and 2.48 do not apply. Moreover, if $\Delta \neq 0$, the only solution is $x1 = x2 = x3 = 0$ by Cramer's rule.

In order to generalize this result we prove first the important

.52 Theorem *The necessary and sufficient condition that the row (column) vectors of a matrix A be linearly dependent is that* $|A| = 0$.
Proof. The necessity follows immediately since the linear relation

$$A_1 a_1 + A_2 a_2 + \ldots + A_n a_n = 0$$

where Ai is the ith row vector of A and not all the ai vanish, implies that $\Delta = 0$ by 2.34. Similarly for the columns, if αi is the ith column vector of A,

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_n a_n = 0$$

implies that $\Delta = 0$.

Conversely, if $\Delta = 0$ and not all cofactors Aij vanish, let $Alk \neq 0$. Then

$$\sum_{i=1}^{n} a_{ij}A_{ik} = 0 \qquad\qquad j = 1,2,\ldots n$$

by 2.35, so that

$$A_1 A_{1k} + A_2 A_{2k} + \ldots + A_n A_{nk} = 0$$

Thus the row vectors of A are linearly dependent and the row rank of A is $< n$. By an exactly similar argument, the column rank of A is $<n$.

On the other hand, if $\Delta = 0$ and all $Aij = 0$ we must use induction. Let us take as our inductive assumption that the vanishing of a determinant of order $n - 1$ implies the linear dependence of its $n - 1$ rows. Certainly this is true for $n = 2$, 3. Now consider the matrix B made up of the first $n - 1$ rows of A. If the row rank of B were equal to $n - 1$, we would have a contradiction since, by assumption, every minor of order $n - 1$ vanishes and this implies that its row rank is $<n - 1$. We conclude that the row rank of B is $< n - 1$ so that of A must be $<n$. Similarly, the column rank of A is $<n$, which proves the theorem.

We can now prove that

.53 The necessary and sufficient condition that a system of n homogeneous linear equations in n unknowns x1, x2, ... xn should have a solution other than x1 = x2 = ... = xn = 0 is that the determinant of the coefficients have the value zero.

Proof. The necessity of the condition follows immediately since $\Delta x_i = \Delta_i = 0$ for all i, by Cramer's rule, and if some $x_i \neq 0$ we must have $\Delta = 0$.

On the other hand, $\Delta = 0$ implies that the column vectors of A are linearly dependent by 2.52, so that there exist numbers $x1, x2, \ldots xn$, not all zero, such that

$$\boldsymbol{\alpha}_1 x_1 + \boldsymbol{\alpha}_2 x_2 + \ldots + \boldsymbol{\alpha}_n x_n = 0$$

or in scalar form

2.531
$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0$$
$$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0$$
$$\ldots$$
$$a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = 0$$

but these $x1, x2, \ldots xn$ provide the nontrivial solution of the homogeneous equations 2.531 which we are seeking.

The study of analytical geometry is rewarding if we do things elegantly; otherwise it leads to a morass of ugly calculation. To suggest the elegant way, we solve the following two problems.

.54 *Find the equation of a linear subspace through n given points.* We could take $n = 2$ or $n = 3$, but since the argument holds for any n, we consider the general case. Let the n points be $Z_k(z_1^k, z_2^k, \ldots z_n^k)$ and the required equation be

2.541
$$a_1 x_1 + a_2 x_2 + \ldots + a_n x_n + a_{n+1} = 0$$

so that

2.542
$$a_1 z_1^k + a_2 z_2^k + \ldots + a_n z_n^k + a_{n+1} = 0$$
$$(k = 1, 2, \ldots n)$$

We may consider 2.541 along with 2.542 as a set of $n + 1$ homogeneous equations in the $n + 1$ unknowns $a1, a2, \ldots an, an + 1$ so that we must have

2.543
$$\begin{vmatrix} x_1 & x_2 \ldots x_n & 1 \\ z_1^1 & z_2^1 \ldots z_n^1 & 1 \\ z_1^2 & z_2^2 \ldots z_n^2 & 1 \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ z_1^n & z_2^n \ldots z_n^n & 1 \end{vmatrix} = 0$$

Any examiner would accept this!—but by subtracting the last row from each of the others the determinant can be reduced to one of order n,

2.544
$$\begin{vmatrix} x_1 - z_1^n & x_2 - z_2^n \ldots x_n - z_n^n \\ z_1^1 - z_1^n & z_2^1 - z_2^n \ldots z_n^1 - z_n^n \\ \cdot & \\ \cdot & \\ \cdot & \\ z_1^{n-1} - z_1^n & z_2^{n-1} - z_2^n \ldots z_n^{n-1} - z_n^n \end{vmatrix} = 0$$

This result should be compared with 1.67.

.55 *Find the equation of the plane through two intersecting lines*

2.551
$$a_{11} x_1 + a_{12} x_2 + a_{13} x_3 = 0 = a_{21} x_1 + a_{22} x_2 + a_{23} x_3$$

2.552
$$b_{11} x_1 + b_{12} x_2 + b_{13} x_3 = 0 = b_{21} x_1 + b_{22} x_2 + b_{23} x_3$$

The important thing is to find the vector $V(v_1, v_2, v_3)$ normal to each line. But direction numbers of 2.551 are

$$\textbf{2.56} \qquad l_1 = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}, \qquad l_2 = \begin{vmatrix} a_{13} & a_{11} \\ a_{23} & a_{21} \end{vmatrix}, \qquad l_3 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

as in 2.422, and similar expressions yield direction numbers $m1$, $m2$, $m3$ of 2.552. Thus if we assume the equation of the required plane is

$$v_1 x_1 + v_2 x_2 + v_3 x_3 = 0$$

we must have

$$v_1 l_1 + v_2 l_2 + v_3 l_3 = 0$$

and

$$v_1 m_1 + v_2 m_2 + v_3 m_3 = 0$$

Eliminating, according to 2.53 we have the required equation:

$$\textbf{2.57} \qquad \begin{vmatrix} x_1 & x_2 & x_3 \\ l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{vmatrix} = 0$$

EXERCISES

1. Discuss the solutions of the system of equations

$$x_1 + 2x_2 + 3x_3 = 0, \qquad 2x_1 - x_2 + x_3 = 0$$
$$x_1 + x_2 + 2x_3 = 0$$

and their geometrical significance.

2. Give the equation of the plane through the three points $(-1,2,0)$, $(1,0,3)$, $(-1,2,-2)$ in determinantal form and expand the determinant.

3. Write down the equation of the plane through the points $(1,1,0)$ and $(-1,0,2)$ parallel to the line $x1 = x2 = -x3$.

4. Find the equation of the plane through the point $(1,2,3)$ and the line

$$x_1 - x_2 - x_3 = 0, \qquad 2x_1 + x_2 - x_3 = 2$$

5. Find the equation of the plane through the two intersecting lines

$$x_1 - x_2 + x_3 = 0 = x_1$$
$$2x_1 + x_2 - x_3 = 0 = x_2$$

6. Prove that the determinantal equation 2.543 of the linear subspace is equivalent to the vector equation

$$X = a_1 Z_1 + a_2 Z_2 + \ldots + a_n Z_n$$

with the condition that $a1 + a2 + \ldots + an = 1$. How would you interpret this vector equation as a set of parametric equations for the linear subspace in question?

3

MATRICES

3.1 MATRIX ADDITION AND MULTIPLICATION

In the preceding discussion of determinants and their applications to algebra and geometry, we have often found it convenient to speak of a square "array" or *matrix* apart from the calculations involved in evaluating its determinant. We also have seen that a matrix need not be square, but in this case a determinant is not defined.

Let us now consider matrices of this general form: matrices which have m rows and n columns, where $m \lesseqgtr n$. Such a matrix A is best thought of as a rectangular array of n column vectors ($a1j$, $a2j$, ... amj) or m row vectors ($ai1$, $ai2$, ... ain), and we write

3.11 $$A = (a_{ij}) \qquad \begin{aligned} i &= 1,2, \ldots m \\ j &= 1,2, \ldots n \end{aligned}$$

If $B = (bij)$, we shall write $A = B$ if and only if $aij = bij$ for all i, j.

As in the case of vectors, we can *add* matrices by simply adding corresponding row and column vectors, i.e., by adding *elements:*

3.12 $$A + B = (a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$$

Also, we can define the multiplication of a matrix A by a scalar k:

3.13 $$kA = Ak = (ka_{ij})$$

In particular, for $k = 0$ we have the *zero matrix,* all of whose elements are zero.

In order to see the significance that can be attached to a matrix and to suggest how we may define the multiplication of two matrices, we take the following simple calculation from everyday life.

Example. A housewife goes to market to buy 1 lb. coffee at 75¢/lb., $\frac{1}{2}$ lb. cheese at 60¢/lb., 2 lb. butter at 50¢/lb., 1 doz. oranges at 40¢/doz., and 3 loaves bread at 20¢/loaf. Her total bill is

$$1 \times 75 + \tfrac{1}{2} \times 60 + 2 \times 50 + 1 \times 40 + 3 \times 20 = \$3.05$$

which we have written as the *inner product A·B* of the two *vectors* $A(1,\frac{1}{2},2,1,3)$ and $B(75,60,50,40,20)$.

If we follow this suggestion by defining the *product* of two matrices A, B in terms of the inner products of their row and column vectors, then such vectors must obviously have the same number of components. Returning to the foregoing example, convention decrees that we write:

$$(1,\tfrac{1}{2},2,1,3) \begin{pmatrix} 75 \\ 60 \\ 50 \\ 40 \\ 20 \end{pmatrix} = (305)$$

i.e., we take the inner product of a *row* on the left with a *column* on the right.

.14 Definition The *product AB* of an $m \times n$ matrix A and a $p \times q$ matrix B is an $m \times q$ matrix C if and only if $p = n$. Setting

$$A = (a_{ij}), \quad B = (b_{jk}), \quad C = (c_{ik})$$

we have

$$AB = (a_{ij})(b_{jk}) = (c_{ik}) = C$$

where

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \ldots + a_{in}b_{nk}$$

is the *inner product* of the ith row vector of A and the kth column vector of B.

The product AB of an $m \times n$ matrix A and a $p \times q$ matrix B is *not* defined if $p \neq n$; thus the existence of AB does not imply the existence of BA. Even if *both AB and BA* are defined, as in the case of square matrices, they are in general different, e.g. :

3.15

$$\begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} -2 & -1 \\ 5 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} -2 & -5 \\ 1 & 0 \end{pmatrix}$$

.16 Definition If $AB = BA$, then A and B are said to *commute*.

One might imagine that definitions hedged about with so many conditions

would not lead to very significant ideas. On the contrary, though we shall not be concerned very much with nonsquare matrices apart from vectors, square matrices A, B and their products will play a major role in what follows. Curiously enough, we must postpone the proof that $|AB| = |A||B|$ until we have developed more machinery.

3.2 TRANSPOSE OF A MATRIX

We have already introduced the notion of the transpose of a square matrix A obtained by interchanging the matrix rows and columns, proving in 2.27 that $|A| = |At|$. What is the significance of transposition for the multiplication of matrices? We prove the following important result:

.21 If At and Bt are the transposed matrices of A and B, then

$$(AB)^t = B^t A^t$$

The following example illustrates what is going on:

$$\begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} -2 & -1 \\ 4 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} -2 & 4 \\ -1 & 0 \end{pmatrix}$$

$$\quad A \qquad\qquad B \qquad\qquad AB \qquad\qquad\qquad B^t \qquad\qquad A^t \qquad\qquad (AB)^t$$

Proof. Let us suppose that $A = (aij)$, $B = (bkl)$ are $n \times n$ square matrices so that the product $AB = C$ is defined. Setting $C = (crs)$ we have

$$c_{rs} = a_{r1}b_{1s} + a_{r2}b_{2s} + \ldots + a_{rn}b_{ns} = b_{1s}a_{r1} + b_{2s}a_{r2} + \ldots + b_{ns}a_{rn}$$

Since this is the element in the sth row and rth column of $BtAt$, we have proved that $BtAt = Ct = (AB)t$, as desired. By successive applications, we have

3.22 $$(ABC \ldots D)^t = D^t \ldots C^t B^t A^t$$

EXERCISES

1. If

$$A = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & -2 & -5 \\ -2 & 2 & -1 \end{pmatrix}$$

  (a) Show that $2A - B + C = 0$.

(b) Verify that $ABt = \begin{pmatrix} -2 & 4 \\ -2 & 1 \end{pmatrix} = (BAt)t.$

(c) Calculate $ABtC$ and verify that

$$AB'C = AB'B - 2AB'A$$

2. If

$$X = \begin{pmatrix} 0 & 1 & -2 \\ -1 & 0 & 1 \\ 2 & -1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 1 & 1 & 0 \\ -2 & 0 & 1 \\ 0 & -2 & 1 \end{pmatrix}$$

calculate $X2$, XY, YX, $Y2$.

3. Any matrix A is called *symmetric* if $A = At$. Prove that $S = \frac{1}{2}(A + At)$ is symmetric, and determine the matrix K such that

$$A = S + K$$

Prove that $Kt = -K$; such a matrix is called *skew-symmetric*.

4. Express each of the matrices X, Y in Exercise 2 as a sum of a symmetric matrix S and a skew-symmetric matrix K.

5. Prove that every integral power of a symmetric matrix is symmetric.

6. Prove that every *even* positive integral power of a skew-symmetric matrix is symmetric, but every *odd* positive integral power is skew-symmetric.

7. If K is an $n \times n$ skew-symmetric matrix, prove that $|K| = 0$ if n is odd.

3.3 INVERSE OF A MATRIX

If we can multiply $n \times n$ matrices, it is natural to look for an $n \times n$ matrix which, under multiplication, produces no change. Clearly, such a matrix is the *unit matrix*

3.31
$$I = \begin{pmatrix} 1 & 0 & 0 \dots 0 \\ 0 & 1 & 0 \dots 0 \\ & \cdot & \\ & \cdot & \\ & \cdot & \\ 0 & 0 & 0 \dots 1 \end{pmatrix}$$

and $IA = AI = A$ for every $n \times n$ matrix A.

With the analogy of ordinary arithmetic in mind, it would be natural to designate the matrices

$$\begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}$$

as inverses of each other, since

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

If we denote the *inverse* of A by $A-1$, the question arises, does every $n \times n$ matrix A have an inverse, and is this inverse unique?

Consider the simple case where we assume that

$$A = \begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}, \quad AB = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

so that

3.32
$$\begin{aligned} b_{11} - b_{21} &= 1, & b_{12} - b_{22} &= 0 \\ -2b_{11} + 2b_{21} &= 0, & -2b_{12} + 2b_{22} &= 1 \end{aligned}$$

Since $|A| = 0$, these equations are inconsistent and the matrix A has no inverse; A is said to be *singular*. Conversely, by Cramer's rule, the inverse $A-1$ will certainly exist if $|A| \neq 0$, in which case A is said to be *nonsingular*.

We shall give two methods of constructing the inverse $A-1$ of a nonsingular matrix A, the first along the lines of the above example. Let us suppose that $A = (aij)$, $B = (bkl)$ are two $n \times n$ matrices such that $AB = I$ where I is the unit $n \times n$ matrix 3.31; then the following n equations determine the jth column vector $(b1j, b2j, \ldots bnj)$ of B:

$$a_{11}b_{1j} + a_{12}b_{2j} + \ldots + a_{1n}b_{nj} = 0$$
$$\cdot$$
$$\cdot$$
3.33
$$a_{j1}b_{1j} + a_{j2}b_{2j} + \ldots + a_{jn}b_{nj} = 1$$
$$\cdot$$
$$a_{n1}b_{1j} + a_{n2}b_{2j} + \ldots + a_{nn}b_{nj} = 0$$

In order to solve these equations 3.33 we multiply the first by $A1i$, the second by $A2i$, ..., and the last by Ani where Aij is the cofactor of aij in A; adding, every sum on the left vanishes except one:

3.34
$$(a_{1i}A_{1i} + a_{2i}A_{2i} + \ldots + a_{ni}A_{ni})b_{ij} = A_{ji}$$

But the inner product on the left is just $\Delta = |A|$ by 2.29, so that

3.35
$$b_{ij} = A_{ji}/\Delta$$

assuming that $\Delta \neq 0$. Thus, if we define the *adjoint* of A to be the matrix of cofactors

$$(A_{ij})$$

we have the inverse matrix A–l of A given by

3.36
$$A^{-1} = \frac{1}{\Delta} (A_{ij})^t$$

from which we conclude that

3.37
$$(A^{-1})^t = \frac{1}{\Delta} (A_{ij}) = (A^t)^{-1}$$

Again, we encounter the problem of evaluating the determinant of a matrix. If the matrix A is large, as happens in many practical applications, lthe construction of $A - 1$ by this method is difficult, and so we have recourse to quite a different line of thought in the following section. Before proceeding further, however, it is important to show that

3.38
$$AA^{-1} = A^{-1}A = I$$

Let us suppose that A–1 $= A_R^{-1}$, called the *right inverse* of A, and imagine that solving a different set of equations 3.33 would lead to a *left inverse* A_L^{-1}, where $AA_R^{-1} = A_L^{-1}A = I$. Clearly

3.39
$$A_L^{-1} = A_L^{-1}I = A_L^{-1}AA_R^{-1} = IA_R^{-1} = A_R^{-1}$$

and the *inverse A*–1 is uniquely defined.

EXERCISES

1. Discuss the significance of the products

$$(1,2)\begin{pmatrix} 3 \\ -1 \end{pmatrix} = (1,2)\begin{pmatrix} 5 \\ -2 \end{pmatrix} = (1,2)\begin{pmatrix} 7 \\ -3 \end{pmatrix} = (1)$$

for the existence of the inverse of a nonsquare matrix. Is there any matrix A

such that $A(1,2) = (1)$?

2. Has the matrix X in Exercise 2 of the preceding section an inverse? Find the inverse of the matrix Y and verify that 3.38 holds.

3. If

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

calculate $A{-}1$ and $(At){-}1$ and verify that $(A{-}1)t = (At){-}1$.

3.4 REDUCTION OF A MATRIX TO CANONICAL FORM

In Chapter 2 we proved a sequence of theorems 2.31–2.34 whereby we were able to make some progress in the evaluation of a determinant. Let us try to organize the steps which were suggested in the first example of Section 2.3 into a sequence of *elementary operations*. We propose to define them with reference to a matrix A; *it is not necessary here to assume that A is square.* As a by-product of the discussion, we shall obtain a second construction for $A{-}1$ in this special case.

.41 *To interchange the first two row (column) vectors of A we multiply on the left (right) by the matrix*

3.411
$$\left(\begin{array}{cc|ccc} 0 & 1 & 0 & & 0 \\ 1 & 0 & 0 & & 0 \\ \hline 0 & 0 & 1 & & \\ & & & 1 & \\ 0 & 0 & & & 1 \end{array} \right)$$

If the 2×2 matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is Properly placed, we may interchange any two row (column) vectors of A.

.42 *To multiply the first row (column) vector of A by k we multiply on the left (right) by the matrix*

3.421

$$\begin{pmatrix} k & 0 & 0 & & & 0 \\ 0 & 1 & 0 & & & 0 \\ \hline 0 & 0 & 1 & & & \\ & & & 1 & & \\ 0 & 0 & & & & 1 \end{pmatrix}$$

By placing k properly we may multiply any row (column) vector of A by k.

.43 *To add k times the first row or column vector of A to the second row or column vector we multiply on the left or right by the matrix*

3.431

$$\begin{pmatrix} 1 & 0 & 0 & & & 0 \\ k & 1 & 0 & & & 0 \\ \hline 0 & 0 & 1 & & & \\ & & & 1 & & \\ 0 & 0 & & & & 1 \end{pmatrix} \quad or \quad \begin{pmatrix} 1 & k & 0 & & & 0 \\ 0 & 1 & 0 & & & 0 \\ \hline 0 & 0 & 1 & & & \\ & & & 1 & & \\ 0 & 0 & & & & 1 \end{pmatrix}$$

If the 2×2 matrix $\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$ or $\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$ is Properly placed, the addition of k times any given row or column vector can be made to any other row or column vector.

Matrices of the form 3.411, 3.421, 3.431 are called *elementary* and the operation accomplished by multiplying A on the left or right by such a matrix is called an *elementary operation.*

Provided not every $aij = 0$, we may, by multiplying on the left and on the right by 3.411 and 3.421, arrange that $a11 = 1$. When this has been accomplished, we may, by successive multiplication by matrices 3.431 on the left, arrange that $ai1 = 0$ for $i > 1$. After similarly arranging that $a22 = 1$, we may again arrange that $aj2 = 0$ for $j > 2$ and so on until all elements *below* the diagonal vanish. By further multiplication on the right we may arrange that every element above the diagonal vanishes also. Thus we may write

$(Ps \ldots P2P1)A(Q1Q2 \ldots Qt) = PAQ$

3.44

$$= \begin{pmatrix} 1 & & & & & & 0 \\ & 1 & & & & & \\ & & 1 & & & & \\ & & & \cdot & & & \\ & & & & \cdot & 1 & \\ & & & & & 0 & \\ 0 & & & & & & 0 \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} r_1 \end{matrix}$$

$$\underbrace{}_{r_2}$$

where $P = Ps \ldots P2P1$ and $Q = Q1Q2 \ldots Qt$ The matrix PAQ is now said to be

in *canonical form.*

Let us consider carefully the significance of what we have done. Certainly, multiplication on the left (right) by 3.411 or 3.421 will not change the *number* of linear relations holding between the row (column) vectors of A, though the relations themselves will change. Nor will multiplication on the left (right) by 3.431 affect the number of such linear relations, though again the form of the relations will change. Thus we have proved that the *row rank and the column rank of PAQ are the same as the row rank $r1$ and column rank $r2$ of A.* But it follows from 3.44 that $r1 = r2 = r$, which is called the *rank* of A. We sum up these conclusions in the

.45 Theorem By multiplying on the left and right by suitably chosen elementary matrices Ps ... P2P1 = P and Q1Q2 ... Qt = Q, any n × n matrix A may be reduced to canonical form PAQ = Ir, where the number of 1's in the diagonal of Ir is equal to the rank r ($\leqslant$ n) of A.
We illustrate this important result by the following
Example. If we suppose that

$$A = \begin{pmatrix} 2 & 4 & 0 \\ 3 & 7 & 3 \\ 0 & 2 & 6 \end{pmatrix}$$

then

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}}_{P_3} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{P_2} \underbrace{\begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{P_1} A = \underbrace{\begin{pmatrix} \frac{1}{2} & 0 & 0 \\ -\frac{3}{2} & 1 & 0 \\ 3 & -2 & 1 \end{pmatrix}}_{P} A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

so that

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix} \underbrace{\begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{Q_1} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix}}_{Q_2} = \underbrace{\begin{pmatrix} \frac{1}{2} & 0 & 0 \\ -\frac{3}{2} & 1 & 0 \\ 3 & -2 & 1 \end{pmatrix}}_{P} A \underbrace{\begin{pmatrix} 1 & -2 & 6 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix}}_{Q}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = I_2$$

and the rank of A is 2.

EXERCISES

1. Determine the rank of each of the matrices

$$\begin{pmatrix} 1 & -1 & 2 \\ 2 & -2 & 4 \end{pmatrix}, \quad \begin{pmatrix} 1 & -1 & 2 \\ 2 & -2 & 1 \end{pmatrix}$$

by applying elementary operations as described above.

2. How much of the reduction in Exercise 1 could you accomplish by operating (a) on the left *only*, (b) on the right *only*?

3. Reduce the matrix

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

to canonical form by applying elementary operations (a) on both sides, (b) on the left *only*, (c) on the right *only*.

4. Prove that a matrix A can be reduced to canonical form by elementary operations (a) on the left *only*, (b) on the right *only*, if A is square and also nonsingular. Are these conditions necessary as well as sufficient?

3.5 INVERSE OF A MATRIX (SECOND METHOD)

In Section 3.3 we agreed that the inverse of an $n \times n$ matrix A exists if and only if $|A| \neq 0$. Now each of the elementary matrices 3.411, 3.421, 3.431 satisfies this condition, so we may write its inverse and conclude that

3.51 $\qquad P^{-1} = P_1^{-1}P_2^{-1}\dots P_s^{-1}, \qquad Q^{-1} = Q_t^{-1}\dots Q_2^{-1}Q_1^{-1}$

Thus if A is nonsingular, $PAQ = Ir = I$ so that

3.52 $\qquad\qquad\qquad\qquad A = P^{-1}IQ^{-1} = P^{-1}Q^{-1}$

Since the inverse of an elementary matrix is again elementary, 3.52 expresses A as a product of elementary matrices. On the other hand, we can take the inverse of each side of 3.52 to obtain the important result

3.53 $\qquad\qquad\qquad\qquad A^{-1} = QP$

which yields a second and more practical method of calculating A–1. However, if A is singular, $PAQ = Ir$ of rank $r < n$. Certainly $A = P$–IrQ–1, but we cannot

take the inverse of I_r, so that *no expression corresponding* to 3.53 exists if A is singular.

Example. Consider the case where

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

Then

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{P_2} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}}_{P_1} \underbrace{\begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 3 \\ 1 & 0 & 2 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{Q_1} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}}_{Q_2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

so that

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

The advantage of this method of calculating $A{-}1$ is just that each "elimination" is explicit and, though the method is based on Theorems 2.31 and 2.34, *no evaluations of determinants are required*. We shall have more to say on the practical computations involved after we have proved that

.54 If A and B are both n × n matrices, then |AB| = |A||B|.
Proof. If $|B| = 0$, then the row vectors of B are linearly dependent by 2.52, and since the row vectors of AB are just linear combinations of those of B, these must also be linearly dependent and $|AB| = 0$, again by 2.52. If $|A| = 0$, then $|At = 0$ by 2.27 and we can apply the same argument to $(AB)t = BtAt$ to conclude that $|(AB)t| = |AB| = 0$ as before.

If neither A nor B is singular, we base our discussion on the possibility of expressing a nonsingular matrix A as a product of elementary matrices, as in 3.52. If we could prove that

3.541 $$|CB| = |C||B|$$

for C an elementary matrix, then by breaking up and recombining, we could deduce the general result. But 3.541 follows immediately for:

(i) C of type 3.411, from 2.31 since

$$\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1$$

(ii) C of type 3.421, from 2.33 since

$$\begin{vmatrix} k & 0 \\ 0 & 1 \end{vmatrix} = k$$

(iii) C of type 3.431, from 2.34 since

$$\begin{vmatrix} 1 & 0 \\ k & 1 \end{vmatrix} = \begin{vmatrix} 1 & k \\ 0 & 1 \end{vmatrix} = 1$$

Tackling the general case, we write

$$\begin{aligned} |AB| &= |C_1 C_2 \ldots C_t B| \\ &= |C_1||C_2 \ldots C_t B| = |C_1||C_2||C_3 \ldots C_t B| \\ &= |C_1 C_2||C_3 \ldots C_t B| = \ldots\ldots \\ &= |C_1 C_2 \ldots C_t||B| = |A||B| \end{aligned}$$

as we desired to prove.

Since $|AA{-}1| = |I| = 1 = |A||A{-}1|$, we have

3.55 $$|A^{-1}| = 1/|A|$$

EXERCISES

1. Find the inverse of the matrix Y in Exercise 2 of Section 3.2 by the method of this section.
2. Find the inverses of the matrices

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 & 1 & -1 \\ 1 & -1 & 2 & 0 \\ -1 & 2 & 0 & 1 \\ 0 & 1 & -1 & 2 \end{pmatrix}$$

by the method of Section 3.3 and also by the method of this section.

3.6 THE APPROXIMATE INVERSE OF A MATRIX

In order to see the distance we have traveled since we first introduced the problem of solving the system of linear equations 2.181, let us write this system in matrix form. We have two possibilities: (i) We may consider the vector with components $x1$ $x2$, … xn as a *column vector*, so that

$$3.61 \quad \begin{pmatrix} a_{11} & a_{12} \ldots a_{1n} \\ a_{21} & a_{22} \ldots a_{2n} \\ \cdot \\ \cdot \\ \cdot \\ a_{n1}a_{n2} \ldots a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = \begin{pmatrix} a_{10} \\ a_{20} \\ \cdot \\ \cdot \\ \cdot \\ a_{n0} \end{pmatrix}$$

or (ii) we may consider the vector with components $x1$, $x2$, … xn as a *row vector*, so that

$$3.62 \quad (x_1, x_2, \ldots x_n) \begin{pmatrix} a_{11} & a_{21} & a_{n1} \\ a_{12} & a_{22} & a_{n2} \\ \cdot \\ \cdot \\ \cdot \\ a_{1n} & a_{2n} & a_{nn} \end{pmatrix} = (a_{10}, a_{20}, \ldots a_{n0})$$

It is important to observe that 3.62 is just the *transpose* of 3.61 according to 3.21, and *there is nothing to choose between the two methods of writing the equdtions except ds convenience mdy dictdte.*

If we denote the column vectors appearing in 3.61 by X and $\alpha 0$, we may write the vector equation in the form

$$3.63 \qquad\qquad AX = \alpha_0$$

and its solution in the form

$$3.64 \qquad\qquad X = A^{-1}\alpha_0$$

assuming that A is nonsingular. We illustrate this *second method* of solution in the following
Exdmple. Consider the set of equations

$$x_1 + x_2 = 1$$
$$x_1 + 1.01x_2 = 2$$

By Cramer's rule,

$$x_1 = \frac{\begin{vmatrix} 1 & 1 \\ 2 & 1.01 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 1 & 1.01 \end{vmatrix}} = \frac{-0.99}{0.01}, \qquad x_2 = \frac{\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 1 & 1.01 \end{vmatrix}} = \frac{1}{0.01}$$

We could write the matrix equation in the form

$$\begin{pmatrix} 1 & 1 \\ 1 & 1.01 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

and by constructing the inverse matrix, obtain its solution in the form

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 101 & -100 \\ -100 & 100 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -99 \\ 100 \end{pmatrix}$$

From the practical point of view this second method of determining $A{-}1$ is preferable for large values of n, but there is another consideration which enters into the problem. In the above example only one coefficient has been chosen to be nonintegral. If such coefficients were obtained experimentally, or were subject to an assigned "error," one might very well ask for the effect of such an error on the solution. *The answer has to do with how the inverse $A{-}1$ is calculated.*

In the previous section, we expressed $A{-}1$ as a product of elementary matrices, some on the right and some on the left. But we could have restricted ourselves to the right or the left *only* (cf. Exercises 3 and 4 of). Take again the matrix A of the equations of the preceding example:

$$3.65 \qquad \underbrace{\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 100 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}}_{A_L^{-1}} \underbrace{\begin{pmatrix} 1 & 1 \\ 1 & 1.01 \end{pmatrix}}_{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and similarly on the right. In this case $A{-}1 = A_L^{-1} = A_R^{-1}$.

If we were working approximately, we might have arrived at

$$3.66 \qquad \underbrace{\begin{pmatrix} 1 & 1 \\ 1 & 1.01 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} 103.01 & -99.99 \\ -102 & 100 \end{pmatrix}}_{A_R^{-1}} = \begin{pmatrix} 1.01 & 0.01 \\ -0.01 & 1.01 \end{pmatrix}$$

which is approximately I, whereas

$$3.67 \qquad \begin{pmatrix} 103.01 & -99.99 \\ -102 & 100 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1.01 \end{pmatrix} = \begin{pmatrix} 3.02 & 2.02 \\ -2 & 1.01 \end{pmatrix}$$

$$\underset{A_R^{-1}}{} \qquad\qquad \underset{A}{}$$

which is very different from I.

The foregoing example illustrates several things in a striking way:

(a) Matrix multiplication is *not* commutative in general.

(b) The *exact inverse* of a matrix is the same however it may be calculated (cf. 3.39).

(c) The *approximate inverse*, which is that used in actual computation, depends on *how* it is calculated, i.e., whether from the right or left. Approximate inverses calculated from both sides are particularly likely to introduce errors.*

3.7 LINEAR TRANSFORMATIONS

In 3.63, we used matrices to express a system of n linear equations as one equation. The vector $\alpha 0$ was supposed fixed and we sought the solution vector X. If we write the equation in the form

$$3.71 \qquad\qquad\qquad Y = AX$$

we have a *linear transformation* or *mapping* of the vectors X of υ n onto the vectors Y; usually we shall assume that Y lies in υn, but it could well be a vector X' of another vector space υ'_n.

There are several important remarks concerning linear transformations which we shall make here, leaving their illustration and detailed development to the next and subsequent chapters. In the first place, the equation 3.71 could equally well be written in the form

$$3.72 \qquad\qquad\qquad Y^t = X^t A^t$$

where $Y t$ and $X t$ are *row* vectors. The two equations 3.71 and 3.72 correspond to 3.61 and 3.62.

If we think of applying the two linear transformations

$$3.73 \qquad\qquad\qquad Z = AY, \qquad Y = BX$$

in succession, we obtain as a result the linear transformation

$$3.74 \qquad\qquad\qquad Z = (AB)X$$

Actually, we have a theorem here, but the proof of 3.74 is more a matter of understanding the definition of matrix multiplication than of performing any

additional mathematical operation. To clarify what is going on, we write out the steps in detail in the 2×2 case.

$$z_1 = a_{11}y_1 + a_{12}y_2, \qquad y_1 = b_{11}x_1 + b_{12}x_2$$
$$z_2 = a_{21}y_1 + a_{22}y_2, \qquad y_2 = b_{21}x_1 + b_{22}x_2$$

so that

$$z_1 = a_{11}(b_{11}x_1 + b_{12}x_2) + a_{12}(b_{21}x_1 + b_{22}x_2)$$
$$= (a_{11}b_{11} + a_{12}b_{21})x_1 + (a_{11}b_{12} + a_{12}b_{22})x_2$$
$$z_2 = a_{21}(b_{11}x_1 + b_{12}x_2) + a_{22}(b_{21}x_1 + b_{22}x_2)$$
$$= (a_{21}b_{11} + a_{22}b_{21})x_1 + (a_{21}b_{12} + a_{22}b_{22})x_2$$

If we write the linear transformations 3.73 in transposed form we have:

3.75 $$\boldsymbol{Z}^t = \boldsymbol{Y}^t \boldsymbol{A}^t, \qquad \boldsymbol{Y}^t = \boldsymbol{X}^t \boldsymbol{B}^t$$

so that

3.76 $$\boldsymbol{Z}^t = \boldsymbol{X}^t(\boldsymbol{B}^t \boldsymbol{A}^t) = \boldsymbol{X}^t(\boldsymbol{AB})^t$$

No ambiguity can arise if we drop the superscript t on the vectors $\boldsymbol{X}t$, $\boldsymbol{Y}t$ etc., writing 3.72 as

3.77 $$\boldsymbol{Y} = \boldsymbol{X} \boldsymbol{A}^t$$

since *the form of writing the vectors is determined by the rule of matrix multiplication.* We shall insist, however, on calling A the *matrix* of the transformation, whether this is written in the form 3.71 or 3.77.

In Chapter 1 we introduced the notion of the basis of an n-dimen-sional vector space υn:

$$\boldsymbol{E}_1 = (1,0, \ldots 0), \quad \boldsymbol{E}_2 = (0,1, \ldots 0), \quad \ldots, \quad \boldsymbol{E}_n = (0,0, \ldots 1)$$

It is natural to ask how such vectors are transformed by the linear transformation $\boldsymbol{Y} = \boldsymbol{AX}$. We have

$$AE_i = \begin{pmatrix} a_{11}a_{12} & \cdots & a_{1n} \\ a_{21}a_{22} & \cdots & a_{2n} \\ \cdot & & \\ & & \\ a_{i1}a_{i2} & \cdots & a_{in} \\ \cdot & & \\ a_{n1}a_{n2} & \cdots & a_{nn} \end{pmatrix} E_i = \alpha_i$$

where α_i is the ith *column* vector of A. Conversely, if we require that E_i ($i =$ 1,2, … n) be transformed into α_n, then the matrix A is completely determineIt is natural to call thed. We collect together these ideas in the following theorem.

.78 A linear transformation $Y = AX$ maps the basis vectors E_i ($i = 1,2, \dots n$) on the column vectors α_i of A. These vectors α_i are linearly independent if, and only if, A is nonsingular. Conversely, the linear transformation is completely determined when the vectors α_i are given.

    It is natural to call the linear transformation 3.71 or 3.77 singular or nonsingular according to whether A is singular or nonsingular. It follows immediately that the inverse of $Y = AX$ is $X = A{-1}Y$ and, in transposed form, that the inverse of $Y = XAt$ is $X = Y(At){-1} = Y(A{-1})t$ by 3.37.

EXERCISES

1. Write the following system of equations in the form 3.61 and solve by constructing the matrix $A{-1}$.

$$x_1 - x_2 + x_3 + x_4 = 2, \qquad 3x_1 + 2x_2 + x_3 + x_4 = 0$$
$$2x_1 + x_2 + x_3 + x_4 = 1, \qquad 4x_1 \qquad + 2x_3 + 3x_4 = 3$$

2. If basis vectors E_i ($i = 1,2,3,4$) are transformed by a linear transformation $Y = AX$ into vectors $\alpha_1(1,2,3,4)$, $\alpha_2(-1,1,2,0)$, $\alpha_4(1,1,1,3)$, respectively, find the vector Y into which the vector $X(1,0,1,0)$ is transformed.

3. If the matrix A is defined as in Exercise 2, find the vector Y into which $X(1,0,1,0)$ is transformed by the linear transformation (a) $Y = A2X$, (b) $Y = XA$.

    * Mendelsohn, N. S., "Some Elementary Properties of Ill-Conditioned Matrices and Linear Equations," *Am. Math. Monthly,* 63, 285–295 (1956).

4

GROUPS AND LINEAR TRANSFORMATIONS

4.1 DEFINITION OF A GROUP*

If everyone in a gathering of students were asked to give his definition of mathematics, many would define *algebra* and some *geometry*; to the economist, mathematics is *arithmetic* or *statistics*, and to the engineer it is almost certainly the *calculus*. Possibly they might all agree on the definition: Mathematics is the study of *numbers* and their properties. But what are numbers but abstractions from the world around us? The *twoness* of a pair of apples or a pair of oranges provides a starting point for a satisfactory definition of number! So perhaps we had better settle on the statement: Mathematics is the science of *abstraction*. For example, the equation of a parabola can yield the path of a projectile or the shape of a reflector on a motor headlamp. Were it not for the abstracting process, i.e., mathematization, we might not have recognized the essential identity of many apparently different phenomena in the world around us.

Let us return to the most elementary mathematics and write out the laws governing the processes of *addition* and *multiplication*.

.11 Addition Take the set of all integral numbers, including zero. If a, b are integral, then

(1)	$a + b = c$ is also integral	
(2)	$(a + b) + c = a + (b + c)$	(the associative law)
(3)	$a + 0 = 0 + a = a$	
(4)	$a + (-a) = (-a) + a = 0$	

.12 Multiplication Take the set of all rational numbers. If a, b are rational, then

(1)	$a \times b = c$ is also rational	
(2)	$(a \times b) \times c = a \times (b \times c)$	(the associative law)
(3)	$a \times 1 = 1 \times a = a$	
(4)	$a \times a^{-1} = a^{-1} \times a = 1$	

We observe also that both addition and multiplication are *commutative*.

The important thing to note is that except for the interchanges of signs + and ×, 0 and 1, *the laws are the same in both cases*. If we abstract again, we may set up a more general system called an

.13 Abstract Group $\mathfrak{g}$ This has as elements $G1$, $G2$,..., with a law of combination indicated by (.), such that if Gi, Gj are elements of $\mathfrak{g}$, then:

(1) $G_i.G_j = G_k$ is also an element of $\mathfrak{g}$
(2) $(G_i.G_j).G_k = G_i.(G_j.G_k)$ (the associative law)
(3) $G_i.G_1 = G_1.G_i = G_i$
(4) $G_i.G_i^{-1} = G_i^{-1}.G_i = G_1$

The number of elements Gi in $\mathfrak{g}$ is called the order g of $\mathfrak{g}$, and $G1$ the identity element of $\mathfrak{g}$. We do not assume commutativity; if, however, $Gi.Gj = Gj.Gi$, then the group is called *Abelian* after the mathematician Abel who first studied such a system. Could we distinguish the two Abelian groups of addition and multiplication in some way?

Suppose we define the operation S as that of adding 1 to the number 0; then we have a 1−1 correspondence,

4.14 $n \leftrightarrow S^n$

and we can say that the additive group of the integers is *cyclically generated* by S. On the other hand, if we denote multiplication by the prime p by the operator Sp, then every rational number can be uniquely represented in the form

4.15 $\ldots (S_p)^\alpha (S_q)^\beta (S_r)^\gamma \ldots 1$

for suitably chosen p, q, r, ... and α, β, γ, ... This is the fundamental theorem of arithmetic. The multiplicative group of the rational numbers is generated by an *infinite* number of independent generators Sp. The order in which these are applied is unimportant, so that the group is Abelian and the direct product of cyclical groups $\{Sp\}$, $\{Sq\}$,... .

One might imagine that these simple abstractions are very old, but they were first stated in this form by Cayley only a little more than 100 years ago. Those groups which we have considered so far have been of infinite order. Let us turn now to some familiar geometrical figures whose groups of rotations are of finite order.

.16 *The Groups of the Regular Solids* Symmetry of form appealed very much to the Greeks. In Plato's cosmogony, atoms of earth are represented as cubes; of fire, as tetrahedra; of air, as octahedra; and of water, as icosahedra. The dodecahedron seems to have symbolized the universe. This mystical interweaving of ideas was characteristic of Greek philosophy in which abstractions such as "the good" and "the beautiful" were objects of constant thought and discussion. The surprising thing is that having gone so far, the Greeks did not take the next step and abstract the notion of a *group*.

If we think of the rotations of a regular tetrahedron about axes (1) through

the midpoints of any opposite edges, (2) through any vertex and the centroid of the opposite face, preserving the *positions* occupied by the vertices though the latter may be interchanged, we find:

$$
\begin{aligned}
&3 \text{ rotations of type (1)} \\
&8 = 4 \times 2 \text{ rotations of type (2)} \\
&\underline{1} \text{ the identity (i.e., no rotation at all)} \\
&12
\end{aligned}
$$

and we say that the group of rotations of the tetrahedron is of order 12. It is worth verifying that the group of rotations of the octahedron, or the cube, is of order 24; while that of the icosahedron, or the dodecahedron, is of order 60.

There is one feature of all the groups we have considered so far which is important, namely, their *discreteness*. If we think of the motions of a chair over the floor, we can suppose the chair moved from A to B and denote this by MAB; if it is then moved from B to C we could imagine it moved directly from A to C and write
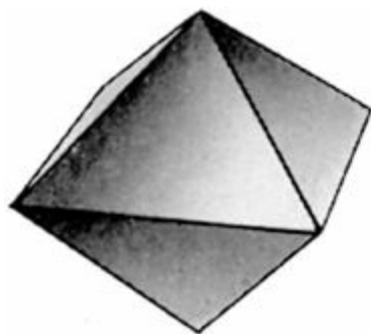
Octahedron
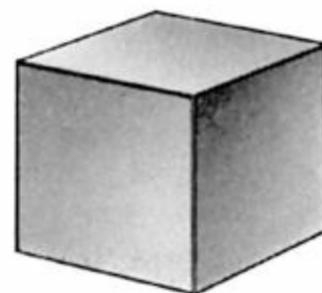
Cube
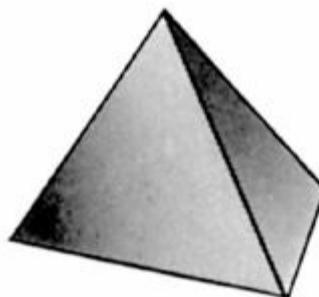
Tetrahedron
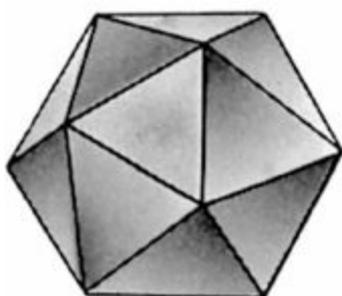
Icosahedron
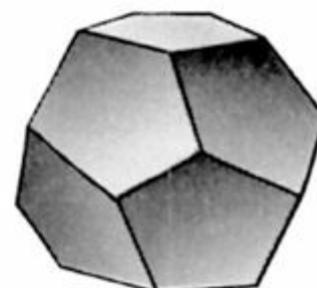
Dodecahedron

FIG. 4.1

$$M_{AB}M_{BC} = M_{AC}$$

These operations satisfy all our postulates if we let the identity operation be *MAA*, the operation of not moving the chair at all! But here there is a difference from the groups previously considered, since we could choose *B* as close to *A* as we like. This introduces the notion of continuity, and such a group is said to be *continuous*.

4.2 THE SYMMETRIC GROUP S_n

In Chapter 2 we saw the significance of permutations in defining determinants and deriving their properties. In particular, in 2.24 we gave the 3! permutations

on 3 symbols, and it is easy to verify that they form a group denoted by $\mathfrak{S}_3$. Since

$$(12)(23) = \begin{pmatrix} 123 \\ 132 \\ 231 \end{pmatrix} = (123), \qquad (23)(12) = \begin{pmatrix} 123 \\ 213 \\ 312 \end{pmatrix} = (132)$$

we conclude that $\mathfrak{S}_3$ is *non-Abelian*, and this is true of $\mathfrak{S}_n$ of order $n!$ for all $n > 2$.

If we define a *subgroup* of a given group $\mathfrak{g}$ as a subset of the g elements which satisfy the conditions 4.13, it can be verified that the following list of subgroups of $\mathfrak{S}_3$ is exhaustive:

4.21	I	of order 1
4.22	$I, (12); I, (13); I, (23)$	each of order 2
4.23	$I, (123), (132)$	of order 3

If $\mathfrak{K}$ is any subgroup of $\mathfrak{g}$ with elements $L, M, N, \ldots$, then we define the *transforms* of these elements by an element A of $\mathfrak{g}$ to be $ALA^{-1}, AMA^{-1}, ANA^{-1}, \ldots$, omitting the (.) signifying the law of combination. If $LM = N$, then

$$(ALA^{-1})(AMA^{-1}) = ALMA^{-1}$$

so that these transforms constitute a subgroup *conjugate* to $\mathfrak{K}$ which we may denote $A\mathfrak{K}A^{-1}$. The three subgroups 4.22 are all conjugates of one another, e.g.,

$$4.24 \qquad\qquad (13)I(13) = I, \qquad (13)(12)(13) = (23)$$

If $A\mathfrak{K}A^{-1} = \mathfrak{K}$ for all A in $\mathfrak{g}$, $\mathfrak{K}$ is said to be self-conjugate or *normal* in $\mathfrak{g}$. To say that $A\mathfrak{K}A^{-1} = \mathfrak{K}$ means that not each element of $\mathfrak{K}$ but only $\mathfrak{K}$ as a set is invariant under transformation by A. For example,

$$4.25 \qquad\qquad (12)I(12) = I, \qquad (12)(123)(12) = (132)$$
$$(12)(132)(12) = (123)$$

and similarly for transformation by (13) and (23), so that the *cyclic* subgroup $\mathfrak{K}$ = I, (123), (132) of order 3 is normal in $\mathfrak{S}_3$. No subgroup 4.22 of order 2 is normal, while the identity subgroup is always normal in any group.

In Chapter 2 we divided the $n!$ permutations on n symbols into two sets, one consisting of all the *even* permutations and the other of all the *odd* permutations. From the definition, the product of two even permutations is an even permutation. Moreover, I is even, and the inverse of an even permutation must

also be even, since their product is I. Thus all the even permutations on n symbols form a subgroup of $\mathcal{S}_n$ called the *alternating group*, denoted $\mathcal{Q}_n$.

On the other hand, the odd permutations do not form a subgroup of $\mathcal{S}_n$. If we multiply any even permutation of $\mathcal{Q}_n$ by a single transposition, say (12), we obtain an odd permutation, e.g.,

$$\mathcal{Q}_3 = I, (123), (132)$$

and

$$\mathcal{Q}_3(12) = (12), (13), (23)$$

It remains to show that the number of odd permutations is exactly equal to the number of even permutations. This will follow if we can show that (i) if $E1$, $E2$ are even permutations and O is odd, then $E1O \neq E2O$ unless $E1 = E2$, and (ii) every odd permutation can be written in the form EO when E is even and O is a *fixed* odd permutation.

The proof of (i) is immediate, since if

$$E_1 O = E_2 O$$

then

$$E_1 O . O^{-1} = E_2 O . O^{-1}$$

so that

$$E_1 = E_2$$

assuming only that the axioms for a group are satisfied, as we have seen to be the case.

The proof of (ii) is equally easy. Let us assume $O1$ to be any odd permutation and let us suppose that

$$O_1 = EO$$

as required. Since we are dealing with a group, $O-1$ exists and

$$O_1 O^{-1} = EO . O^{-1} = E$$

As in (i), E is uniquely defined.

We conclude that the order of $\mathcal{Q}_n$ is $\frac{1}{2}n!$ Moreover, since

$$PEP^{-1}$$

is even for all E belonging to $\mathcal{Q}_n$ and *any* permutation P of $\mathcal{S}_n$ we conclude that $\mathcal{Q}_n$ is normal in $\mathcal{S}_n$, generalizing 4.25. We gather together all this information in

.26 *All even permutations on n symbols form a subgroup* $\mathcal{Q}_n$ *of* $\mathcal{S}_n$ *known as the alternating group. The order of* $\mathcal{Q}_n$ *is* $\frac{1}{2}n!$ *and* $\mathcal{Q}_n$ *is normal in* $\mathcal{S}_n$.

It can be proven that $\mathcal{Q}_n$ contains no normal subgroup other than I for $n \neq 4$. Such a group is said to be *simple*.

EXERCISES

1. Write out all 24 permutations of the four symbols 1,2,3,4.

2. Which ones of these are even? Verify that they form a subgroup $\mathcal{Q}_4$ of $\mathcal{S}_4$.

3. Determine all subgroups of $\mathcal{Q}_4$. Which ones of these are normal (a) in $\mathcal{Q}_4$, (b) in $\mathcal{S}_4$?

4.3 THE GROUP OF A SQUARE

Let us see if we can attach a geometrical significance to the notion of a group. To this end, consider the square in which the coordinates of the vertices AB' $A'B$ are as indicated in Figure 4.2. We say that such a geometrical figure has *symmetry*, but what precisely do we mean by this? Apart from a somewhat vague interpretation of the word, we can analyze our idea by saying that *symmetry* is characterized by the property of *invariance* under *reflection* and (or) *rotation*.
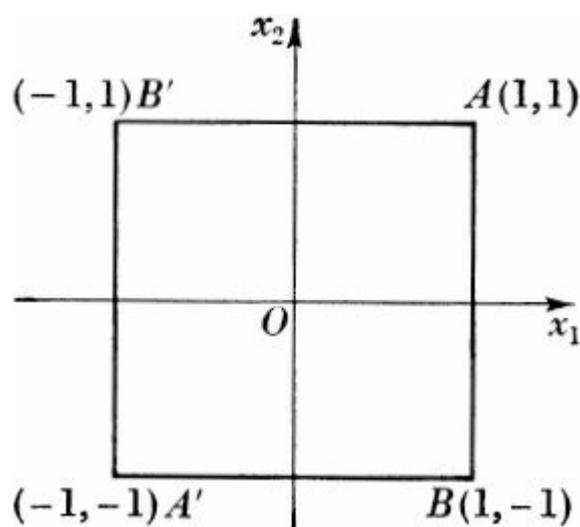

FɪG. 4.2

How could we describe the operation we have called *reflection*, say in the

coordinate axis $Ox2$? This can be done in two ways: either as the *linear transformation*

4.311 $$\left.\begin{array}{l} y_1 = -x_1 \\ y_2 = x_2 \end{array}\right\} \qquad Y = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} X$$

or relative to the square, as the *permutation*

4.312 $$(AB')(A'B)$$

Similarly, reflection in $Ox1$ can be written

4.321 $$\left.\begin{array}{l} y_1 = x_1 \\ y_2 = -x_2 \end{array}\right\} \qquad Y = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} X$$

or

4.322 $$(AB)(A'B')$$

The "product" of these two transformations in either order is easily seen to be

4.331 $$\left.\begin{array}{l} y_1 = -x_1 \\ y_2 = -x_2 \end{array}\right\} \qquad Y = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} X$$

or

4.332 $$(AB')(A'B).(AB)(A'B') = (AA')(BB') = (AB)(A'B').(AB')(A'B)$$

and this is a *rotation* about the origin of coordinates. It can be verified that the remaining symmetries of the square are the reflections

4.341 $$Y = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} X, \qquad Y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} X$$

or

4.342 $$(AA'), \qquad (BB')$$

and the rotations

4.351 $$Y = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} X, \qquad Y = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} X$$

or

$$4.352 \qquad (ABA'B'), \qquad (AB'A'B)$$

which, along with I,

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

make the eight "operations" of the group or the square.

Not all permutations on the four symbols A, B, A', B' have geometrical significance; e.g., (AB) does not correspond to any geometrical operation valid for all points in the plane. This definition of a group of operations under which a geometrical configuration remains invariant has wide application and great importance.

EXERCISES

1. Prove that the points $A(1,0)$, $B(-1/2, \sqrt{3/2})$, $C(-1/2, -\sqrt{3/2})$ are the vertices of an equilateral triangle.

2. By the method of 3.78, construct the linear transformations which effect the permutations (AB) and (AC) of the vertices of the triangle in Exercise 1. Thence, construct all rotations and reflections of the group of the triangle ABC.

4.4 ROTATIONS AND REFLECTIONS

Pursuing the line of thought of the preceding section, let us determine first the form of a *rotation* through an arbitrary angle θ about the origin. Having reference to Figure 4.3:

$$
\begin{aligned}
y_1 &= r \cos (\alpha + \theta) \\
&= r \cos \alpha \cos \theta - r \sin \alpha \sin \theta \\
&= x_1 \cos \theta - x_2 \sin \theta \\
y_2 &= r \sin (\alpha + \theta) \\
&= r \cos \alpha \sin \theta + r \sin \alpha \cos \theta \\
&= x_1 \sin \theta + x_2 \cos \theta
\end{aligned}
$$
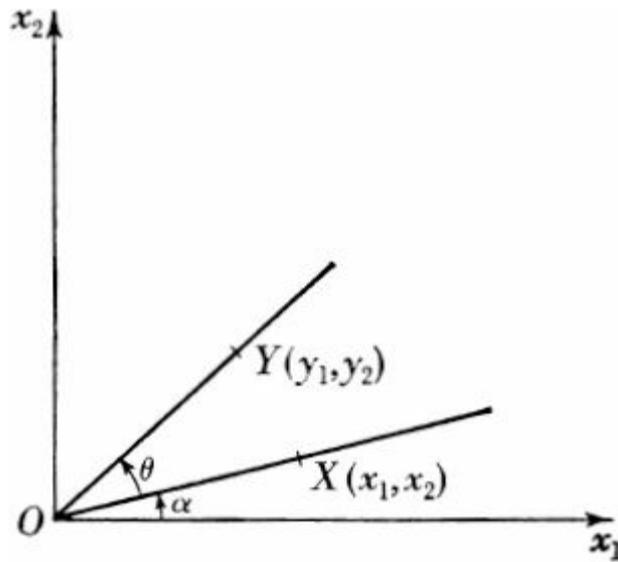
F<small>IG</small>. 4.3

which we can write in either one of the two ways:

4.41 $\qquad Y = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} X, \quad \text{or} \quad Y = X \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$

corresponding to 3.71 or 3.77. Observe that these formulas are independent of r and α, i.e., of the position of X. That the distance of a point from the center of rotation remains fixed is expressed in

4.42 $\qquad\qquad\qquad\qquad y_1^2 + y_2^2 = x_1^2 + x_2^2$

as the *invariance* of the *quadratic* form $x_1^2 + x_2^2$ under the transformation 4.41.

If we combine two rotations, A through θ and B through ϕ, as in 3.73 and 3.74 we obtain

4.43
$$Z = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} X$$
$$= \begin{pmatrix} \cos(\theta+\phi) & -\sin(\theta+\phi) \\ \sin(\theta+\phi) & \cos(\theta+\phi) \end{pmatrix} X$$

Since it would be enough if we wrote merely the *matrices* of the transformation, we are led to consider these as the elements of another group. We thus have three ways of describing the group of the square in Figure 2: (i) as a group of permutations, (ii) as a group of linear transformations, or (iii) as a group of matrices. There is an obvious one-to-one correspondence between any two of these groups, and they are said to be *isomorphic**

There is one important point which should be emphasized. *Matrices* must be

multiplied from left to right according to our definition 3.14. Consider the matrix product

$$\textbf{4.44} \qquad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$
$$\quad\quad\quad\;\; \pi_1 \qquad\qquad\quad \pi_2 \qquad\qquad\qquad \pi_3$$

If we write down the corresponding permutations $\pi_1 = (ABA'B')$ and $\pi_2 = (BB')$ of the preceding section, we ask the question, in what order should they be applied to yield the permutation $\pi_3 = (AB)(A'B')$? It is easy, in fact, to verify that π_2 must be applied *first* and π_1 *second* to yield π_3. If π_1 were applied first followed by π_2 we would obtain $(AB')(A'B)$ whose corresponding matrix is given in 4.311. This explains the reason for the convention of Section 2.2 that *permutations shall always be multiplied from right to left*.

Incidentally, we observe that neither the permutations nor the matrices representing them all commute, so that the group of the square is non-Abelian.

To construct the general form of a *reflection* in an arbitrary line through the origin, we refer to Figure 4.4.
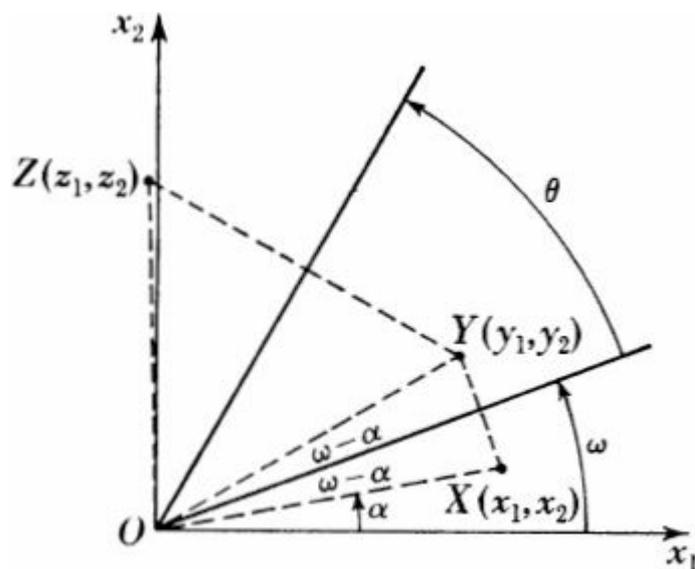

FIG. 4.4

$$y_1 = r \cos (2\omega - \alpha)$$
$$\quad = r \cos 2\omega \cos \alpha + r \sin 2\omega \sin \alpha$$
$$\quad = x_1 \cos 2\omega + x_2 \sin 2\omega$$
$$y_2 = r \sin (2\omega - \alpha)$$
$$\quad = r \sin 2\omega \cos \alpha - r \cos 2\omega \sin \alpha$$
$$\quad = x_1 \sin 2\omega - x_2 \cos 2\omega$$

which becomes in matrix form

4.45
$$Y = \begin{pmatrix} \cos 2\omega & \sin 2\omega \\ \sin 2\omega & -\cos 2\omega \end{pmatrix} X$$

If we reflect successively in two lines inclined to each other at an angle $\theta >$ 0, the second reflection could be written:

4.46
$$Z = \begin{pmatrix} \cos(2\omega + 2\theta) & \sin(2\omega + 2\theta) \\ \sin(2\omega + 2\theta) & -\cos(2\omega + 2\theta) \end{pmatrix} Y$$

so

4.47
$$Z = \begin{pmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix} X$$

proving that:

.48 *Successive reflection in two lines inclined at an angle $\theta > 0$ amounts to the same thing as rotating through 2θ about their common point.*

In particular, 4.311, 4.321, and 4.341 are special cases of 4.45, as is 4.331 of 4.47.

4.5 THE GROUP OF THE CUBE

All that we have said with reference to the square can easily be generalized to apply to the cube with vertices as indicated in Figure 4.5. Beginning with the reflections in the coordinate planes, we have

$$(AD')(BC')(CB')(DA'): \quad Y = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} X$$

$$(AC')(BD')(CA')(DB'): \quad Y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} X$$

$$(AB')(BA')(CD')(DC'): \quad Y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} X$$

Note that no *one* of the four transpositions multiplied together to yield one of these permutations is a symmetry of the cube—it does not belong to the group of the cube although it does belong to the larger group $\mathfrak{S}_8$ of all permutations on the eight symbols, of which the group of the cube is a subgroup.
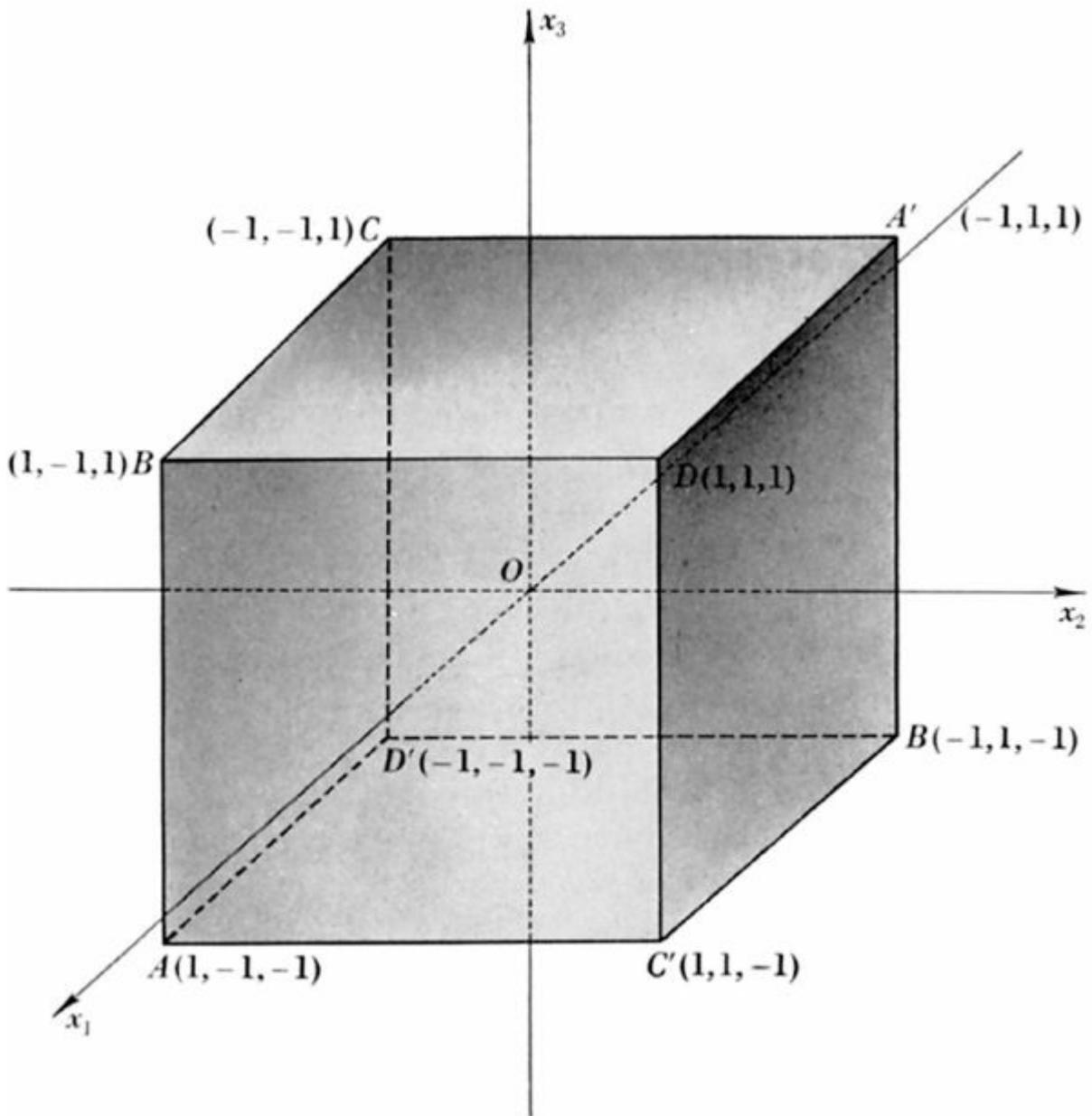

FIG. 4.5

Similarly, we may reflect in planes through pairs of opposite edges of the cube, which yields

$$(AB)(A'B'): \quad \mathbf{Y} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{X}$$

$$(CD)(C'D'): \quad \mathbf{Y} = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{X}$$

Each of these reflections, and all those others that leave the cube invariant, can be constructed by applying 4.45 in the appropriate manner, i.e., by considering a reflection in space to be in a *plane* through the origin. By combining such reflections in planes we could prove the following analogue of 4.48 in space:

.51 *Successive reflections in two planes inclined at an angle* $\theta > 0$ *amounts to the same thing as rotating through* 2θ *about their line of intersection.*

EXERCISE

1. Construct all 48 symmetries of the cube as permutations and also as linear transformations. How many of them leave invariant the regular tetrahedron *ABCD* in Figure 4.5? Prove that these form a subgroup of the group of the cube.

4.6 EULER'S FORMULA

In Figure 4.5 of the preceding section, imagine a sphere Σ drawn with center at the origin O, so as to pass through the vertices of the cube. If we project the edges of the cube from O into great circles on Σ, the resulting set of points and great circular arcs is called a *graph* and, including the faces, a *map* $\mathfrak{M}$ on Σ What we are interested in here is not the relationship of $\mathfrak{M}$ to the cube, which is well defined in space, but the relations between the elements of $\mathfrak{M}$ on Σ That certain of these relations remain invariant under a continuous *deformation* of the surface, suggests a new emphasis in geometry. The situation may be visualized by supposing that Σ is made of rubber and that, after the projection of the cube has been drawn on it, Σ is stretched and folded at will but not torn. The first such "topological" relation was found by Euler, but the subject was not put on a systematic basis until Poincaré's classic work of 1895.

If we denote the number of vertices of a spherical map $\mathfrak{M}$ by V, the number of edges by E, and the number of faces by F, then Euler's famous formula is that

$$V - E + F = 2$$

For 4.61 to hold for 𝔐, (i) the graph of 𝔐 must be *connected* (i.e., every vertex must be connected to every other vertex by a sequence of edges), (ii) no edge may intersect itself or any other edge except at a vertex, and (iii) every edge must be incident in exactly two faces.

The proof will be by induction. To start things off we verify the truth of 4.61 in the case of a map 𝔐0 containing *one* vertex, *one* edge, and *two* faces. Such a map is obtained by taking any circle σ in Σ and a point P on σ. The circle σ divides the sphere into two faces and is itself the one edge incident in P, the one vertex of the map.

To obtain a map 𝔐1 from 𝔐0 we may insert a new vertex in σ, thereby *increasing E and V* each by 1. The values of V, E, F for 𝔐1 would thus be $V = 2, E = 2, F = 2$.

The map 𝔐1 is more symmetrical than 𝔐0, in that we can add not only further vertices in the manner described above without increasing F, but also further edges, by joining pairs of vertices incident with the same face without increasing V. Such a new edge would divide the face in question so that E and F would each *increase* by 1. Each of these changes leaves 4.61 unaltered, so that any map obtainable by applying such changes in any order would satisfy 4.61 and the conditions (i) through (iii).

On the other hand, if we have given a spherical map 𝔐 satisfying the required conditions, we may successively remove edges which separate two faces, thereby *decreasing E* and F each by 1 until $F = 2$, always making sure that the graph remains connected. If this condition is satisfied at every stage, we may subsequently remove vertices, decreasing V and E each by 1, and reach the map 𝔐1, the conditions (ii) and (iii) remaining satisfied throughout the process. It follows that:

.62 *If a spherical map has V vertices, E edges, and F faces, and satisfies conditions* (i) *through* (iii) *above, then*

$$V - E + F = 2$$

4.7 THE REGULAR POLYHEDRA

In proving Euler's formula, we made no use of the regularity of the figure.

.71 Definition A convex polyhedron is said to be *regular* if all its faces are regular polygons, p edges surrounding each face and q meeting in each vertex.

The notion of "convexity" is important in mathematics. Here it means that no plane containing any face of the polyhedron penetrates the interior of the

polyhedron. It is natural to extend this definition of a regular polyhedron to a spherical map, and in so doing "convexity" is taken care of. The condition that p edges surround each face and q edges meet at each vertex can be written

4.72
$$pF = 2E = qV$$

since the edges are counted twice as they surround faces or meet in vertices. If we put these conditions into Euler's formula, we have

$$\frac{2E}{q} - E + \frac{2E}{p} = 2$$

or

4.73
$$\frac{1}{p} + \frac{1}{q} - \frac{1}{2} = \frac{1}{E}$$

In order to solve this Diophantine equation (i.e., an equation in more than one variable whose integral solutions are sought), we observe that $E > 0$. This limits the number of solutions to those given in the accompanying table. We recognize as corresponding to the last five solutions

4.74

	A	B	T	O	C	I	D
p	E	2	3	3	4	3	5
q	2	E	3	4	3	5	3
E	E	E	6	12	12	30	30
V	2	E	4	8	6	20	12
F	E	2	4	6	8	12	20

o f 4.73 the *tetrahedron, octahedron, cube, icosahedron, dodecahedron* of Figure 1. The first two solutions A and B are degenerate in the sense that we have a regular map for any value of E. The first has two vertices which may be taken as north and south poles, with the E lines of longitude, equally spaced if we wish, as edges. The second is obtained by taking $V = E$ points equally spaced around an equator and counting the E intervening arcs as edges.

There is a noticeable property of the equation 4.73, namely, that from any solution we can obtain another by interchanging p and q; such *dual* solutions coincide in the case of T, for which $p = q$.

If we take the vertices A, B, C, D of the cube in Figure 5 with coordinates

$$A(1,-1,-1), \quad B(-1,1,-1), \quad C(-1,-1,1), \quad D(1,1,1)$$

it is easy to verify that $AB = AC = AD = BC = BD = 2\sqrt{2}$. The equation of the

plane *ABC* is given by (cf. Exercise 1 of)

$$4.75 \qquad \begin{vmatrix} x_1 & x_2 & x_. & 1 \\ 1 & -1 & -1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \end{vmatrix} = 0 = \begin{vmatrix} x_1 + 1 & x_2 + 1 & x_3 - 1 \\ 2 & 0 & -2 \\ 0 & 2 & -2 \end{vmatrix}$$

and the equations of the other three faces of the regular tetrahedron are easily found.

EXERCISES

1. Construct a cardboard model of each of the five regular polyhedra.

2. Show that the *conjugate* tetrahedron *A'B'C'D'* has a face *A'B'C* parallel to *ABC*, by applying the symmetry of the cube:

$$\mathbf{X'} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \mathbf{X}$$

3. Verify that the points with coordinates

$$(0, \pm\tau, \pm1), \quad (\pm1, 0, \pm\tau), \quad (\pm\tau, \pm1, 0)$$

are the vertices of a regular icosahedron if

$$\tau^2 - \tau - 1 = 0$$

4. Verify that the points with coordinates

$$(0, \pm\tau^{-1}, \pm\tau), \quad (\pm\tau, 0, \pm\tau^{-1}), \quad (\pm\tau^{-1}, \pm\tau, 0), \quad (\pm1, \pm1, \pm1)$$

are the vertices of a regular dodecahedron.

5. Pick out the vertices of the five cubes which can be inscribed in the regular dodecahedron of Exercise 4.

4.8 POLYTOPES

Each regular polygon in the plane yields a regular, though degenerate, *polytope* as in the first column of the . One might well ask what regular polyhedra exist in a space of 4 dimensions—or, more generally, in *n*

dimensions? Curiously enough, there are *six* regular polyhedra in 4 dimensions, but for $n > 4$ there are only *three*, namely, the analogues of T, C, O.

It is easy to give the coordinates of the $2n$ vertices of the analogue of the octahedron in n dimensions:

4.81
$$(0, 0, \ldots, 0, \pm 1, 0, \ldots, 0, 0)$$

The analogue of the cube has the $2n$ vertices:

4.82
$$(\pm 1, \pm 1, \ldots, \pm 1)$$

where all combinations of sign are allowed. It is best to think of the analogue of the tetrahedron or regular *simplex* in n dimensions as consisting of the $n + 1$ points

4.83
$$(0, 0, \ldots, 0, 1, 0, \ldots, 0, 0)$$

in a space of $n + 1$ dimensions. That this figure really lies in a subspace of n dimensions is shown by the fact that its vertices satisfy the linear equation

4.84
$$x_1 + x_2 + \ldots + x_n + x_{n+1} = 1$$

It is awkward to describe the regular simplex in n dimensions by n coordinates, though this can be done also.

Two such figures make up the analogue of the octahedron in $n + 1$ dimensions, just as the corresponding two triangles

$$(1,0,0), \quad (0,1,0), \quad (0,0,1)$$

and

$$(-1,0,0), \quad (0,-1,0), \quad (0,0,-1)$$

make up the octahedron in three dimensions.

The group of symmetries of the regular simplex in n dimensions is the symmetric group $\mathfrak{S}_n + 1$ of order $(n + 1)!$, while that of the analogue of the cube or octahedron is the *hyperoctahedral* group of order $2^n n!$.

---

* This section is taken from a lecture given by the author to students in Australia and published in the Year Book (1959) of the Sydney University Science Association. Permission to quote is gratefully acknowledged.

* Two groups $\mathfrak{g}$, $\mathfrak{g}'$ are *isomorphic* if the one-to-one correspondence between the elements $Gi \leftrightarrow G'i$, $Gj \leftrightarrow G'j$ extends to their combinations
$$G_i.G_j = G_k \leftrightarrow G'_k = G'_i.G'_j$$
under the appropriate laws.

5

VECTORS AND VECTOR SPACES

5.1 BASIS VECTORS

In Chapter 1 we introduced the notion of a vector $X = \overrightarrow{OX}$ having n comonents $(x1, x2, \ldots xn)$, and the expression

$$\textbf{5.11} \qquad\qquad X = x_1 E_1 + x_2 E_2 + \ldots + x_n E_n$$

in terms of the basis vectors Ei. The following result is an important consequence of 5.11:

.12 *Any n + 1 vectors X, Y, ... Z which lie in an n-dimensional vector space υn must be linearly dependent.*
Proof. If we write each vector in the form 5.11,

$$X = x_1 E_1 + x_2 E_2 + \ldots + x_n E_n$$
$$Y = y_1 E_1 + y_2 E_2 + \ldots + y_n E_n$$
$$\cdot$$
$$\cdot$$
$$\cdot$$
$$Z = z_1 E_1 + z_2 E_2 + \ldots + z_n E_n$$

and multiply the first equation by the cofactor of X in the determinant

$$\Delta = \begin{vmatrix} X & x_1 & x_2 & \ldots & x_n \\ Y & y_1 & y_2 & \ldots & y_n \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ Z & z_1 & z_2 & \ldots & z_n \end{vmatrix}$$

the second equation by the cofactor of Y in Δ, and finally the last equation by the cofactor of Z in Δ and add, then on the left side we have the determinant Δ. On the right, the coefficient of Ei is just Δ with X, Y, ... Z replaced by xi, yi, ... zi, so that every such coefficient vanishes by 2.32. Thus the equation $\Delta = 0$ yields the desired linear relation between the vectors X, Y, ... Z so long as not all the

cofactors of X, Y, ... Z vanish. But in this excluded case linear dependence follows also in virtue of 2.52

Suppose now we start with a set of n linearly independent vectors $X1$, $X2$, ... Xn in υn. It follows from 5.12 that any vector X in υn can be expressed

$$5.13 \qquad\qquad X = x_1 X_1 + x_2 X_2 + \ldots + x_n X_n$$

We may describe $X1$, $X2$, ... Xn as *basis vectors* of υn and $(x1, x2, ... xn)$ as the components of X relative to this basis. Putting it otherwise, we might think of the vectors Xi as lying along n coordinate axes so that the *coordinates* of the point X would be $(x1\ x2\ ... xn)$. It is often convenient in analytical geometry to choose such *oblique axes*, since all the familiar intersection properties continue to hold. However, one must be careful not to interpret the Euclidean expressions for angle and distance in the usual way.

EXERCISES

1. Taking the vectors $X1(1,1,0,0)$, $X2(1,0,1,0)$, $X3(1,0,0,1)$, $X4(0,0,1,1)$ as basis, find the components of $V(2,1,3,4)$ by solving the vector equation

$$V = v_1 X_1 + v_2 X_2 + v_3 X_3 + v_4 X_4$$

for $\upsilon 1\ \upsilon 2$, $\upsilon 3$, $\upsilon 4$.

2. Explain how the linear transformation

$$X = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} Y$$

relates the vectors Ei to Xi. Construct the inverse transformation and there-from derive t he components of V relative to the Xi, as in Exercise 1.

3. Taking $e1(1,1)$, $e2(1,0)$ as basis vectors defining two oblique coordinate axes $Ox1$, $0x2$, make a drawing to show the positions of lines with equations

$$x_1 = 1, \quad x_2 = 1, \quad x_1 = x_2, \quad x_1 + x_2 = 1$$

Find all intersections of these lines *graphically* and verify by solving the appropriate equations.

4. Derive the linear relation connecting the four vectors $X1\ (1,1,0)$, $X2(1,0,1)$, $X3(1,0,0)$, $X4(0,0,1)$.

5. There may well be more than one linear relation connecting $n + 1$ vectors in υn. Find those connecting the vectors $X1(1,1,0)$, $X2(2,2,0)$, $X3(-1,0,0)$, $X4(2,0,0)$. What becomes of the equation $\Delta = 0$ in this case?

6. Express the number of linear relations in terms of the rank of the matrix obtained by omitting the first column of Δ.

5.2 GRAM-SCHMIDT ORTHOGONALIZATION PROCESS

Since the vectors E_i satisfy the two conditions

$$E_i \cdot E_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

they are said to constitute a *normal*, *orthogonal* basis of υn. In general, the Xi will not satisfy such conditions and the question arises, how can we construct a normal, orthogonal basis Yi from the Xi?

Suppose we set $Y1 = X1$ and $Y2 = X2 + cY1$ and require that

$$Y_2 \cdot Y_1 = 0 = X_2 \cdot Y_1 + c Y_1 \cdot Y_1$$

It follows that

$$c = -\frac{X_2 \cdot Y_1}{Y_1 \cdot Y_1}$$

so that

5.21
$$Y_2 = X_2 - \frac{X_2 \cdot Y_1}{Y_1 \cdot Y_1} Y_1$$

Again, set $Y3 = X3 + c2Y2 + c1Y1$ and require that

$$Y_3 \cdot Y_1 = 0 = X_3 \cdot Y_1 + c_1 Y_1 \cdot Y_1$$
$$Y_3 \cdot Y_2 = 0 = X_3 \cdot Y_2 + c_2 Y_2 \cdot Y_2$$

so that

5.22
$$Y_3 = X_3 - \frac{X_3 \cdot Y_2}{Y_2 \cdot Y_2} Y_2 - \frac{X_3 \cdot Y_1}{Y_1 \cdot Y_1} Y_1$$

The procedure can be repeated indefinitely so that from any basis $X1, X2, \ldots Xn$ we can always find an orthogonal basis $Y1\ Y2, \ldots Yn$. It should be emphasized, however, that such an *orthogonal* basis is by no means unique.

Example. If $X1(1,1,0)$, $X2(1,0,1)$, $X3(1,0,0)$ be the given basis, choose $Y1 = X1$

$$Y_2 = X_2 - \tfrac{1}{2}Y_1 = (\tfrac{1}{2}, -\tfrac{1}{2}, 1)$$
$$Y_3 = X_3 - \tfrac{1}{3}Y_2 - \tfrac{1}{2}Y_1 = (\tfrac{1}{3}, -\tfrac{1}{3}, -\tfrac{1}{3})$$

and these Yi's are pairwise orthogonal, as desired.

To pass from an orthogonal basis to a *normal* orthogonal basis, it is necessary only to multiply each vector Yi, by the scalar $1/|\ Yi\ |$. Of course, we could have normalized at each successive stage so that each of the denominators in 5.21 and 5.22 would have been 1.

EXERCISES

1. Instead of setting $Y1 = X1$ in the preceding example, set $Y1 = X3$ and complete the construction of an orthonormal basis in the usual way.

2. Set $Y1 = X3$, $Y2 = X1 - X3$ and find $Y3$ so that $Y1, Y2, Y3$ is an orthonormal basis in the example.

3. Construct vectors $Y3, Y4$ so that along with $Y1 = X1$, $Y2 = X4$ the Yi's constitute an *orthogonal* basis of the space of Exercise 1 of Section 5.1.

4. Using the Gram-Schmidt orthogonalization process, find an orthonormal basis Yi for the space defined by the vectors

$$X_1(1,0,1), \quad X_2(2,-1,1), \quad X_3(-1,-1,1)$$

5.3 THE VECTOR PRODUCT $U \times V$

In the special case $n = 3$ of the preceding section,

$$Y_3 \cdot Y_1 = Y_3 \cdot X_1 = 0$$
$$Y_3 \cdot Y_2 = Y_3(X_2 + cY_1) = Y_3 \cdot X_2 = 0$$

If we set $X1 = U$, $X2 = V$, $Y3 = W$, then the components of W must satisfy the scalar equations

$$W \cdot U = w_1 u_1 + w_2 u_2 + w_3 u_3 = 0$$
$$W \cdot V = w_1 v_1 + w_2 v_2 + w_3 v_3 = 0$$

so that (cf. 2.56)

5.31
$$w_1 : w_2 : w_3 = \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} : \begin{vmatrix} u_3 & u_1 \\ v_3 & v_1 \end{vmatrix} : \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}$$

Though the magnitude of $W \neq 0$ is not determined by these proportionalities, its direction is determined up to a factor of ± 1. It is customary to fix the positive direction, as in the case of the coordinate axes in Chapter 1, to be that of a "right-handed screw," under which U is rotated into V. This amounts to the following

.32 Definition The vector $W = U \times V$ has components

$$w_1 = \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix}, \qquad w_2 = \begin{vmatrix} u_3 & u_1 \\ v_3 & v_1 \end{vmatrix}, \qquad w_3 = \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}$$

It follows immediately that $(V \times U) = - (U \times V)$.

.33 $U \times V = 0$ *if and only if the rank of the matrix*

$$\begin{pmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{pmatrix} \text{ is 0 or 1.}$$

Proof. Certainly, $U \times V = 0$ if either $U = 0$ or $V = 0$ or $U = kV$ when $r = 1$; or if $U = V = 0$, then $r = 0$. Conversely, $U \times V = 0$ implies that one of these conditions must be satisfied.

In 3-space it is often convenient to set $i = E1 = (1,0,0), j = E2 = (0,1,0), k = E3 (0,0,1)$ so that, symbolically,

5.34
$$U \times V = \begin{vmatrix} i & j & k \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

For any W,

5.35
$$(U \times V) \cdot W = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = W \cdot (U \times V)$$

and by permuting U, V, W cyclically,

$$(U \times V) \cdot W = (V \times W) \cdot U = (W \times U) \cdot V$$
$$\{(U + U') \times V\} \cdot W = (U + U') \cdot (V \times W)$$
$$= U \cdot (V \times W) + U' \cdot (V \times W)$$
$$= (U \times V) \cdot W + (U' \times V) \cdot W$$
$$= \{(U \times V) + (U' \times V)\} \cdot W$$

Since this relation is true for all vectors W, we can choose $W = i, j, k$ in turn and, from the equal scalars on left and right, conclude that

5.36 $$(U + U') \times V = (U \times V) + (U' \times V)$$

Thus *vector multiplication is distributive.* It is to be noted that in general

$$(U \times V) \times W \neq U \times (V \times W)$$

so that vector multiplication is *not* associative.

Since the square of the *magnitude* of $U \times V$ is given by the Lagrange identity

$$(u_2 v_3 - u_3 v_2)^2 + (u_3 v_1 - u_1 v_3)^2 + (u_1 v_2 - u_2 v_1)^2$$
$$= (u_1^2 + u_2^2 + u_3^2)(v_1^2 + v_2^2 + v_3^2) - (u_1 v_1 + u_2 v_2 + u_3 v_3)^2$$

we conclude that

5.371 $$|U \times V|^2 = |U|^2 |V|^2 - (U \cdot V)^2 = |U|^2 |V|^2 (1 - \cos^2 \theta)$$

so that

5.372 $$|U \times V| = |U||V| \sin \theta$$

where θ is the angle between U and V.

While the vector product is defined only in 3-space, we may arrive at more general ideas by observing that the Lagrange identity may be written in the form

5.373 $$|U \times V|^2 = \begin{vmatrix} U \cdot U & U \cdot V \\ V \cdot U & V \cdot V \end{vmatrix}$$

Since 5.372 leads to the interpretation of $|U \times V|$ as the *area* of the parallelogram formed by the vectors U and V, it is tempting to coniine our attention to this plane π and suppose that

$$U = X(x_1, x_2), \qquad V = Y(y_1, y_2)$$

relative to some orthonormal basis in π. With such an assumption,

5.38

$$\begin{vmatrix} U \cdot U & U \cdot V \\ V \cdot U & V \cdot V \end{vmatrix} = \begin{vmatrix} X \cdot X & X \cdot Y \\ Y \cdot X & Y \cdot Y \end{vmatrix}$$

$$= \begin{vmatrix} x_1^2 + x_2^2 & x_1 y_1 + x_2 y_2 \\ y_1 x_1 + y_2 x_2 & y_1^2 + y_2^2 \end{vmatrix} = \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}$$

$$= \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}^2$$

by 2.27. We conclude from 5.372 and 5.373 that

.39 *The area of the parallelogram defined by $U = X$ and $V = Y$ in π is given by*

$$|U \times V| = \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \qquad taken \ positive.$$

This remarkable simplification, which arises through consideration of the problem in a subspace of the proper dimension, *does* generalize, as we shall see shortly.

EXERCISES

1. Find the area of the parallelogram whose vertices are $(-1,1)$, $(0,0)$, $(1,3)$, $(2,2)$.

2. Find the components of $U \times V$ when $U = (1,1,0)$, $V = (1,0,1)$. Find also the area of the parallelogram defined by U and V.

3. Determine the fourth vertex of a parallelogram of which the first three vertices are

$$A(1,0,1), \quad B(-1,1,1), \quad C(2,-1,2)$$

How many such vertices are there? What is the area of the parallelogram in each case? What is the area of the triangle ABC?

4. Find the vector $W = U \times V$ when $U = (1, -1, 1)$ and $V = (1, 1, -1)$, and calculate the volume of the parallelepiped defined by U, V, and W. What is the volume of the tetrahedron $OUVW$?

5. Derive the more general form of the Lagrange identity,

$$(U \times V) \cdot (U' \times V') = \begin{vmatrix} U \cdot U' & U \cdot V' \\ V \cdot U' & V \cdot V' \end{vmatrix}$$

6. Prove the Jacobi identity

$$(U \times V) \times W + (V \times W) \times U + (W \times U) \times V = 0$$

5.4 DISTANCE BETWEEN TWO SKEW LINES

From the geometrical point of view, the vector product $U \times V$ is just the normal vector to the plane determined by U and V and, as such, it has arisen on numerous previous occasions. By assigning a definite direction to this normal according to the definition 5.32, we have introduced a refinement which is chiefly useful in mechanics and in the study of electrical phenomena. Nevertheless, it is worth utilizing these ideas to determine the shortest distance between two skew, nonparallel lines in 3-space.
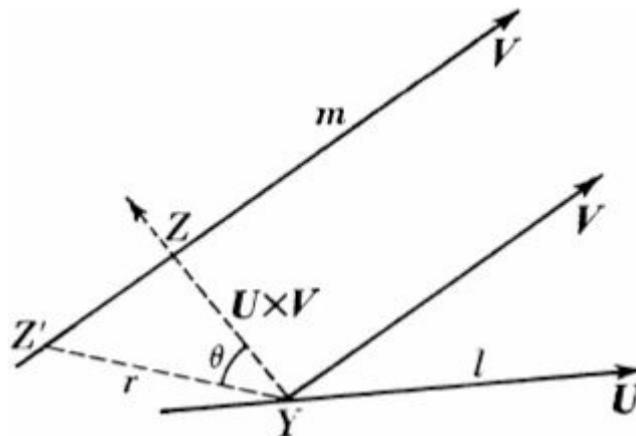
FIG. 5.1

Let us take the equations of l and m in parametric form to be

$$5.41 \qquad l: \begin{aligned} x_1 &= y_1 + su_1, \\ x_2 &= y_2 + su_2, \\ x_3 &= y_3 + su_3, \end{aligned} \qquad m: \begin{aligned} x_1 &= z_1 + lv_1 \\ x_2 &= z_2 + lv_2 \\ x_3 &= z_3 + lv_3 \end{aligned}$$

and let $r = \overrightarrow{YZ}$, so that

$$5.42 \qquad r \cdot (U \times V) = \begin{vmatrix} z_1 - y_1 & z_2 - y_2 & z_3 - y_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

By adding multiples of the second (third) row to the first, it follows from 5.41 that the choice of Y (Z) on I (m) does not affect the value of $r \cdot (U \times V)$. Thus if we take Y to be the foot of the common perpendicular and Z to be a variable point, the expression 5.42 set equal to zero yields the equation of a *plane* through l, parallel to m. Hence, for Z on m the required perpendicular distance δ between l and m is given by

5.43
$$\delta = \frac{r \cdot (U \times V)}{|U \times V|}$$

according to 1.68. An alternative approach to the same result is indicated by writing

5.44
$$\delta = |r| \cos \theta = |r| \frac{r \cdot (U \times V)}{|r||U \times V|} = \frac{r \cdot (U \times V)}{|U \times V|}$$

where $Z = Z'$ is any point on m, as in Figure 5.1.

EXERCISES

1. Find the equations of a line drawn from the origin O to intersect each of the lines

$$l: \quad x_1 - x_2 + x_3 - 1 = 0 = 2x_1 - x_3 - 2$$
$$m: \quad x_1 - 2x_2 - 1 = 0 = x_1 - x_2 - x_3 + 5$$

(Hint: Find a plane through each of l, m which contains $\mathbf{0}$.)

2. Obtain the equations of the lines l, m in Exercise 1 in parametric form and also in symmetric form.

3. Find the shortest distance between the lines l and m in Exercise 1.

5.5 *n*-DIMENSIONAL VOLUME

At the end of Section 5.3 we saw that the area of a parallelogram defined by two vectors X, Y in $\upsilon 2$ is given by the determinant

5.51
$$\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}$$

In order to generalize this result we shall reconsider it from a slightly different

point of view.

If W is a vector coplanar with X, Y and such that

$$|W| = |Y| \cos \phi = |Y| \sin\left(\frac{\pi}{2} - \phi\right)$$

with

5.521 $\qquad\qquad W \cdot X = 0$

then

5.522 $\qquad W \cdot Y = |W| \cdot |Y| \cos \phi = |W|^2$

which we may write in detail as follows:

$$w_1 x_1 + w_2 x_2 = 0$$
$$w_1 y_1 + w_2 y_2 = |W|^2$$

Solving these equations, we have by Cramer's rule

5.53 $\qquad \dfrac{w_1}{\begin{vmatrix} 0 & x_2 \\ |W|^2 & y_2 \end{vmatrix}} = \dfrac{w_2}{\begin{vmatrix} x_1 & 0 \\ y_1 & |W|^2 \end{vmatrix}} = \dfrac{1}{\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}} = \dfrac{1}{\begin{vmatrix} x_1 & x_2 \\ w_1 & w_2 \end{vmatrix}}$
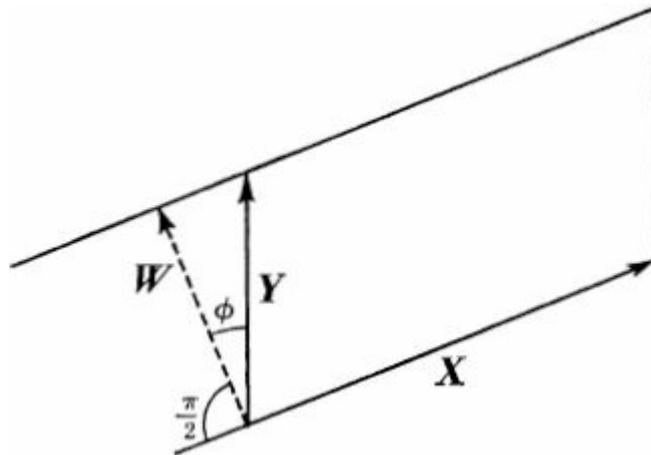
FIG. 5.2

the last fraction being obtained by multiplying numerator and denominator of the first by $w1$ and the second by $w2$, adding, and dividing out the factor $|W|2$.

It is important to realize that *area is invariant under rotation*. To see this, we suppose the axes transformed as in 4.41 so that

$$\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \begin{vmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{vmatrix} = \begin{vmatrix} x_1' & x_2' \\ y_1' & y_2' \end{vmatrix}$$

and also

$$\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = \begin{vmatrix} x_1 & x_2 \\ w_1 & w_2 \end{vmatrix} \begin{vmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{vmatrix} = \begin{vmatrix} x_1' & x_2' \\ 0 & w \end{vmatrix} = wx_1'$$

where θ is determined by the condition that

$$w_1 \cos\theta - w_2 \sin\theta = 0$$

and $w = w1 \sin\theta + w2 \cos\theta$. Since wx_1' is the usual expression for the area of the parallelogram in question, we may use induction to obtain the required generalization of 5.51, if we assume that *volume and generalized volume are similarly invariant under rotation*. Further properties and applications of the important "orthogonal" transformation involved here will be found in Chapter 9.

As before, let W be a vector cospatial with X, Y, Z and such that $|W| = |Z|$ cos ø,

5.541
$$W \cdot X = 0$$

5.542
$$W \cdot Y = 0$$

and

5.543
$$W \cdot Z = |W| \cdot |Z| \cos\phi = |W|^2$$

Writing out these equations in full:

$$w_1 x_1 + w_2 x_2 + w_3 x_3 = 0$$
$$w_1 y_1 + w_2 y_2 + w_3 y_3 = 0$$
$$w_1 z_1 + w_2 z_2 + w_3 z_3 = |W|^2$$

we solve to obtain

$$\frac{w_1}{\begin{vmatrix} 0 & x_2 & x_3 \\ 0 & y_2 & y_3 \\ |W|^2 & z_2 & z_3 \end{vmatrix}} = \frac{w_2}{\begin{vmatrix} x_1 & 0 & x_3 \\ y_1 & 0 & y_3 \\ z_1 & |W|^2 & z_3 \end{vmatrix}} = \frac{w_3}{\begin{vmatrix} x_1 & x_2 & 0 \\ y_1 & y_2 & 0 \\ z_1 & z_2 & |W|^2 \end{vmatrix}}$$

5.55
$$= \frac{1}{\begin{vmatrix} x_1 & x_2 & x_3 \\ x_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}} = \frac{1}{\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ w_1 & w_2 & w_3 \end{vmatrix}}$$

Again the denominator of the last fraction is independent of Z, and after a suitable transformation we may suppose that $W' = (0,0,w)$, so that this denominator becomes

$$w \begin{vmatrix} x_1' & x_2' \\ y_1' & y_2' \end{vmatrix}$$

Since this is, by definition, the volume of the parallelepiped defined by the vectors X, Y, Z, we conclude that this volume is given by the determinant

5.56
$$\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}$$

in every case.

Having established the basis of our induction and the definition of a generalized volume in υn in terms of that in $\upsilon n{-}1$, we have proved the following

*.57 **Theorem** The generalized volume of the generalized parallelepiped defined by n linearly independent vectors Xi ($i = 1,2, \dots n$) in υn with components (Xi1, Xi2, … Xin) is given by the determinant*

$$\begin{vmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{vmatrix}$$

It follows from its determinantal expression that the generalized volume is zero if the n vectors are linearly dependent. On the other hand, if the n linearly independent vectors lie in υm ($m > n$), the matrix of components is no longer square and the determinant is no longer defined. As in 5.38, however, we observe that

$$
5.58 \quad
\begin{vmatrix}
X_1 \cdot X_1 & X_1 \cdot X_2 & \cdots & X_1 \cdot X_n \\
X_2 \cdot X_1 & X_2 \cdot X_2 & \cdots & X_2 \cdot X_n \\
\cdot & & & \\
\cdot & & & \\
\cdot & & & \\
X_n \cdot X_1 & X_n \cdot X_2 & \cdots & X_n \cdot X_n
\end{vmatrix}
$$

$$
=
\begin{vmatrix}
x_{11} & x_{12} & \cdots & x_{1n} \\
x_{21} & x_{22} & \cdots & x_{2n} \\
\cdot & & & \\
\cdot & & & \\
\cdot & & & \\
x_{n1} & x_{n2} & \cdots & x_{nn}
\end{vmatrix}
\begin{vmatrix}
x_{11} & x_{21} & \cdots & x_{n1} \\
x_{12} & x_{22} & \cdots & x_{n2} \\
\cdot & & & \\
\cdot & & & \\
\cdot & & & \\
x_{1n} & x_{2n} & \cdots & x_{nn}
\end{vmatrix}
$$

and the determinant of inner products 5.58 is defined in υm for every $m \geq n$. Thus:

.59 *The square of the generalized volume of the generalized parallelepiped defined by n linearly independent vectors Xi ($i = 1,2, \ldots n$) in υm ($m \geq n$) is given by the determinant*

$$
\begin{vmatrix}
X_1 \cdot X_1 & X_1 \cdot X_2 & \cdots & X_1 \cdot X_n \\
X_2 \cdot X_1 & X_2 \cdot X_2 & \cdots & X_2 \cdot X_n \\
\cdot & & & \\
\cdot & & & \\
\cdot & & & \\
X_n \cdot X_1 & X_n \cdot X_2 & \cdots & X_n \cdot X_n
\end{vmatrix}
$$

As before, this determinantal expression vanishes if the n vectors are linearly dependent.

EXERCISES

1. Find the volume of the parallelepiped defined by the three vectors $X(0,0,1)$, $Y(0,1,2)$, $Z(1,2,3)$.

2. Make a drawing of the figure in Exercise 1; find the area A of the parallelogram in the plane $x1 = 0$, and calculate the required volume by means of the formula $V = A|Z| \cos\theta1$.

3. Find the volume of the parallelepiped defined by the three vectors $X(0,0,1,1)$, $Y(0,1,2,2)$, $Z(1,2,3,3)$.

4. Find the area of the face defined by X and Y in Exercise 3, and using the method of Exercise 2 calculate the required volume.

5. What is the volume of the tetrahedron *OXYZ* in Exercises 1 and 3?

6. Could you set up an induction which would yield the volume of an n-dimensional simplex in terms of the volume of the generalized parallelepiped in 5.59?

5.6 SUBSPACES OF υ*n*

Though a geometrical entity such as the *volume* of a figure may not change its value, its description relative to the space in which it is embedded may change, as we have seen. We have this phenomenon arising in a simpler form in the case of a line, which is defined by one linear equation in 2-space, by two linear equations in 3-space, … by $n - 1$ linear equations in *n*-space. Thus the *dimension* of the space in which a geometrical figure is embedded is important in describing it analytically.

Let $X1, X2, \ldots Xm$ be any m linearly independent vectors of υ*n* so that $m \le n$ by 5.12. If $0 < m < n$, we say that these m vectors define a subspace υ*m* of υ*n* made up of all vectors

$$5.61 \qquad\qquad X = \sum_{i=1}^{m} a_i X_i$$

Any vector Y which cannot be written in this form does not belong to υ*m*. Since there are just n linearly independent vectors in υ*n*, we may choose as basis

$$5.62 \qquad\qquad X_1, X_2, \ldots X_m, \quad Y_1, Y_2, \ldots Y_{n-m}$$

and we may suppose that the Gram-Schmidt orthogonalization process has been applied so that they are all pairwise orthogonal. Since every $Xi \cdot Yj = 0$, it follows that

$$5.63 \qquad\qquad \left(\sum_{i=1}^{m} a_i X_i \right) \cdot \left(\sum_{1}^{n-m} b_j Y_j \right) = 0$$

and *every vector in* υ*m is orthogonal to every vector in the subspace* υ*n–m defined by the vectors Y1, Y2, … Yn–m.*

The two subspaces υ*m* and υ*n–m* are said to be *orthogonal complements* of each other relative to υ*n*

Example. For $n = 2$, any two distinct vectors are linearly independent, and if orthogonal they are complementary relative to υ2. But two orthogonal vectors

are not complementary in a three-dimensional vector space $\upsilon 3$, whereas a plane and its normal vector define complementary subspaces in $\upsilon 3$. In $\upsilon 4$ the orthogonal complement of a $\upsilon 1$ is a $\upsilon 3$ and that of a $\upsilon 2$ is another $\upsilon 2$.

Just as we broke down υn into orthogonal complementary subspaces υm and $\upsilon n–m$, so we could break down υm. The particular case in which $\upsilon 2n$ is broken down into n pairwise orthogonal planes, or $\upsilon 2n+1$ into n pairwise orthogonal planes and a line orthogonal to each plane, is of special interest. If we think of a rotation in each such plane about the complementary subspace we have, taking them all together, the most general rotation in $\upsilon 2n$ or $\upsilon 2n+1$, as we shall see later on.

Two subspaces υr and υs of υn may intersect. We define the subspace made up of all vectors common to υr and υs as the *intersection* $\upsilon r \cap \upsilon s$, and the subspace made up of all vectors linearly dependent on vectors of υr and υs as the *union* $\upsilon r \cup \upsilon s$. If we denote the dimensions of $\upsilon r \cap \upsilon s$ and $\upsilon r \cup \upsilon s$ by $d(\upsilon r \cap \upsilon s)$ and $d(\upsilon r \cup \upsilon s)$, then

$$5.64 \qquad d(\mathcal{V}_r \cup \mathcal{V}_s) + d(\mathcal{V}_r \cap \mathcal{V}_s) = r + s$$

Proof. If υr and υs have no vectors in common, then $d(\upsilon r \cap \upsilon s) = 0$ and the number of linearly independent vectors in $\upsilon r \cup \upsilon s$ is just $r + s$, as claimed. If, however, $d(\upsilon r \cap \upsilon s) > 0$, then we may suppose that the $d(\upsilon r \cap \upsilon s) = d$ linearly independent vectors $Z1\ Z2, \dots Zd$ of $\upsilon r \cap \upsilon s$ form part of the basis of each of υr and υs:

$$5.65 \qquad \overbrace{X_1, X_2, \dots X_{n-d},}^{\mathcal{V}_r} \underbrace{Z_1, Z_2, \dots Z_d,}_{} \; Y_1, Y_2, \dots Y_{s-d}$$
$$\underbrace{\phantom{Z_1, Z_2, \dots Z_d, Y_1, Y_2, \dots Y_{s-d}}}_{\mathcal{V}_s}$$

so that in the enumeration every vector Zi is counted twice on each side of 5.64, proving the result.

If we replace υr and υs by their orthogonal complements $\upsilon n–r$ and $\upsilon n–s$, 5.64 becomes

$$5.66 \qquad [n - d(\mathcal{V}_r \cup \mathcal{V}_s)] + [n - d(\mathcal{V}_r \cap \mathcal{V}_s)] = (n - r) + (n - s)$$

as was to be expected.

5.7 EQUATIONS OF A SUBSPACE

From the point of view of analytical geometry, all the subspaces of υn pass through the origin and so are defined by one or more *homogeneous* linear equations. For example, υm is defined by the $n − m$ linear equations

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0$$
$$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0$$

5.71

$$\cdot$$
$$\cdot$$
$$\cdot$$

$$a_{n-m,1}x_1 + a_{n-m,2}x_2 + \ldots + a_{n-m,n}x_n = 0$$

To see that this is so, it is only necessary to adjoin the *m* further equations

$$x_{n-m+1} = t_1$$
$$x_{n-m+2} = t_2$$
$$\cdot$$
$$\cdot$$
$$\cdot$$
$$x_n \qquad = t_m$$

and solve by Cramer's rule to obtain (x1, x2, … xn) expressed in terms of the *parameters* t1 t2, … tm. Since these expressions are linear, just *m* solutions are linearly independent and we can associate them with the values of the parameters,

$$1,0,0, \ldots 0; \quad 0,1,0, \ldots 0; \quad \ldots ; \quad 0,0,0, \ldots 1$$

If we call the corresponding vectors X1 X2, … Xm, we have recovered υm as the *solution space* of the set of linear equations 5.71.

Example. Let us suppose that n = 3, and we wish to find a basis of the solution space of the equation

5.72 $\qquad\qquad\qquad 2x_1 - 3x_2 + x_3 = 0$

To this we adjoin the equations

$$x_2 = t_1, \qquad x_3 = t_2$$

so that

5.73 $\qquad\qquad\qquad x_1 = \tfrac{3}{2}t_1 - \tfrac{1}{2}t_2$

for all t1, t2. As above, the basis vectors of υ2 could be taken to be

5.74 $\qquad\qquad X_1 = (\tfrac{3}{2},1,0), \qquad X_2 = (-\tfrac{1}{2},0,1)$

Certainly these are linearly independent and any solution of 5.72 defines a vector

$$X = t_1 X_1 + t_2 X_2$$

Conversely, every such vector yields a solution of 5.72.

Thus the dimension of the solution space is 2, i.e., 5.72 represents a plane $\upsilon 2$ through the origin containing the vectors $X1$, $X2$ and every vector linearly dependent on them. These vectors form a basis of $\upsilon 2$ from which an orthonormal basis could be constructed by the Gram-Schmidt orthogonalization process.

If we wish to study the intersection $\upsilon 2 \cap \upsilon_2'$, where υ_2' is defined by the equation

5.75 $\hspace{4cm} x_1 + x_2 - x_3 = 0$

we should adjoin *one* further equation, $x3 = t1$ and solve 5.72 and 5.75 to obtain

5.76 $\hspace{3cm} x_1 = \tfrac{2}{5} l_1, \quad x_2 = \tfrac{3}{5} l_1, \quad x_3 = l_1$

Thus the intersection space has dimension 1 with basis vector $Z = (2,3,5)$. To bring this vector into evidence as in 5.65, we observe that it arises by setting $t1 = 3$, $t2 = 5$ in 5.73. Clearly, any vector of $\upsilon 2$ may be written in the form

5.77 $\hspace{3cm} x_1 = -l_1' + l_2', \quad x_2 = l_1', \quad x_3 = l_2'$

from which we obtain Z again by setting $l_1' = 3$, $l_2' = 5$. As a basis for $\upsilon 2 \cup \upsilon_2' = \upsilon 3$, we could choose

5.78 $\hspace{2cm} X_1 = (\tfrac{3}{2}, 1, 0), \quad Z = (2,3,5), \quad Y_1 = (1,0,1)$

since $Y1$ does not lie in $\upsilon 2$. We note in passing that the zero vector lies in every subspace, since the defining equations are all homogeneous.

5.8 ORTHOGONAL PROJECTION

If we take $E1 = (1,0)$, $E2 = (0,1)$ as basis vectors in $\upsilon 2$, then the coordinate axes are two mutually orthogonal subspaces and any vector

$$X = x_1 E_1 + x_2 E_2$$

is said to have *orthogonal projections* $x1 E1$ and $x2 E2$ on these axes. More generally, if $X1$ and $Y1$ are orthogonal, then any vector X may similarly be written

$$X = a_1 X_1 + b_1 Y_1$$

and X has orthogonal projections $aX1$ and $b1Y1$ on $X1$ and $Y1$.

If we take the basis 5.62 of υn, then any vector X of υn may be written in the form

$$X = \sum_{i=1}^{m} a_i X_i + \sum_{j=1}^{n-m} b_j Y_j$$

and we say that X has orthogonal projections

$$\sum_{i=1}^{m} a_i X_i \quad \text{on} \quad \upsilon_m$$

and

$$\sum_{j=1}^{n-m} b_j Y_j \quad \text{on} \quad \upsilon_{n-m}$$

Since any subspace of υn is defined in terms of its basis vectors, we may similarly speak of its orthogonal projections on υn and $\upsilon n\!-\!m$, but we shall not pursue the matter further.

EXERCISES

1. Find an orthogonal basis for vectors which lie in the plane $x1 - x2 + x3 = 0$. Find also the orthogonal projection of the vector $(1,0,1)$ on this plane and on the normal to the plane.

2. Find parametric equations and orthogonal basis vectors for the solution space of the system of equations

$$x_1 - x_2 - x_3 + 2x_4 = 0$$
$$2x_1 - x_2 - 2x_3 + x_4 = 0$$

3. Prove that the vector $(1,1,0,0)$ does *not* lie in the solution space of Exercise 2, and find its orthogonal projection (a) on that space, (b) on each of the primes represented by the two equations which define the space.

6

CONICS AND QUADRICS

6.1 CIRCLES AND SPHERES

So far we have concerned ourselves with linear properties of algebra and geometry, but Pythagoras' theorem suggests that we should study *quadratic* properties also, since the relation between the sides of a triangle is not between their lengths but between the *squares* of these lengths. In fact, this theorem yields the analytic definition of a *circle* of radius r and center $(a1, a2)$ to be

6.11
$$(x_1 - a_1)^2 + (x_2 - a_2)^2 = r^2$$

and that of a *sphere* with center $(a1, a2, a3)$ to be

6.12
$$(x_1 - a_1)^2 + (x_2 - a_2)^2 + (x_3 - a_3)^2 = r^2$$

Expanding these expressions we observe that (i) *the coefficients of the square terms* x_i^2 *are all equal*, and (ii) *no cross terms* $x_i x_j$ *appear*. Conversely, if these conditions are satisfied the squares may be completed and the equation takes the form 6.11 in the plane and 6.12 in space.

It follows immediately that circles and spheres, like lines and planes, have the property that any two define a *linear pencil*. Consider, for example, the two circles

$$C_a: \quad (x_1 - a_1)^2 + (x_2 - a_2)^2 - r_a^2 = 0$$
$$C_b: \quad (x_1 - b_1)^2 + (x_2 - b_2)^2 - r_b^2 = 0$$

Any linear combination

6.13
$$C_a + \rho C_b = 0$$

satisfies the conditions (i) and (ii) above, so that 6.13 represents a circle. Moreover, any point satisfying $C_a = 0 = C_b$ also satisfies $C_a + \rho C_b = 0$ so that if the two circles $C_a = 0$ and $C_b = 0$ intersect in P, Q then every circle of the pencil passes through P, Q. Two circles intersect in *two* points, which may coincide, or not at all since elimination leads to a quadratic equation which has two real roots or two complex roots. In the latter case no two circles of the pencil intersect. Such a linear pencil of circles is said to be *coaxal*, and we

illustrate the three possibilities in Figure 6.1.

Case 1   Case 2   Case 3
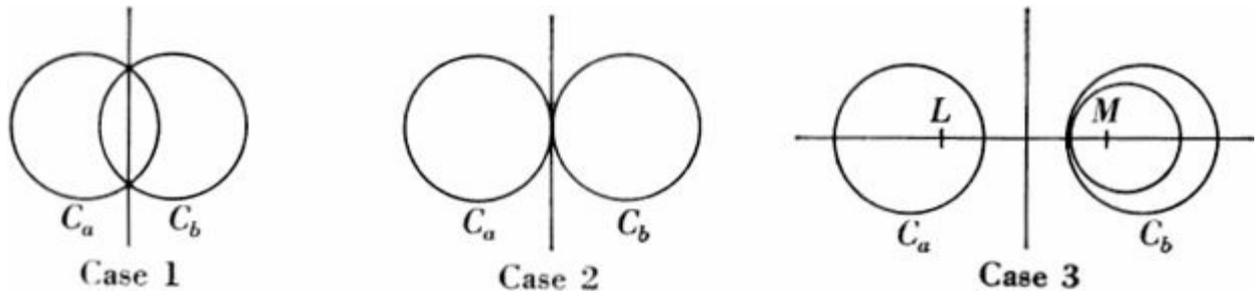
FIG. 6.1

If $\rho = -1$, 6.13 reduces to a linear equation which represents the line PQ in Case 1 and Case 2 and a line not meeting any circle of the coaxal pencil in Case 3; this line is called the *radical axis* of the system.

In Case 1 the diameter of a circle belonging to the linear pencil cannot be less than $|PQ|$, whereas no such restriction holds in Case 2 or Case 3. In the latter case we may let the radius tend to zero and define two *limit points L, M* which are equidistant from the radical axis. The following example illustrates Case 3.

Example. Take the two circles in Case 3 to be

$$C_a: \quad (x_1 - 2)^2 + x_2^2 = 1$$
$$C_b: \quad (x_1 + 1)^2 + x_2^2 = 1$$

The equation $C_a + \rho C_b = 0$ when simplified becomes

$$(1 + \rho)x_1^2 + (-4 + 2\rho)x_1 + (1 + \rho)x_2^2 + 3 = 0$$

and if we require the radius of such a circle to be zero, then $\rho^2 - 7\rho + 1 = 0$ so that $\rho = \frac{1}{2}(7 \pm 3\sqrt{5})$. It follows that the x_1 coordinates of the *limit points L, M* (circles of zero radius) are $-\frac{1}{2}(-1 + \sqrt{5}) = -0.62$ and $\frac{1}{2}(1 + \sqrt{5}) = 1.62$, with $x_2 = 0$ in each case. The equation of the radical axis is $x_1 = \frac{1}{2}$.

All this generalizes to spheres in space and indeed to hyperspheres in any number of dimensions. Any three circles, spheres, or hyperspheres of a coaxal system are linearly *dependent* in the sense we have so often used the term.

If we suppose that $a_1 = a_2 = 0$ in 6.11, we have the equation of a circle with center the origin,

6.14 $$x_1^2 + x_2^2 = r^2$$

Alternatively, we could have applied the linear transformation

$$6.15 \qquad \begin{aligned} y_1 &= x_1 - a_1 \\ y_2 &= x_2 - a_2 \end{aligned}$$

called a "parallel translation" or simply a *translation* to achieve the simplification of the equation. We can think of this as moving the circle 6.11 while keeping the axes fixed, or as a change of axes. We shall study such transformations in Chapter 9, but we note in passing that the equation 6.14 remains *invariant* under any rotation or reflection 4.41 or 4.45.

More generally, we may write the equation of a hypersphere in υn, with center the origin, in the form

$$6.16 \qquad X'IX \equiv x_1^2 + x_2^2 + \ldots + x_n^2 = r^2$$

Any linear transformation which leaves 6.16 invariant is said to be *orthogonal*, and we shall see in Chapter 9 that such a transformation may be thought of as a succession of rotations and reflections, when these are suitably defined. We prove here the important theorem

.17 *The necessary arid sufficient condition that a linear transformation* **Y** = **AX** *should be orthogonal is that A*–1 = *A*t.
Proof. If we assume that

$$Y'Y = X'A'AX = X'IX$$

then *A*t *A* = *I* and *A*–1 = *A*t as required. Conversely, this condition is sufficient as well as necessary for the invariance of the *quadratic form XtIX* = *XtX*.

EXERCISES

1. Prove that the centers of the circles of the coaxal system 6.13 are collinear and find the equation of the line in question.

2. Prove that the radical axis of a coaxal system of nonintersecting circles is the right bisector of *LM*, where *L, M* are the limit points of the system.

3. Prove that every circle with center on the radical axis and passing through the limit points cuts every circle of the system orthogonally.

4. Find the point of intersection *P* of the radical axes of the three circles *C*a, *C*b of the above example along with *C*c: $x_1^2 + (x2 - 2)2 - 1 = 0$, taken two by two. The point *P* is called the *radical center* of the system of circles. Prove that the circle with center *P*, passing through the limit points, cuts any circle

$$C_a + \rho C_b + \sigma C_c = 0$$

orthogonally.

5. Write the equation of a circle through the three points $(0,1)$, $(1,0)$, $(-2,3)$ as a determinant.

6. Prove that the equation of a circle through the three points $(0,1)$, $(1,0)$, $(\frac{1}{2},\frac{1}{2})$, when written in determinantal form, reduces to $x1 + x2 = 1$. Explain.

7. Write the equation of a sphere through the four points $(0,1,1)$, $(1,0,1)$, $(1,1,0)$, $(1,1,1)$ in determinantal form and find its radius.

8. If

$$1 + a^2 - b^2 - c^2 = \alpha, \qquad 1 - a^2 + b^2 - c^2 = \beta$$
$$1 - a^2 - b^2 + c^2 = \gamma, \qquad 1 + a^2 + b^2 + c^2 = \delta$$

verify that the matrix

$$A = \delta^{-1} \begin{pmatrix} \alpha & 2(ab + c) & 2(ac - b) \\ 2(ab - c) & \beta & 2(bc + a) \\ 2(ac + b) & 2(bc - a) & \gamma \end{pmatrix}$$

due to Rodrigues and Euler, is *orthogonal* for all real values of *a, b, c*.

9. Show by actual substitution that the linear transformation $Y = AX$, with $n = 3$, obtained by setting $a = b = 1$, $c = 0$ in Exercise 8, leaves XtX invariant.

6.2 CONICS IN CARTESIAN COORDINATES

Besides the circle in the plane, there are two other central *conics* called the *ellipse,*

6.21
$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1$$

and the *hyperbola,*

6.22
$$\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = 1$$

The matrix equations for these curves are easily seen to be

6.23
$$X^t \begin{pmatrix} \dfrac{1}{a^2} & 0 \\ 0 & \dfrac{1}{a^2} \end{pmatrix} X = 1, \quad X^t \begin{pmatrix} \dfrac{1}{a^2} & 0 \\ 0 & \dfrac{-1}{b^2} \end{pmatrix} X = 1$$

which correspond to 6.16 for a circle.

It is not our purpose here to investigate the properties of these *central conics* in great detail. Suffice it to say that they are *symmetrical* with respect to each coordinate axis and so with respect to the origin which is called the *center* of the conic. The ellipse 6.21 cuts each axis in two pairs of *vertices* ($\pm a$,0), (0, $\pm b$), the lines joining these pairs of vertices being called the *principle axes* of the conic. The length of the *semimajor* axis is a and that of the *semiminor* axis is b, if $a > b$.
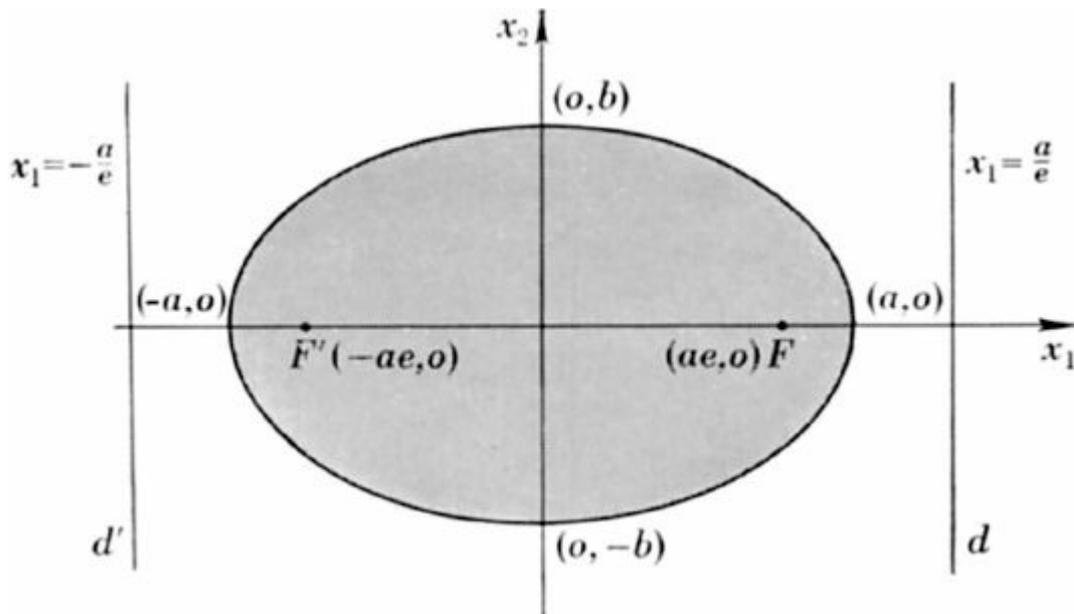
Fig. 6.2

In the case of the hyperbola there are two vertices ($\pm a$,0) on the *transverse* axis, but the curve does not meet the *conjugate* axis $x1 = 0$ in real points. The term *"principle axes"* is, however, still applicable.

If we think of an arbitrary point P on either an ellipse or an hyperbola and $\rho = \overrightarrow{OP}$ as a radius vector, then the vertices may be defined in terms of the magnitude of ρ. In the case of an ellipse, ρ takes its maximum value on the major axis and its minimum value on the minor axis, while in the case of an hyperbola ρ is a minimum on the transverse axis and has no maximum value.

There is one further conic called the *parabola*, whose equation in simplest form is

6.24
$$x_2^2 = 4px_1$$

The device we used before, of writing the equation in matrix form, is no longer applicable, but we shall see in a later chapter how our coordinate system can be modified to recover this convenience of expression.
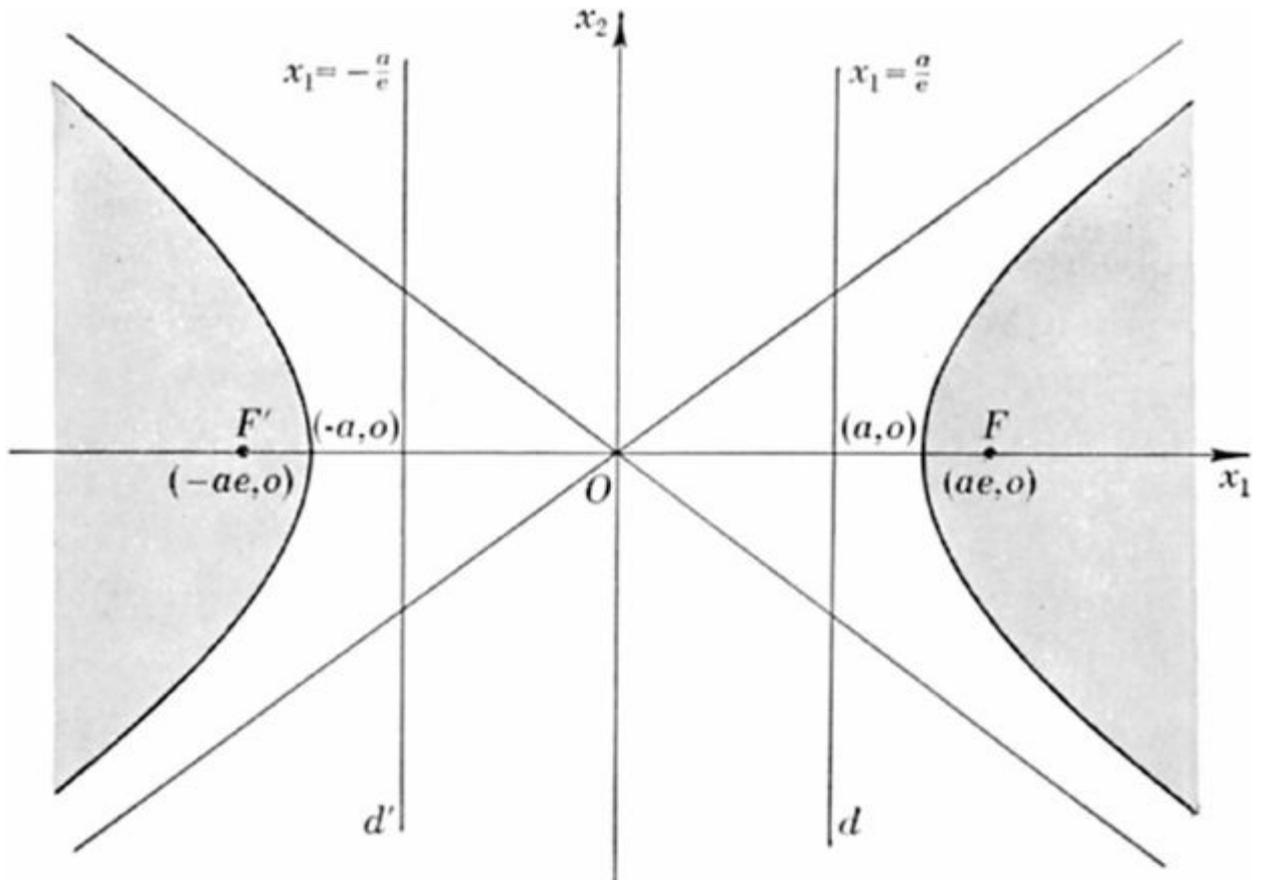
FIG. 6.3

Besides these analytical definitions, a conic can be defined synthetically. To this end we introduce the notion of the *eccentricity e*, where $e < 1$ for an ellipse, $e = 1$ for a parabola, and $e > 1$ for an hyperbola. The *foci F, F'* of an ellipse or an hyperbola have coordinates $(\pm ae, 0)$ and these, along with corresponding *directrices d, d'*, with equations $x1 = \pm a/e$, are indicated in Figures 6.2 and 6.3. In the case of the parabola there is only one focus $(p, 0)$ and one directrix $x1 = -p$, as in Figure 6.4.
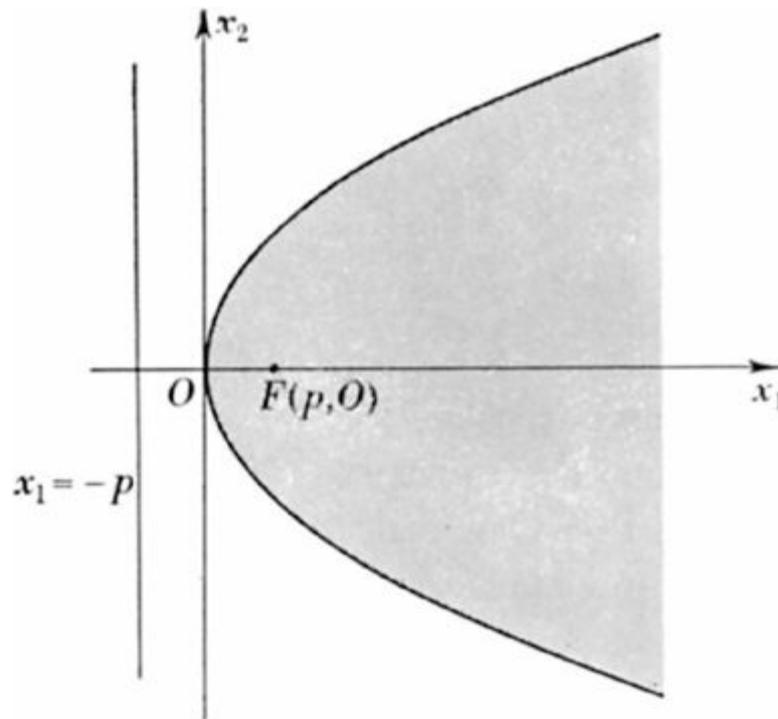
FIG. 6.4

EXERCISES

1. Derive the equation 6.24 of the parabola as the locus of points P equidistant from the focus F and the directrix d.

2. Find the equation of the parabola with focus the origin and directrix the line $x1 + x2 = 1$

3. Derive the equation 6.21 of the ellipse as the locus of points P such that PF is e times the perpendicular distance from P to the directrix d. Express b in terms of a and e. For what value of e is the locus a circle?

4. Show that the locus of Exercise 3 becomes the hyperbola 6.22 for $e > 1$.

5. Find the equation of an hyperbola with focus $(1,1)$ and corresponding directrix $x1 + x2 = 1$ for which $e = 2$.

6. Write the equation of a conic through the five points $(1,0)$, $(0,1)$, $(1,1)$, $(-2,0)$, $(0,-2)$ in determinantal form and expand to obtain the equation $x_1^2 - 2x1x2 + x_2^2 + x1 + x2 - 2 = 0$.

7. Prove that a conic is uniquely determined when *five* points on it are given.

6.3 QUADRICS AND THE LINES ON THEM

If we translate the origin to the point $(a1,a2,a3)$, the sphere 6.12 becomes

6.31
$$x_1^2 + x_2^2 + x_3^2 = r^2$$

The *normal form* of the equation of the *ellipsoid* (Figure 6.5) is

6.32
$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} + \frac{x_3^2}{c^2} = 1$$
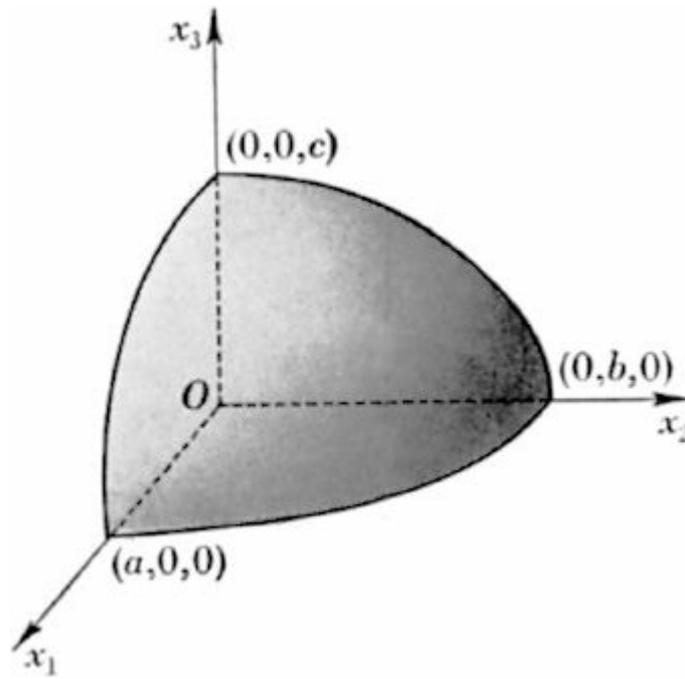
FIG. 6.5

while that of the *hyperboloid of one sheet* (Figure 6.6) is

6.33
$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} - \frac{x_3^2}{c^2} = 1$$

and that of the *hyperboloid of two sheets* (Figure 6.7) is

16.34
$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} - \frac{x_3^2}{c^2} = -$$

Each of these *quadric surfaces* is symmetrical with respect to each coordinate plane, and so with respect to the origin which is here the *center* of the quadric.
The matrix forms of the equations analogous to 6.23 are:

$$6.321 \qquad X^t \begin{pmatrix} \dfrac{1}{a^2} & 0 & 0 \\[2mm] 0 & \dfrac{1}{b^2} & 0 \\[2mm] 0 & 0 & \dfrac{1}{c^2} \end{pmatrix} X = 1$$

$$6.331 \qquad X^t \begin{pmatrix} \dfrac{1}{a^2} & 0 & 0 \\[2mm] 0 & \dfrac{1}{b^2} & 0 \\[2mm] 0 & 0 & -\dfrac{1}{c^2} \end{pmatrix} X = 1$$

$$6.341 \qquad X^t \begin{pmatrix} -\dfrac{1}{a^2} & 0 & 0 \\[2mm] 0 & -\dfrac{1}{b^2} & 0 \\[2mm] 0 & 0 & \dfrac{1}{c^2} \end{pmatrix} X = 1$$
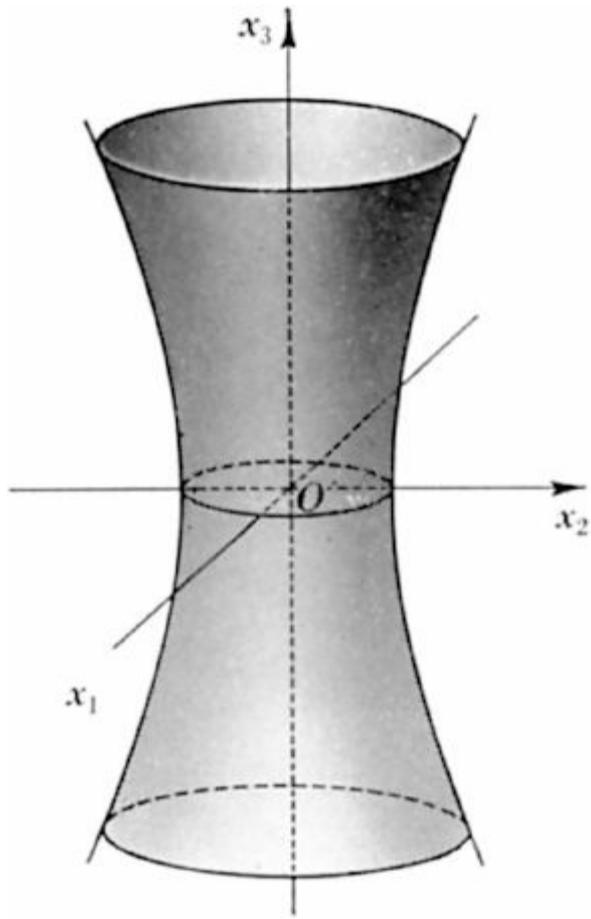
FIG. 6.6
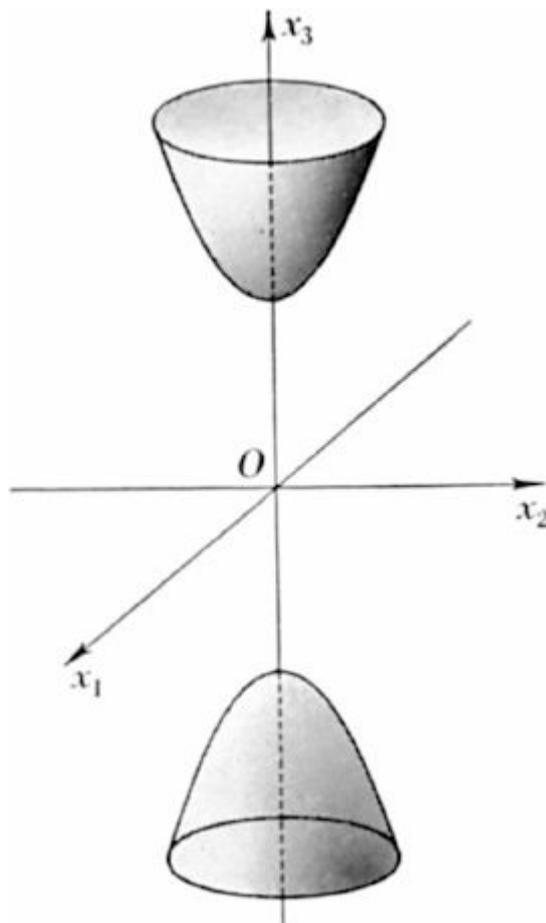
FIG. 6.7

By setting each coordinate equal to zero in turn, i.e., by taking the intersection of the surface with the corresponding coordinate plane, we may determine the shape of the surface. There are two noncentral quadrics: the *elliptic paraboloid* (Figure 6.8) with normal equation

6.35
$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = x_3$$

and the *hyperbolic paraboloid* (Figure 6.9) with equation

6.36
$$\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = x_3$$

Again we have no matrix forms of these equations until we suitably modify our coordinate system.

There is a remarkable property of the hyperboloid of one sheet 6.33
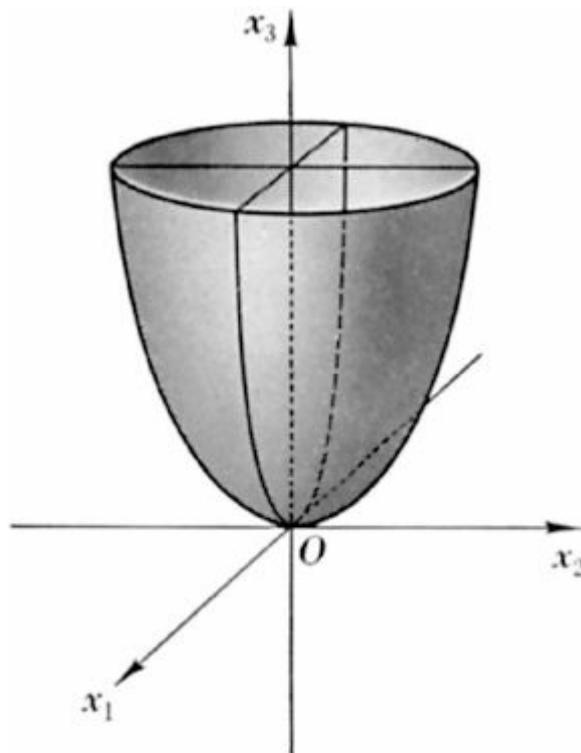
FIG. 6.8
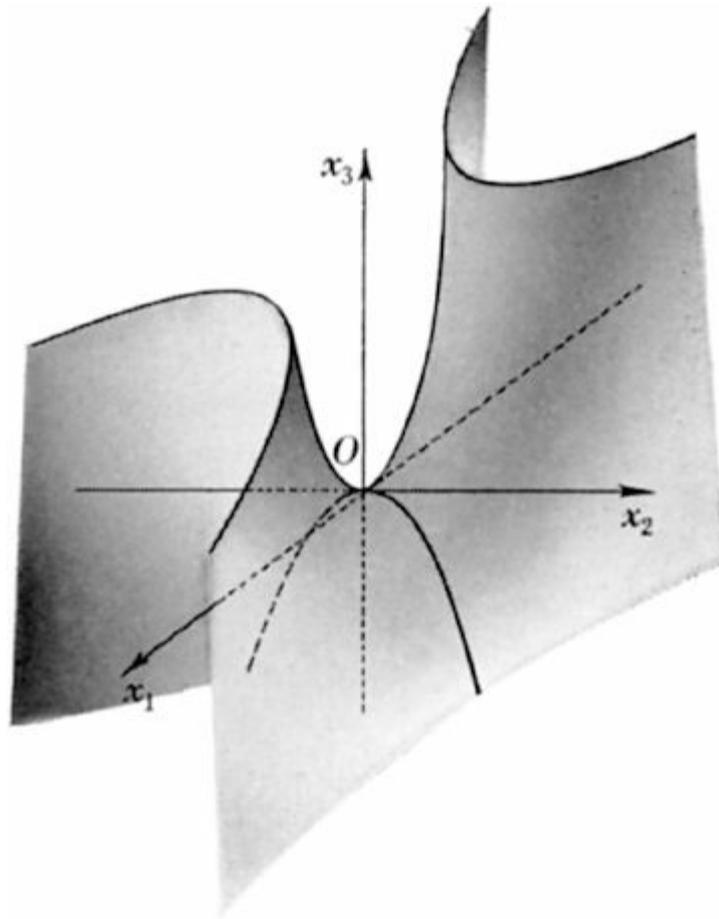
FIG. 6.9

which is also shared by the hyperbolic paraboloid. Let us consider first the surface 6.33, writing the equation in the form

6.331
$$\frac{x_1^2}{a^2} - \frac{x_3^2}{c^2} = 1 - \frac{x_2^2}{b^2}$$

Factoring both sides we see that every point on the surface lies on a line

6.332
$$\frac{x_1}{a} + \frac{x_3}{c} = \lambda \left(1 + \frac{x_2}{b} \right), \qquad \frac{x_1}{a} - \frac{x_3}{c} = \frac{1}{\lambda} \left(1 - \frac{x_2}{b} \right)$$

and also on a line

6.333
$$\frac{x_1}{a} + \frac{x_3}{c} = \frac{1}{\mu} \left(1 - \frac{x_2}{b} \right), \qquad \frac{x_1}{a} - \frac{x_3}{c} = \mu \left(1 + \frac{x_2}{b} \right)$$

Conversely, every point on each of these lines lies in the surface, so that the surface is *ruled*. We call the lines 6.332 *λ-generators* and the lines 6.333 *μ-generators*, and prove two interesting theorems:

.37 Every λ-generator meets every μ-generator.

Proof. If we write the determinant of the coefficients of $x1/a$, $x2/b$, $x3/c$, 1 in the four equations 6.332 and 6.333, we must show that

$$\begin{vmatrix} 1 & -\lambda & 1 & \lambda \\ 1 & \dfrac{1}{\lambda} & -1 & \dfrac{1}{\lambda} \\ 1 & \dfrac{1}{\mu} & 1 & \dfrac{1}{\mu} \\ 1 & -\mu & -1 & \mu \end{vmatrix} = 0$$

and this can easily be verified.

.38 No two generators of the same family intersect.

Proof. If we set $\lambda = \lambda 1$, $\lambda 2$ ($\lambda 1 \neq \lambda 2$) in 6.332, the result follows from the fact that

$$\begin{vmatrix} 1 & -\lambda_1 & 1 & \lambda_1 \\ 1 & \dfrac{1}{\lambda_1} & -1 & \dfrac{1}{\lambda_1} \\ 1 & -\lambda_2 & 1 & \lambda_2 \\ 1 & \dfrac{1}{\lambda_2} & -1 & \dfrac{1}{\lambda_2} \end{vmatrix} \neq 0$$

so that the corresponding equations are inconsistent.

These two properties hold also for lines of the two families

6.361
$$\frac{x_1}{a} + \frac{x_2}{b} = \lambda x_3, \qquad \frac{x_1}{a} - \frac{x_2}{b} = \frac{1}{\lambda}$$

6.362
$$\frac{x_1}{a} + \frac{x_2}{b} = \frac{1}{\mu}, \qquad \frac{x_1}{a} - \frac{x_2}{b} = \mu x_3$$

on the hyperbolic paraboloid 6.36, since the appropriate determinants can readily be constructed and evaluated.

If we take a fixed point P on, say, the hyperboloid of one sheet 6.33, then the two generators through P will define a plane which will be obtained by identifying the two equations

6.391 $k_1 \left\{ \dfrac{x_1}{a} + \dfrac{x_3}{c} - \lambda \left(1 + \dfrac{x_2}{b} \right) \right\} + k_2 \left\{ \dfrac{x_1}{a} - \dfrac{x_3}{c} - \dfrac{1}{\lambda} \left(1 - \dfrac{x_2}{b} \right) \right\} = 0$

6.392 $l_1 \left\{ \dfrac{x_1}{a} - \dfrac{x_3}{c} - \mu \left(1 + \dfrac{x_2}{b} \right) \right\} + l_2 \left\{ \dfrac{x_1}{a} + \dfrac{x_3}{c} - \dfrac{1}{\mu} \left(1 - \dfrac{x_2}{b} \right) \right\} = 0$

This identification yields $k2 = \lambda = l1$, $k1 = \mu = l2$, so that the equation of the plane in question is

6.393 $\qquad \dfrac{x_1}{a} (\lambda + \mu) + \dfrac{x_2}{b} (1 - \lambda\mu) - \dfrac{x_3}{c} (\lambda - \mu) = 1 + \lambda\mu$

where λ, μ are the parameters associated with the two generators of the surface through P. The plane so defined is the *tangent plane* at P, which can also be obtained by using the calculus, as we shall explain in the Appendix.

EXERCISES

1. Find the equations of the two families of generators of the hyperbolic paraboloid $x_1^2 - x_2^2 = x_3$ and in particular the equations of these generators through the point $(2,1,3)$ on the surface.

2. Find the equation of the tangent plane to the surface at the point $(2,1,3)$ in Exercise 1.

3. Find the equation of the tangent plane defined by a λ-generator and μ-generator of the hyperbolic paraboloid 6.36, corresponding to the equation 6.393.

4. Solve the equations 6.332 and 6.333 to obtain the coordinates

$$\left(\dfrac{a(\lambda + \mu)}{1 + \lambda\mu}, \quad \dfrac{b(1 - \lambda\mu)}{1 + \lambda\mu}, \quad \dfrac{c(\lambda - \mu)}{1 + \lambda\mu} \right)$$

of the point of intersection of a λ-generator and a μ-generator of the hyperboloid 6.331. Verify that this point lies on the tangent plane 6.393. Explain why $1 + \lambda\mu \neq 0$.

6.4 CONES, CYLINDERS, AND SURFACES OF REVOLUTION

What locus is represented by the equation 6.11 in 3-space? Clearly $x3$ is unrestricted, so that any point P whose first two coordinates satisfy 6.11 lies on the surface; thus any point on the line through P parallel to $Ox3$ lies on the

surface. This is the simplest example of a *right cylinder*; on such a surface there is one family of generators, each generator being orthogonal to the x_1x_2 plane. It is easy to see that any plane curve defines such a cylinder whose equation is that of the plane curve considered as a locus in 3-space.

A more interesting surface than the cylinder is the *cone*. For example, the locus represented by the quadratic equation

6.41 $$x_1^2 + x_2^2 = x_3^2$$

has the properties that (i) the origin lies on it; (ii) if (x_1, x_2, x_3) satisfies 6.41 then so also does $(\lambda x_1, \lambda x_2, \lambda x_3)$ for every λ; and also, every section of the surface by a plane $x_3 = k$ yields a circle. The surface represented by 6.41 is called a *right circular cone*. Indeed, the argument we have just given proves that

.42 *Every homogeneous equation in x_1, x_2, x_3 represents a cone with vertex the origin.*

We can readily deduce other interesting properties of the right circular cone 6.41. If we take its intersection with the plane $x_1 = k$, we obtain

$$x_3^2 - x_2^2 = k^2$$

which is a *rectangular hyperbola* ($a = b = k$ in 6.22) with transverse axis parallel to $0x_3$. If the secant plane intersects only that part of the cone above the x_1x_2 plane we have an *ellipse*, and if it is parallel to a generator, a *parabola*, as in Figure 6.10. It is these properties which led the Greeks to call these curves *conic sections* or *conics*.
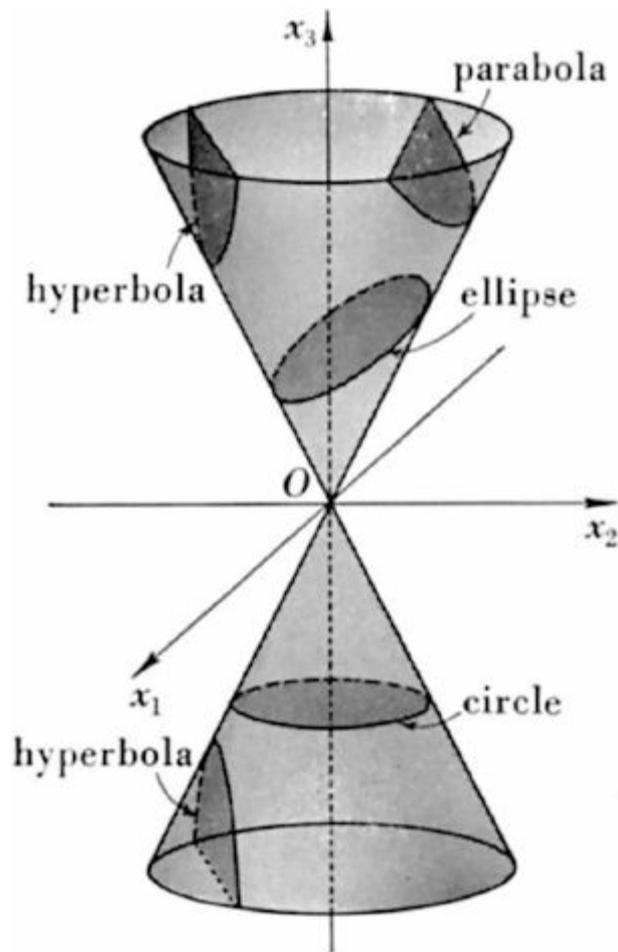
F**IG**. 6.10

A cylinder is a special case of a cone with vertex 'at infinity', and an arbitrary plane will intersect a circular cylinder in an ellipse, a circle, or a pair of parallel lines which may coincide. Of course the plane may not intersect the cylinder at all. In an exactly analogous manner a secant plane through the vertex O of the cone 6.41 may not intersect the cone in any other point, or if it does, the intersection will consist of two intersecting lines through O which again may coincide.

But we may also look at the right circular cylinder represented by the equation

6.43 $$x_1^2 + x_2^2 = r^2$$

and the right circular cone represented by the equation 6.41 as *surfaces of revolution*. In order to study such surfaces in general, consider the plane curve represented by the equation $f(x_2, x_3) = 0$ and imagine it rotated about the axis Ox_3 as in Figure 6.11. If we replace x_2 by $\sqrt{x_1^2 + x_2^2}$ in $f(x_2, x_3) = 0$ and rationalize, we have the desired equation.
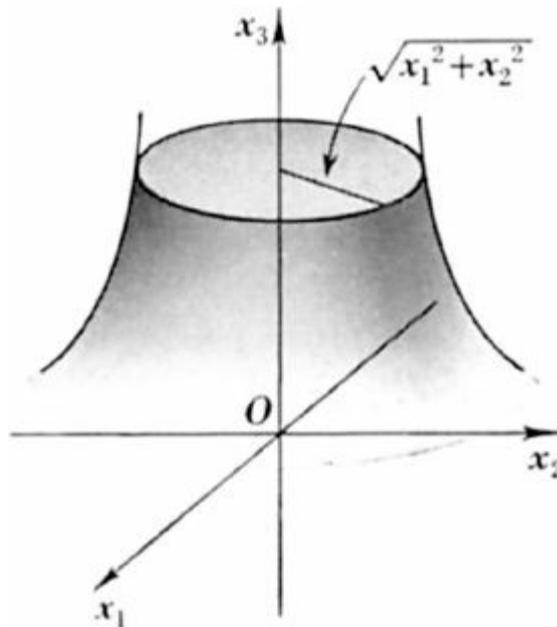
FIG. 6.11

Example. The line $x2 = r$ when rotated yields the equation 6.43 and the line $x2 = x3$ yields the equation 6.41.

The representation of a quadric cone with vertex the origin by a matrix equation is interesting and instructive. The equation of the most general such cone could be written in the form

6.44 $$ax_1^2 + bx_2^2 + cx_3^2 + 2hx_1x_2 + 2gx_1x_3 + 2fx_2x_3 = 0$$

which we can rewrite

6.45 $$X^t \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix} X = 0$$

The form of the matrix in 6.45 is significant since, denoting it by M, we have

6.46 $$M^t = M$$

and such a matrix is said to be *symmetric* (cf. Exercises 3–5 of Section 3.2). Conversely, any symmetric matrix yields a cone with vertex the origin, unless it degenerates (cf. Exercise 5 of Section 6.5).

EXERCISES

1. $\quad \dfrac{x_1^2}{2} + \dfrac{x_2}{3} = 1, \quad \dfrac{x_1}{2} - \dfrac{x_3^2}{3} = 1$

represent two conics in the plane $x3 = 0$. Find the equations of the surfaces of revolution obtained by rotating each conic about each of its principle axes. Roughly sketch the four surfaces so obtained.

2. Find the equations of the cones formed by rotating the line $x3 = 0$, $x2 = 2x1$ about (a) the $x1$ axis, (b) the $x2$ axis.

3. Which of the different types of quadric surface can be realized as surfaces of revolution? Make rough sketches to illustrate your answer.

4. Derive the equation of the *torus* (anchor ring) generated by rotating the circle

$$x_1^2 + x_2^2 + 2ax_1 + b^2 = 0 = x_3 \qquad\qquad (a^2 > b^2)$$

about the $x2$ axis.

5. Discuss the intersection of the plane $x3 = k$ with the torus in Exercise 3, for all values of k.

6.5 PAIRS OF LINES AND PLANES

We may well ask when a quadratic equation in two variables,

6.51
$$ax_1^2 + bx_2^2 + 2hx_1x_2 + 2gx_1 + 2fx_2 + c = 0$$

factorizes, and so represents a pair of lines. One may approach the problem directly and identify the coefficients of the equation 6.51 with those of the equation

6.52
$$(lx_1 + mx_2 + n)(l'x_1 + m'x_2 + n') = 0$$

to obtain

$$ll' = a, \quad mm' = b, \quad nn' = c$$
$$lm' + l'm = 2h, \quad ln' + l'n = 2g, \quad mn' + m'n = 2f$$

Substituting,

$$8fgh = 2ll'mm'nn' + ll'(m'^2n^2 + m^2n'^2)$$
$$+ mm'(n'^2l^2 + n^2l'^2) + nn'(l'^2m^2 + l^2m'^2)$$
$$= 2abc + a(4f^2 - 2bc) + b(4g^2 - 2ca) + c(4h^2 - 2ab)$$

Collecting terms and dividing out the factor 4 we obtain as the required condition:

$$6.53 \qquad \begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix} = 0$$

A more significant approach to the condition 6.53 uses the calculus, as will be explained in the Appendix.

The curve represented by the general equation 6.51 is met by an arbitrary line in two points which may coincide, or in no points at all. This follows immediately by eliminating one of the variables in 6.52 by substituting from the general linear equation

$$6.54 \qquad u_1 x_1 + u_2 x_2 + u_0 = 0$$

and considering the possible roots of the resulting quadratic equation. Alternatively, we may suppose the line 6.54 to be defined parametrically by equations

$$x_1 = y_1 + l_1 l$$

$$x_2 = y_2 + l_2 l$$

as in Section 1.2. Substituting in 6.51 we have a quadratic equation in t and the same argument applies.

It would be possible to study the effect of a linear transformation on the equation 6.51 and show that by suitably choosing the constants we could bring it into one of the normal forms 6.21, 6.22, 6.24 if it did not factorize into a product of linear factors. A similar procedure is applicable to the general quadratic equation in $x1, x2, x3$ and we conclude that every quadric is reducible to one of the normal forms 6.32–6.36 or is recognizable immediately as a cone, a cylinder, or a pair of planes. The method is the same in each case but we shall postpone its consideration to the final chapter of this book.

Just as in the case of a cone, *an arbitrary plane π meets any quadric Q in a conic C or a pair of lines*. This follows from the fact that two planes define a line and by elimination we conclude that any line in π meets Q, and so C, in two points which may coincide or in no points at all. The locus C is quadratic and so must be a conic or a pair of lines.

EXERCISES

1. Obtain the condition 6.53 by solving 6.51 as a quadratic in $x1$. and insisting that the discriminant be a perfect square.

2. For what values of k does the quadratic equation

$$2x_1^2 - x_2^2 - x_1x_2 + (2 - k)x_1 + (1 + k)x_2 = k$$

represent a pair of lines? Find the lines.

3. Test the quadratic equation

$$x_1^2 + x_1x_2 - 2x_2^2 + 3x_2 = 1$$

to see if it factorizes, and if so, obtain the factors. Plot the locus on a sheet of graph paper.

4. Write each of the equations

$$x_1^2 - x_1x_2 + x_2^2 - x_3^2 = 0$$
$$2x_1^2 - x_2^2 - x_3^2 - x_1x_2 + x_1x_3 + 2x_2x_3 = 0$$
$$(x_1 - x_2 + 2x_3)^2 = 0$$

in matrix form 6.45.

5. What is the significance of the condition 6.53 for the equation of a cone 6.45? Test your answer on the equations of Exercise 4, and describe the loci.

6. Prove that the equation

$$X' \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix} X = 1$$

represents a central quadric.

6.6 A QUADRIC TO CONTAIN THREE SKEW LINES

In Section 6.3 we saw that a hyperboloid of one sheet or a hyperbolic paraboloid has on it two families of generators with the properties 6.37 and 6.38. One is tempted to ask the question: Given three skew lines in space is there a quadric containing them?

To be specific, let us take the three lines $l1$, $l2$, $l3$ to be three nonintersecting edges of a rectangular parallelepiped with center the origin and equations as indicated in Figure 6.12 below. The most general quadric surface to contain $l1$ and $l2$ would have the form

6.61 $\quad \alpha(x_1 - a_1)(x_1 + a_1) + \beta(x_1 - a_1)(x_3 - a_3)$
$$+ \gamma(x_1 + a_1)(x_2 + a_2) + \delta(x_2 + a_2)(x_3 - a_3) = 0$$

and if this is to be satisfied by *every* point of $l3$ then we must have

$$0 = \alpha(x_1 - a_1)(x_1 + a_1) + \beta(x_1 - a_1)(-2a_3)$$
$$+ \gamma(x_1 + a_1)(2a_2) + \delta(2a_2)(-2a_3)$$
$$= \alpha x_1^2 + 2(-a_3\beta + a_2\gamma)x_1 + (-a_1^2\alpha + 2a_1a_3\beta + 2a_1a_2\gamma - 4a_2a_3\delta)$$

so that

$$\alpha = 0, \quad a_3\beta = a_2\gamma, \quad a_1\gamma = a_3\delta$$

from which we conclude that

6.62
$$\alpha = 0, \quad \frac{\beta}{a_2} = \frac{\gamma}{a_3} = \frac{\delta}{a_1}$$

and 6.61 takes the simple form.

6.63
$$a_1x_2x_3 + a_2x_1x_3 + a_3x_1x_2 + a_1a_2a_3 = 0$$
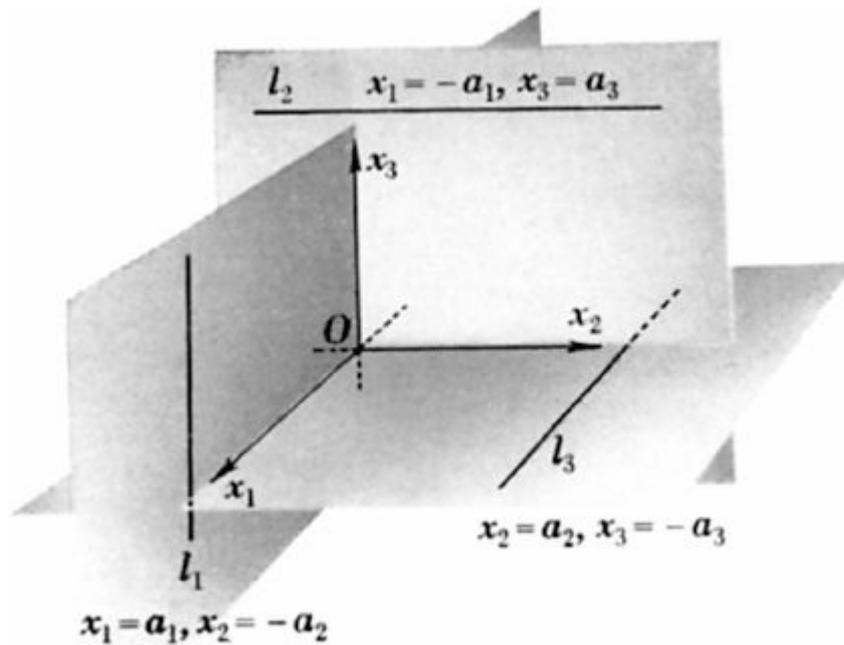
FIG. 6.12

While the choice of the lines $l1$, $l2$, $l3$ would seem specialized, no metrical property is involved and we have only chosen them symmetrically with regard to the basis vectors and this can always be arranged. That the ruled surface in question is uniquely defined is important and since it is centrally symmetrical it must be a hyperboloid of one sheet. If the three lines are parallel to the same plane, we no longer have such central symmetry and the surface is an hyperbolic paraboloid.

Inserting the values of α, β, γ, δ from 6.62, the equation 6.61 takes the form

6.64 $\quad a_1(x_2 + a_2)(x_3 - a_3) + a_2(x_1 - a_1)(x_3 - a_3) + a_3(x_1 + a_1)(x_2 + a_2) = 0$

which factors into

6.65
$$(x_2 + a_2) = \lambda(x_1 - a_1)$$
$$a_1(x_3 - a_3) + a_3(x_1 + a_1) = -\frac{a_2}{\lambda}(x_3 - a_3)$$

and similarly into

6.66
$$(x_2 + a_2) = \mu(x_3 - a_3)$$
$$a_1(x_3 - a_3) + a_3(x_1 + a_1) = -\frac{a_2}{\mu}(x_1 - a_1)$$

if we set $\mu = 0, \infty, -a2/a3$, we obtain the lines $l1, l2, l3$ respectively. On the other hand, we may argue as in Section 6.3, that every line 6.65 meets every line 6.66 since the determinant of the coefficients vanishes identically in λ and μ.

We may look at the two systems of generators on a hyperboloid as yielding a parametric representation of the surface in the sense that each point P is the intersection of a unique λ-generator and a unique μ-generator (cf. Exercise 4 of Section 6.3). Alternatively, we may use trigonometric functions. Just as the ellipse 6.21,

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1$$

is represented parametrically in the form $x1 = a \cos \theta$, $x2 = b \sin \theta$, so the ellipsoid 6.32,

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} + \frac{x_3^2}{c^2} = 1$$

is represented parametrically in the form

6.67 $\qquad x_1 = a \cos \theta \sin \varphi, \quad x_2 = b \sin \theta \sin \varphi, \quad x_3 = c \cos \varphi$

Corresponding representations can be constructed for the other quadrics.

EXERCISES

1. Construct a parametrization for the hyperboloid 6.33 corresponding to 6.67

for the ellipsoid.

2. Verify that the hyperbolic paraboloid 6.36 can be represented parametrically by the equations

$$x_1 = ae^\varphi \cosh\theta, \quad x_2 = be^\varphi \sinh\theta, \quad x_3 = e^{2\varphi}$$

and that $\theta + \varphi$ is constant for points of a given generator of one system and $\theta - \varphi$ is constant for a given generator of the system.

3. The equation of an ellipsoid 6.32 may be written

$$\frac{x_1^2 + x_2^2 + x_3^2}{b^2} - 1 + x_1^2\left(\frac{1}{a^2} - \frac{1}{b^2}\right) + x_3^2\left(\frac{1}{c^2} - \frac{1}{b^2}\right) = 0$$

If $a > b > c$, show that

$$x_1^2\left(\frac{1}{a^2} - \frac{1}{b^2}\right) + x_3^2\left(\frac{1}{c^2} - \frac{1}{b^2}\right) = 0$$

represents two planes which intersect the ellipsoid in circles of radius b.

4. Apply the method of Exercise 3 to find the circular sections of the hyperboloid.

5. If we denote the matrix in 6.45 by M, we may write the equation of Exercise 6 of Section 6.5 in the form

$$X'MX - \lambda X'X + \lambda\left(X'X - \frac{1}{\lambda}\right) = 0$$

For what values of λ does the equation

$$X'MX - \lambda X'X = 0$$

represent a pair of planes? Discuss the intersections of the planes with the quadric.

6.7 THE INTERSECTION OF TWO QUADRICS

Finally, we consider the nature of the intersection of two quadric surfaces $Q1$, $Q2$. The simplest procedure is to consider the intersection of each surface with a given plane π, so that we have two conics $C1$, $C2$ in π. If $C1$ and $C2$ intersect, their common points belong to both $Q1$ and $Q2$. Since $C1$ and $C2$ intersect in at most 4 points, we say that the curve of intersection of $Q1$ and $Q2$ in space is of *order* 4. In particular, this *twisted* quartic curve may degenerate into a common

generator and a twisted cubic curve. We investigate this case briefly.

Before doing so, however, we remark that a *curve C of the second order in space must be a conic*, i.e., it must lie in a plane; since, if not, a plane containing two points of C could be chosen so as to meet it in a third point and C would have order greater than 2.

It is sufficient to consider the simplest case, e.g., the intersection of the cone

6.71
$$x_2^2 = x_1 x_3$$

with the hyperbolic paraboloid

6.72
$$x_1 x_2 = x_3$$

The generators of the cone 6.71 are given by the equations

6.73
$$x_2 = \lambda x_1, \qquad x_2 = \frac{1}{\lambda} x_3$$

and those of the hyperbolic paraboloid 6.72 by the equations

6.74
$$x_1 = \lambda, \quad x_2 = \frac{1}{\lambda} x_3; \qquad x_2 = \mu, \quad x_1 = \frac{1}{\mu} x_3$$

and it is easy to see that the two surfaces have in common the $x1$ axis, i.e., the line

6.75
$$x_2 = 0 = x_3$$

The residual intersection is the twisted cubic curve given parametrically by the equations

6.76
$$x_1 = \theta, \quad x_2 = \theta^2, \quad x_. = \theta^3$$

EXERCISE

Make a sketch of the two surfaces 6.71 and 6.72 and indicate as best you can the nature of the intersection.

7

HOMOGENEOUS COORDINATES AND PROJECTIVE GEOMETRY

7.1 EUCLIDEAN GEOMETRY

So far, we have accepted the two basic features of Euclidean geometry without question: namely, *parallelism* and the *Pythagorean theorem*. These are not unrelated, for a fundamental property of two parallel lines is that they are *equidistant,* and the notion of distance is defined by means of the Pythagorean theorem. Even if we agreed not to use the Pythagorean theorem, however, we could still speak of parallel lines using Euclid's definition that there is a unique line through a given point P, coplanar with a given line l and not meeting it.

But we may look at the matter differently. Consider a Euclidean plane π and a point O not in π. Every point P in π determines a unique line OP and every line l in π a unique plane Ol. Two lines l, l' in π which intersect in P determine two planes Ol, Ol' which intersect in OP; while if l and l' are parallel in π then the planes Ol, Ol' still intersect in a unique line $OP\infty$ parallel to π. All such lines $OP\infty$ lie in the plane $\pi\infty$ through O, parallel to π.

Let us set out this correspondence between the points and lines of π and the lines and planes of the *bundle* with vertex O, in the following fashion:

$$\text{Point } P \text{ of } \pi \leftrightarrow \text{Line } OP$$

$$\text{Line } l \text{ of } \pi \leftrightarrow \text{Plane } Ol$$

$$\left.\begin{array}{l}\text{Two points } P, Q \text{ of } \pi \text{ deter-}\\\text{mine a line } PQ \text{ of } \pi.\end{array}\right\} \leftrightarrow \left\{\begin{array}{l}\text{Two lines } OP, OQ \text{ deter-}\\\text{mine a plane } OPQ.\end{array}\right.$$

$$\left.\begin{array}{l}\text{Two lines } l, l' \text{ of } \pi \text{ either}\\\text{intersect or are parallel.}\end{array}\right\} \leftrightarrow \left\{\begin{array}{l}\text{Two planes } Ol, Ol'\\\text{always intersect.}\end{array}\right.$$

If we associate a line OP through O with a point P in π when OP intersects π in P, why should we not similarly associate a point $P\infty$ *at infinity* in π with the line $OP\infty$? Indeed this association is well defined, since any plane Ol' will contain $OP\infty$ if and only if l' is parallel to l. Moreover, all lines $OP\infty$ lie in the plane $\pi\infty$ parallel to π, so that it is natural to speak of $P\infty$ as lying in the line $l\infty$ "at ∞" in π.

In the following section we shall see how this can all be done analytically.

7.2 HOMOGENEOUS COORDINATES

Let us reconsider the problem of finding the common point of two coplanar lines. Assuming now that these lines are parallel, we may take their equations in the form

7.211
$$a_{11}x_1 + a_{12}x_2 = a_{10}$$
$$a_{11}x_1 + a_{12}x_2 = a_{20}$$

Previously we said that such equations are inconsistent for $a10 \neq a20$ and have no solution. If, however, we introduce a new "variable of homogeneity" $x0$ and write

7.212
$$a_{11}x_1 + a_{12}x_2 = a_{10}x_0$$
$$a_{11}x_1 + a_{12}x_2 = a_{20}x_0$$

then the two equations do have a solution $x1 = ka12$, $x2 = -ka11$, $x0 = 0$ for all values of k. If we call $x1, x2, x0$ the *homogeneous coordinates* of the common point, the two approaches can be reconciled by writing the familiar nonhomogeneous coordinates, which we temporarily denote $\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots$, in the form

7.22
$$\bar{x}_1 = \frac{x_1}{x_0}, \quad \bar{x}_2 = \frac{x_2}{x_0}, \quad \bar{x}_2 = \frac{x_3}{x_0}, \quad \cdots$$

Notice that the homogeneous coordinates $x1, x2, x3, \ldots x0$ are determined up to a constant factor $k \neq 0$ only. For any finite point we may set $x0 = 1$ so that $\bar{x}_i = xi$ ($i = 1,2,3, \ldots$) but for points "*at infinity*," $x0 = 0$, which is the equation of the *space at infinity*. The homogeneous coordinates of the origin can be taken to be $(0,0,0, \ldots , 1)$ while those of the point *at infinity* on $Ox1$ can be taken to be $(1,0,0, \ldots , 0)$, those of the point at infinity on $Ox2$ can be taken to be $(0,1,0, \ldots , 0)$, and so on. There is no ambiguity here since division by zero is not allowed, and multiplication by any $k \neq 0$ does not change the point so represented.

Just as the equation of a line in two dimensions can be taken to be homogeneous as in 7.212, so the equation of a plane in three dimensions can be taken to be homogeneous:

7.23
$$a_1x_1 + a_2x_2 + a_3x_3 = a_0x_0$$

We may consider the arbitrary constant factor k, by which the equation may be multiplied without changing its geometrical significance, to be precisely that k for which the point

$$(x_1, x_2, x_3, x_0) = (kx_1, kx_2, kx_3, kx_0)$$

Thus we have made the designation of a *point* in space conform to the same algebraic convention as holds for the *equation* of any locus.

If we write the general equation of a conic in the form

7.241
$$a\bar{x}_1^2 + b\bar{x}_2^2 + 2h\bar{x}_1\bar{x}_2 + 2g\bar{x}_1 + 2f\bar{x}_2 + c = 0$$

it assumes the homogeneous form

7.242
$$ax_1^2 + bx_2^2 + 2hx_1x_2 + 2gx_1x_0 + 2fx_2x_0 + cx_0^2 = 0$$

when we replace nonhomogeneous by homogeneous coordinates. Similarly, any algebraic equation* can be made homogeneous and, indeed, equations *which are nonhomogeneous in x1, x2, x0 have no geometrical significance.*

On introduction of the coordinate of homogeneity $x0$, the set of numbers ($x1$, $x2$, $x0$) can no longer be considered as a vector in the plane. Nevertheless, there is the considerable advantage that the quadratic equation 7.242 can now be written

7.243
$$(x_1, x_2, x_0)\begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_0 \end{pmatrix} = 0$$

yielding a significant generalization of the equations 6.23. The similarity of 7.243 and 6.45 is suggestive. Indeed, we obtain the nonhomogeneous equation of a conic by setting $x3 = 1$ in 6.45, i.e., by taking the intersection of the cone by this plane. In so doing we have established exactly the correspondence envisaged in Section 7.1. The vector $X(x1,x2,x3)$ in 3-space defines a unique point ($x1,x2,x0$) in the plane $x3 = x0 = 1$, and ($kx1,kx2,kx3$) defines the same point for every $k \neq 0$.

In order to generalize these ideas further, it is convenient to write the equation 7.243 in the form

7.25
$$(x_1, x_2, x_0)\begin{pmatrix} a_{11} & a_{12} & a_{10} \\ a_{21} & a_{22} & a_{20} \\ a_{01} & a_{02} & a_{00} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_0 \end{pmatrix} = 0$$

where $aij = aji$. In this notation the general homogeneous quadratic equation in $x1, x2, x3\ x0$ becomes

$$7.26 \qquad (x_1, x_2, x_{\cdot\cdot}, x_0)\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{10} \\ a_{21} & a_{22} & a_{23} & a_{20} \\ a_{31} & a_{32} & a_{33} & a_{30} \\ a_{01} & a_{02} & a_{03} & a_{00} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_0 \end{pmatrix} = 0$$

where again $a_{ij} = a_{ji}$. Such an equation contains as a special case each of 6.321 through 6.341, and 7.26 represents a general quadric surface. We shall see how to reduce 7.25 and 7.26 to normal form in the last chapter of this book.

EXERCISES

1. Write the equations

$$x_1 - 2x_2 = 3$$
$$2x_1 - 4x_2 = 2$$

in homogeneous form and solve.

2. What is the nonhomogeneous form of the equation of the curve

$$x_0 x_1^2 - x_2(x_2^2 - x_0^2) = 0?$$

Find the homogeneous coordinates of its intersections with the axes and locate the points in question. Roughly sketch the curve.

3. Write the equation

$$x_1^2 - x_2^2 + x_1 x_2 - 2x_1 + 3x_2 = 1$$

in the form 7.25 with *integral* coefficients.

4. Write the equation

$$x_1 x_2 - x_1 x_3 + x_2 x_3 = 1$$

in the form 7.26 with integral coefficients. What difference would it have made if the right side had been 0 instead of 1?

7.3 AXIOMS OF PROJECTIVE GEOMETRY

By introducing homogeneous coordinates we have been able to extend our analytical machinery so as to take account of the behavior of geometric loci *at infinity*. In a very practical sense we have *adjoined* such points, and the line or plane (space) *at infinity* in which they lie, to the ordinary Euclidean plane or

space with which we are familiar. We have, in fact, made it possible to say that *any* two coplanar lines have a point in common (finite or infinite) and *any* two planes in space have a line in common (finite or infinite) and *every* line meets *every* plane in a point (finite or infinite). If we do not involve the notion of *distance* or *length,* i.e., if our space has no *metric* imposed upon it (e.g., by a Pythagorean theorem), we have what is called *affine* geometry. If we do not distinguish between finite and infinite elements, we have *projective* geometry.

In order to clarify these ideas, let us approach the situation from the opposite point of view and give a system of *incidence axioms* which will define a projective space, say of three dimensions. To this end we take a *point* to be undefined and a *line* to be an undefined class of at least two points. It is important to be somewhat vague here so that our system of axioms may be capable of different interpretations. In this way we can include many apparently diverse systems which are subject to the same *relations*, when *point* and *line* are interpreted differently. Concerning these undefined elements we make the following assumptions:

.31 There are at least two distinct points.

.32 Two distinct points *A, B* determine one and only one line *AB* (or *BA*) through both *A* and *B*.
It is not difficult to prove that if *C* and *D* are points on *AB*, then *A* and *B* are points on *CD*. Moreover, two distinct lines cannot have more than one common point.

.33 If *A, B* are distinct points, then there is at least one point *C* distinct from *A, B* on the line *AB*.

.34 If *A, B* are distinct points, then there is at least one point *C* not on the line *AB*.

.35 If *A, B, C* are Three noncollinear points and *D* is a point on *BC* distinct from *B* and *C* and *E* is a point on *CA* distinct from *C* and *A*, then there is a point *F* on *AB* such that *D, E, F* are collinear.
This axiom 7.35, first stated by Pasch in 1880, makes it possible to define a *plane ABC* and to prove that *any two coplanar lines have a point in common.* In order to have a space of at least three dimensions, we assume that:

.36 If *A, B, C* are three noncollinear points, then there is at least one point *D* not on the plane *ABC*.
To exclude the possibility that it have more than three dimensions, we assume finally that

.37 Any two distinct planes have a line in common.
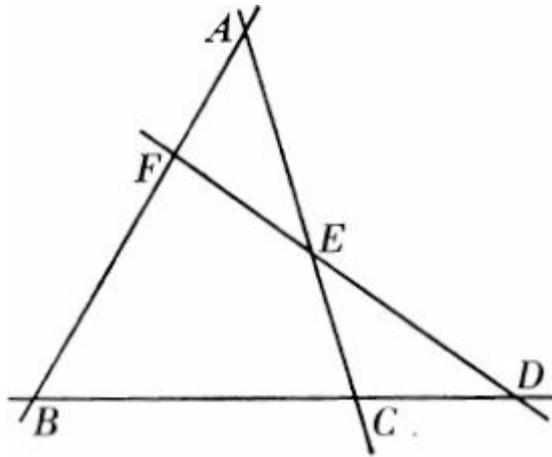
These seven axioms describe how points, lines, and planes are *related* to each other in projective geometry. There must be at least three points on a line; how many more is not specified. Already we have an example of the utility of leaving "point" undefined. In fact, the interpretations of "point" as a "line of the bundle" and "line" as a "plane of the bundle" in Section 7.1 satisfy all our assumptions, and we conclude that the geometry of the bundle is in one-to-one correspondence with the geometry of the projection plane.

7.4 THEOREMS OF DESARGUES AND PAPPUS

Many beautiful theorems can be proved on the basis of these assumptions, but we prove only one due to the French geometer Desargues (1593–1662).

.41 Theorem of Desargues If two triangles ABC, A′B′C′ are situated in the same or in different planes and are such that BC, B′C′ meet in L, CA, C′A′ meet in M, and AB, A′B′ meet in N, where L, M, N are collinear, then AA′, BB′, CC′ are concurrent, and conversely.

Proof, (i) If we assume that the two triangles are in different planes π, π', then the three points L, M, N must lie on the line l common to π and π'. Since A, A′, B, B′ are coplanar, as also are B, B′, C, C′ and C, C′, A, A′, we know that these three planes must have a point O in common so that AA′, BB′, CC′ all pass through O. The converse theorem follows by reversing the argument.

(ii) If the two triangles ABC, A′B′C′ lie in the same plane π, we may choose a plane $\pi1$ through l distinct from π and a point P not in π or $\pi1$. Projecting A′B′C′ from P into a triangle $A1B1C1$ in $\pi1$, we know the theorem is true for ABC and $A1B1C1$, so that there exists a point $O1$ in which $AA1$, $BB1$ $CC1$ all meet. Projecting back again onto π from P, the point $O1$ projects into a point O in which AA′, BB′, CC′ all meet, as required. The converse follows similarly.
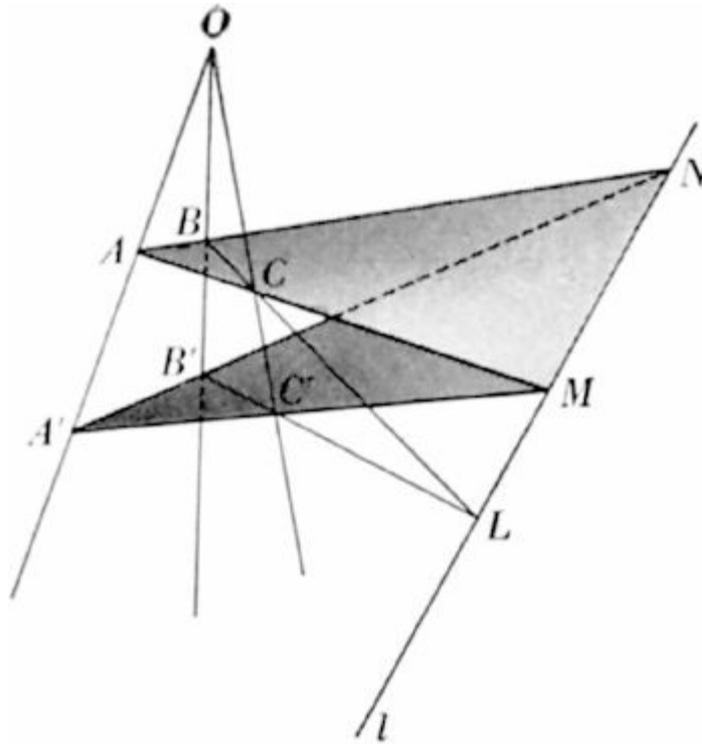
F**IG**. 7.2

The remarkable fact about Desargues' theorem is that it cannot be proved for coplanar triangles unless the plane containing them is embedded in a three-dimensional space. Examples of *non-Desarguesian* planes can be constructed which satisfy axioms 7.31–7.35 without Desargues' theorem being valid (cf. the exercise of Section 7.6). There is thus a notable difference between spaces of two and three or more dimensions which is far from being fully understood.

While the theorem of Pappus (3rd Century A.D.) is expressible in terms of incidence relations, it cannot be proved without further assumptions :

.42 Theorem of Pappus If A, B, C *are any three distinct points on a line l and* A', B', C' *any three points on a line l' intersecting l, then the three points* (BC',B'C), (CA',C'A), (AB',A'B) *are collinear.*

Only relatively recently has the full significance of these theorems of Desargues and Pappus been brought to light in Hilbert's *Foundation of Geometry* (Chicago: Open Court, 1938).
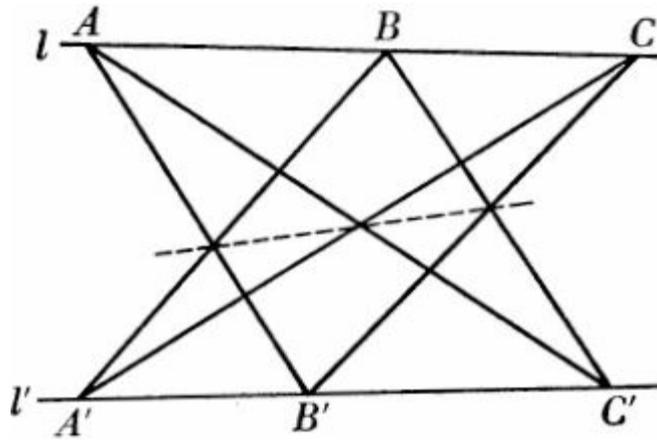
FIG. 7.3

EXERCISES

1. Formulate the theorem of Desargues in the bundle and verify that the proof remains valid.

2. Formulate the theorem of Pappus in the bundle.

7.5 AFFINE AND EUCLIDEAN GEOMETRY

Having defined a projective space, in particular a projective plane π, we can specialize an arbitrary line $l\infty$ in π which we may designate as the "*line at* ∞." Any two lines l, l' which intersect on$l\infty$ will be called *parallel*. If l meets $l\infty$ in $P\infty$, then through any point P not on l there passes one and only one line $PP\infty$ parallel to l, and this is Euclid's axiom of parallelism. Such a definition of parallelism involves no *metric*, i.e., no measure of length or of angle, and yields *affine* geometry.
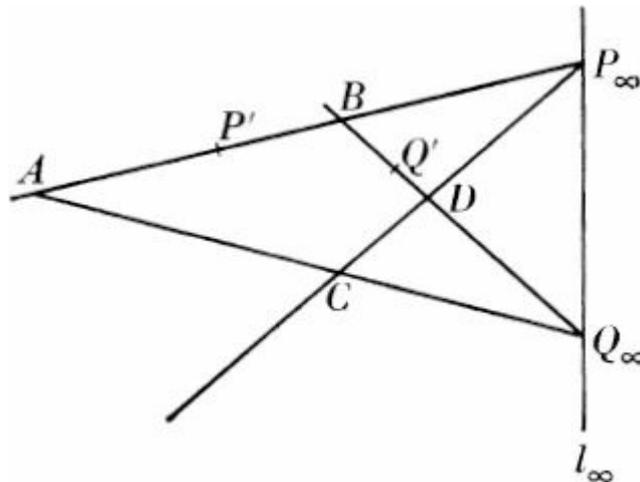
FIG. 7.4

In order to introduce a measure of length it is natural to proceed in two stages: (i) to define equality of segments under *parallel translation* and then (ii) under *rotation*.

(i) Having defined parallelism in affine geometry, we have a parallelogram *ABCD* if *AB*‖*CD* and *AC*‖*BD*, as in Figure 7.4. The statement that *AB* = *CD* and *AC* = *BD* is natural, and leads to all the familiar consequences.

(ii) In order to define equality of segments under *rotation*, we must distinguish a *circle* from the general conic 7.242. The most natural way of doing this is to observe that every circle

$$7.51 \qquad\qquad x_1^2 + x_2^2 + 2gx_1x_0 + 2fx_2x_0 + cx_0^2 = 0$$

intersects the line at infinity $x_0 = 0$ in the two so-called *"circular points at infinity"* given by the equations

$$7.52 \qquad\qquad x_1^2 + x_2^2 = 0 = x_0$$

These points have the conjugate complex coordinates $(1, \pm i, 0)$. Conversely, any conic through these two *circular points* is by definition a circle. By choosing two such points on *l∞* and designating them as "circular points," it is possible to introduce a measure of length and eventually the full Euclidean metric.

Though we cannot go into details here, this building up of Euclidean geometry from projective and affine geometry leads to a clearer understanding of the ideas involved. In the following chapter, we shall study geometry on the surface of a sphere. Besides being our "homeland" this provides the simplest example available of a *non-Euclidean* metric.

EXERCISES

1. Set up equations corresponding to 6.15 to represent the parallel translation of the origin to the point $P(p1,p2)$. What would these become in homogeneous coordinates?

2. Prove that the equation 7.243 of any conic passing through the circular points $(1, \pm i, 0)$ must reduce to the equation 7.51 of a circle.

3. What is the effect of the parallel translation of Exercise 1 on the circular points?

4. Write the general linear transformation in the plane in homogeneous coordinates and consider its effect on the line at infinity. What is the condition that it leave *l∞* (i) invariant, (ii) pointwise invariant? In which of these categories would you place (a) a parallel translation, (b) a rotation

about the origin?

5. Write the general linear transformations of Exercise 4 in nonhomogeneous form. What would be the corresponding transformations in space?

7.6 DESARGUES' THEOREM IN THE EUCLIDEAN PLANE

In order to clarify these ideas still further, let us consider the theorem of Desargues from the point of view of Euclidean geometry. In Figure 7.2 of Section 7.4 we could assume that all the intersections are "finite" and in this case the proof given is applicable in the Euclidean case. Let us consider in particular the two cases illustrated below in which one (or two and so each) side of the triangle $A'B'C'$ is parallel to the corresponding side of the triangle ABC. Both triangles are here assumed to lie in the same plane.
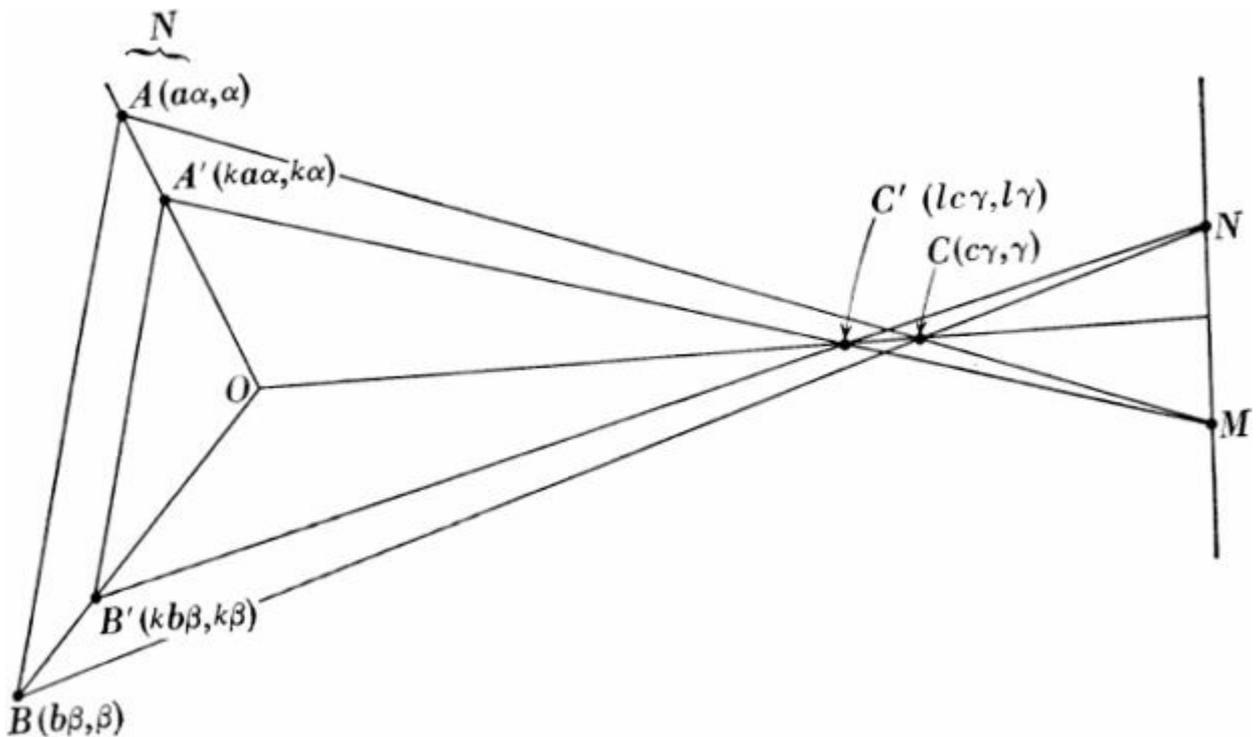
FIG. 7.5

Case (i). We take the center of perspective O as origin and the coordinates and equations as indicated in Figure 7.5. The parallelism of AB and $A'B'$ is expressed by introducing the constant $k \neq 0$ in the manner indicated; we assume $l \neq k$. Writing the equations of AC and $A'C'$ in parametric form, we have

7.61
$$AC : \begin{cases} x_1 = a\alpha + (a\alpha - c\gamma)s \\ x_2 = \alpha + (\alpha - \gamma)s \end{cases}$$

and

7.62
$$A'C' : \begin{cases} x_1 = ka\alpha + (ka\alpha - lc\gamma)t \\ x_2 = ka + (ka - l\gamma)t \end{cases}$$

and if these are identified to obtain the coordinates of M, we have

$$ka\alpha + (ka\alpha - lc\gamma)l = a\alpha + (a\alpha - c\gamma)s$$
$$ka + (ka - l\gamma)l = a + (a - \gamma)s$$

Multiplying the second by a and subtracting, we conclude that $s = t$, so that

$$t = \frac{1 - k}{k - l}, \qquad s = \frac{l(1 - k)}{k - l}$$

By replacing a, α, s, t by b, β, u, v we obtain for the point L the parameter $u = v$ and, as before,

7.63
$$v = \frac{1 - k}{k - l} = t, \qquad u = \frac{l(1 - k)}{k - l} = s$$

The direction numbers of AB are $a\alpha - b\beta$, $\alpha - \beta$, while those of ML are

$$[a\alpha + (a\alpha - c\gamma)s] - [b\beta + (b\beta - c\gamma)u] = (a\alpha - b\beta)(1 + s)$$

and

$$[\alpha + (\alpha - \gamma)s] - [\beta + (\beta - \gamma)u] = (\alpha - \beta)(1 + s)$$

by 7.63; thus ML is parallel to AB as we wished to show.
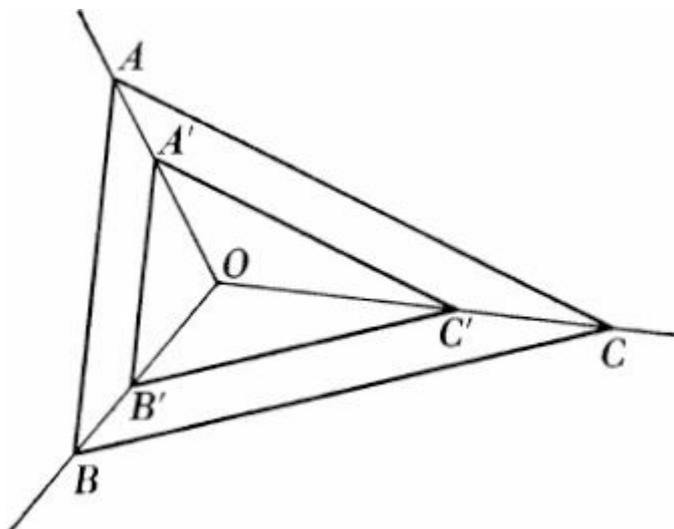
FIG. 7.6

Case (ii). If $A'B'$ is parallel to AB and $B'C'$ is parallel to BC, then it follows immediately that $A'C'$ is parallel to AC, though we cannot set $k = l$ in 7.63 to prove it.

Instead of introducing a different argument in each of the three cases of Desargues' theorem we could have used homogeneous coordinates. We illustrate the method in the following section.

EXERCISES

The simplest example of a non-Desarguesian plane is due to F. R. Moulton. With nonhomogeneous coordinates taken in the Euclidean plane π, a *modified line* is defined by the equation

$$x_2 = m(x_1 - a)f(x_2,m)$$

where the function f is defined as follows:

$$\text{(i) if} \quad m \leqslant 0, \quad f(x_2,m) = 1$$
$$\text{(ii) if} \quad m > 0 \quad \text{and} \quad x_2 \leqslant 0, \quad f(x_2,m) = 1$$
$$\text{(iii) if} \quad m > 0 \quad \text{and} \quad x_2 > 0, \quad f(x_2,m) = \tfrac{1}{2}$$

A modified line is identical with an ordinary line in π in the first case; in the other cases a modified line is made up of two "half-lines." Verify that:

1. Any two points P, Q uniquely determine a modified line PQ;

2. Two modified lines intersect in a unique point or are parallel;

3. Two modified lines are parallel if and only if the corresponding half-lines are parallel;

4. Desargues' theorem is *not* valid for all choices of the two triangles ABC and $A'B'C'$ in Moulton's geometry.

7.7 PAPPUS' THEOREM IN THE EUCLIDEAN PLANE

We have two special cases of Pappus' theorem in the Euclidean plane, but we shall consider first the general theorem 7.42 using homogeneous coordinates.
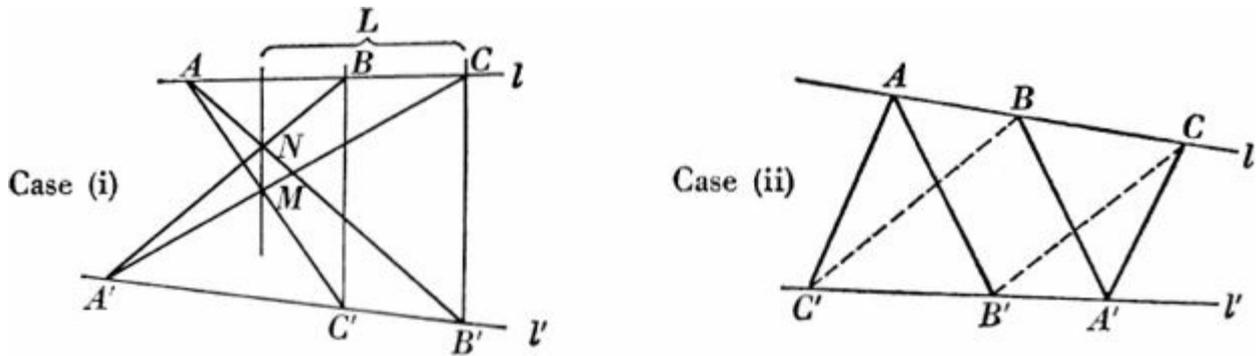
FIG. 7.7

If we take the origin at the intersection of l and l', we may take these lines as coordinate axes with equations $x1 = 0, x2 = 0$ respectively, so that the homogeneous coordinates of A, B, C may be taken to be $(0,a,1)$, $(0,b,1)$, $(0,c,1)$ and those of A',B',C' to be $(a',0,1)$, $(b',0,1)$, $(c',0,1)$ respectively. In such a choice of oblique axes we are utilizing the convenience of expression suggested in Section 5.1.

Setting up the necessary equations, e.g.,

$$\frac{x_1}{c'} + \frac{x_2}{b} = x_0 = \frac{x_1}{b'} + \frac{x_2}{c}$$

for the point L, we readily verify that the *homogeneous* coordinates of L, M, N may be taken to be

7.71 L : $\dfrac{1}{b} - \dfrac{1}{c}, \quad \dfrac{1}{b'} - \dfrac{1}{c'}, \quad \dfrac{1}{bb'} - \dfrac{1}{cc'}$

7.72 M : $\dfrac{1}{c} - \dfrac{1}{a}, \quad \dfrac{1}{c'} - \dfrac{1}{a'}, \quad \dfrac{1}{cc'} - \dfrac{1}{aa'}$

7.73 N : $\dfrac{1}{a} - \dfrac{1}{b}, \quad \dfrac{1}{a'} - \dfrac{1}{b'}, \quad \dfrac{1}{aa'} - \dfrac{1}{bb'}$

Since the third-order determinant Δ made up of these three sets of coordinates vanishes (the sum in each column being zero), we conclude that L, M, N are collinear and *this is independent of the vanishing or nonvanishing of the entries in the third column of Δ* (i.e., of the coordinate of homogeneity). Thus we have proved Pappus' theorem not only when all intersections L, M, N are finite, but also in cases (i) and (ii) of Figure 7.7.

EXERCISE

1. Set up the necessary equations to justify 7.71–7.73.

7.8 CROSS RATIO

There is one remarkable property of the correspondence established in Section 7.1 which again we prove only in the Euclidean plane. If the four lines *OA, OB, OC, OD* in Figure 7.8 have equations

7.81
$$x_2 = ax_1, \quad x_2 = bx_1, \quad x_2 = cx_1, \quad x_2 = dx_1$$

we may suppose that they are met by any transversal

7.82
$$x_2 = mx_1 + n$$

in the points *A, B, C, D*. If *A*1, *B*1, *C*1, *D*1 are the feet of the perpendiculars from *A, B, C, D* on *Ox*1 then, from similar triangles,

7.83
$$\frac{AB \cdot CD}{CB \cdot AD} = \frac{A_1B_1 \cdot C_1D_1}{C_1B_1 \cdot A_1D_1}$$

where

$$A_1B_1 = \frac{n}{a - m} - \frac{n}{b - m} = \frac{n(b - a)}{(a - m)(b - m)}, \quad \text{etc.}$$

Assuming that $n \neq 0$, we conclude that the *cross ratio*

7.84
$$\{AC, BD\} = \frac{AB \cdot CD}{CB \cdot AD} = \frac{(b - a)(d - c)}{(b - c)(d - a)}$$

is independent of m, n and so of the choice of the transversal.
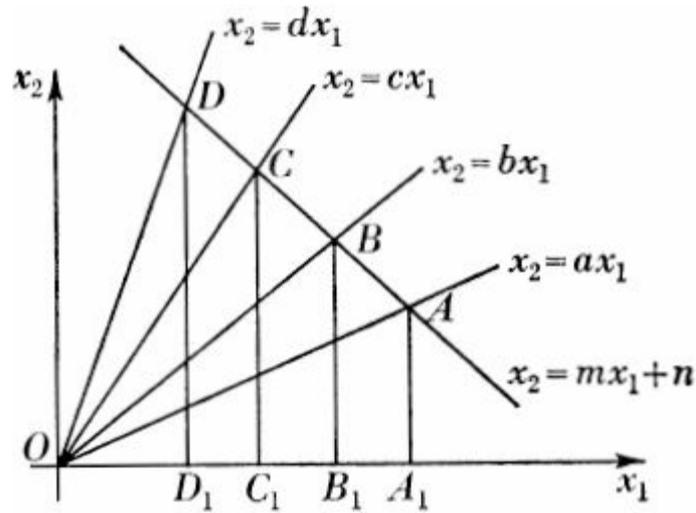
FIG. 7.8
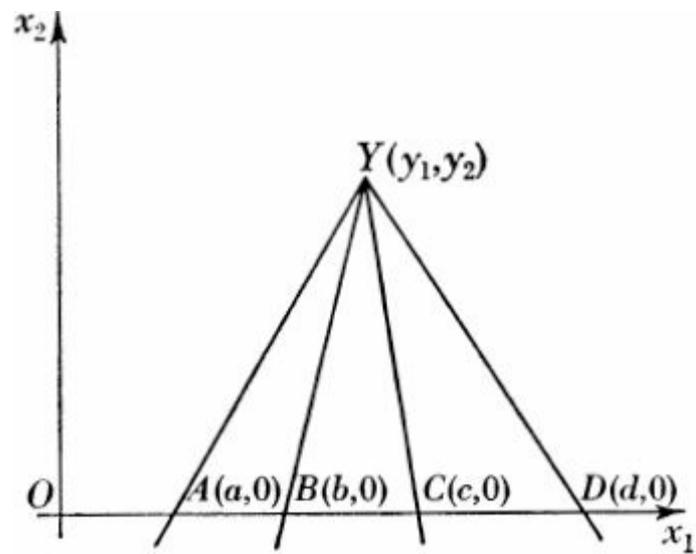
FIG. 7.9

But we can look at the problem slightly differently. Suppose now that the points $A(a,0)$, $B(b,0)$, $C(c,0)$, $D(d,0)$ are fixed and the vertex $Y(y1,y2)$ of the pencil of four lines is arbitrary (Figure 7.9). The slopes of the lines are easily seen to be

$$m_a = \frac{y_2}{y_1 - a}, \quad m_b = \frac{y_2}{y_1 - b}, \quad m_c = \frac{y_2}{y_1 - c}, \quad m_d = \frac{y_2}{y_1 - d}$$

and assuming that $y2 \neq 0$, we have

7.85
$$\frac{(m_b - m_a)(m_d - m_c)}{(m_b - m_c)(m_d - m_a)} = \frac{(b-a)(d-c)}{(b-c)(d-a)}$$

Thus the same *cross ratio* is defined by 7.84 and 7.85 and this is independent not only of the choice of the transversal for a fixed pencil but also of the pencil

for a fixed transversal. The cross ratio is a *projective invariant* and plays an important role in the further development of projective geometry.

In the particular case in which $\{AC, BD\} = -1$, the four points A, B, C, D are said to form an *harmonic range* and YA, YB, YC, YD an *harmonic pencil*. In this case

$$\{AC,BD\} = \{AC,BD\}^{-1} = \{AC,DB\}$$

and since we always have

$$\{AC,BD\} = \{BD,AC\} = \{CA,DB\}$$

the relationship is completely symmetrical with regard to the pairs A, C and B, D, which are called *harmonic conjugates* of each other.

EXERCISES

1. Prove that the internal and external bisectors of any angle form, with the arms of the angle, an harmonic pencil of lines meeting the base of the triangle in an harmonic range of points.

2. What can you say concerning the pencil of lines and range of points in Exercise 1 if the triangle in question is isosceles?

3. Using homogeneous coordinates, verify that 7.84 may be written in the form

$$\{XY,ZT\} = \frac{(z_1 - x_1)(t_1 - y_1)}{(z_1 - y_1)(t_1 - x_1)} = \frac{(z_1 x_0 - x_1 z_0)(t_1 y_0 - y_1 t_0)}{(z_1 y_0 - y_1 z_0)(t_1 x_0 - x_1 t_0)}$$

4. Prove that $\{XY, ZT\}$ is harmonic for any points $X(x1,1)$, $y(y1,1)$, $Z = [\frac{1}{2}(x1 + y1), 1]$, $T = (1,0)$, and show that the point Z is uniquely defined by X, Y.

5. If the midpoint P' of the segment AB in Figure 7.4 is defined by the condition that $\{AB, P'P\infty\} = -1$, and the midpoint Q of BD by the condition that $\{BD, Q'Q\infty\} = -1$, prove that $P',Q\infty$ and $Q'P\infty$ meet the opposite sides of the parallelogram $ABCD$ in their midpoints.
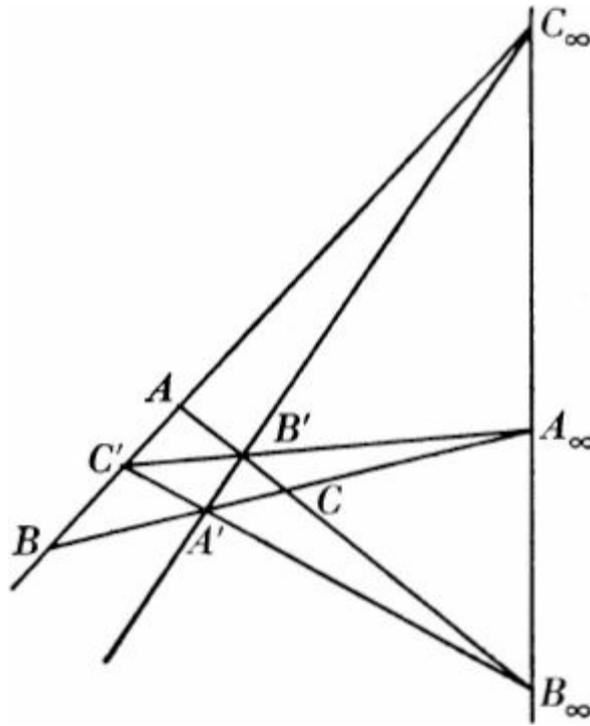
F<small>IG</small>. 7.10

6. If A' is the midpoint of BC, B' the midpoint of CA, and C' the midpoint of AB in the triangle ABC of Figure 10, prove that $C'B' \| BC$ and also that $BA' = C'B' = A'C$ for all $l\infty$ not passing through A, B, C.

   Solution. Assume that $\{BC, A'A\infty\} = -1$, so that projecting from $C\infty$ we have $\{AC, B'B\infty\} = -1$ and from $A\infty$ we have $\{AB, C'C\infty\} = -1 = \{BA, C'C\infty\}$. It follows that $C'A'$ must pass through $B\infty$, so that $BA' = C'B' = A'C$.

7. Prove that the medians AA', BB', CC' of the triangle ABC in Figure 7.10 are concurrent (use Desargues' theorem).

8. By actual substitution, prove that the cross ratio of four points is invariant under any linear transformation of the coordinates.

   * We assume that we have a *rational integral* algebraic function, i.e., a *polynomial*, set equal to zero. Such a polynomial is said to be *homogeneous of weight w* if, when every *xi* is replaced by *xit*, exactly *tw* is a factor of every term.

8

GEOMETRY ON THE SPHERE

8.1 SPHERICAL TRIGONOMETRY

Though we live on the surface of the earth all our days and make maps which guide our automobiles and our airplanes at ever increasing speed, yet most people are not clear as to the relationship between plane and spherical geometry. Undoubtedly the explanation of this lies in the fact that our geometrical ideas stem from Euclid and the concept of a flat world was acceptable long after his time. To understand this mapping process and to introduce the study of a non-Euclidean metric, we examine first *spherical trigonometry* which is the basis of all large-scale surveying.

Euclid's definition of a "straight" line as the shortest distance between two points introduces the idea of a *metric* into geometry. When we think of a surface Σ embedded in 3-dimensional Euclidean space, this notion of the "shortest distance" between two points on Σ leads to a unique curve called a *geodesic* on Σ. If a thread be stretched between two points P and Q on a *sphere* Σ, then this thread will lie along the great circle joining P and Q, and this is a geodesic on the sphere.

For simplicity, let us assume that our sphere Σ has its center at the origin O with unit radius, and so has equation

8.11 $$x_1^2 + x_2^2 + x_3^2 = 1$$

If A, B, C are any three points on Σ, then we call the intersections of the planes OAB, OAC, OBC with Σ a *spherical triangle,* as in the accompanying figure. The metric we adopt on Σ is that of the Euclidean space in which Σ is embedded, so that the "length" of the side AB is determined by the angle $AOB = c$. In fact, these angles a, b, c, which are subtended by the sides BC, CA, AB at O, yield precisely the desired lengths if they are expressed in *radians* (i.e., as fractions of 2π). We define the *angle A* of the spherical triangle ABC to be that between the tangents AD and AE to the great circles AB and AC. We remove the ambiguity as to which of the two possible triangles we are referring to by the convention that *every angle of the triangle ABC shall be less than* π.
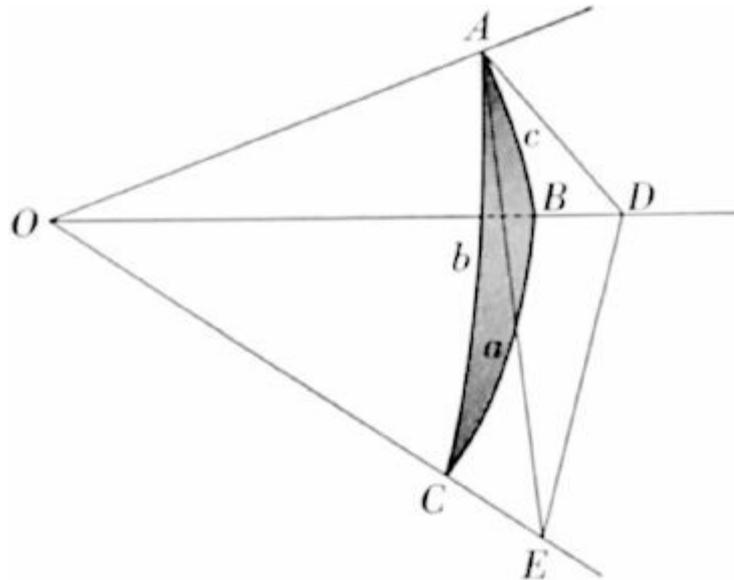
FIG. 8.1

With these definitions it follows from Pythagoras' theorem applied to the triangles *ADE* and *ODE* that

$$AD^2 + AE^2 - 2AD \cdot AE \cos A = OD^2 + OE^2 - 2OD \cdot OE \cos a$$

so that

$$OD \cdot OE \cos a = OA^2 + AD \cdot AE \cos A$$

and

$$\cos a = \frac{OA}{OE} \cdot \frac{OA}{OD} + \frac{AE}{OE} \cdot \frac{AD}{OD} \cos A$$

Referring to Figure 8.1 we have our basic formula:

8.12 $$\cos a = \cos b \cos c + \sin b \sin c \cos A$$

This, and a host of more complicated relations (cf. exercises of Section 8.2) were derived in the sixteenth and seventeenth centuries with the expansion of navigation and the need to fix latitude and longitude on long voyages of exploration. The following consequence of 8.12 resembles a familiar formula of plane trigonometry:

8.13 $$\frac{\sin A}{\sin a} = \frac{\sin B}{\sin b} = \frac{\sin C}{\sin c} = \frac{2\{\sin s \sin (s - a) \sin (s - b) \sin (s - c)\}^{1/2}}{\sin a \sin b \sin c}$$

where $2s = a + b + c$.
Proof. From 8.12 we have*

$$1 - \cos A = 1 - \frac{\cos a - \cos b \cos c}{\sin b \sin c} = \frac{\cos (b - c) - \cos a}{\sin b \sin c}$$

so that

$$\sin^2 \frac{A}{2} = \frac{\sin \frac{1}{2}(a + b - c) \sin \frac{1}{2}(a - b + c)}{\sin b \sin c}$$

Similarly,

$$\cos^2 \frac{A}{2} = \frac{\sin \frac{1}{2}(a + b + c) \sin \frac{1}{2}(b + c - a)}{\sin b \sin c}$$

Thus

$$\sin A = 2 \sin \frac{A}{2} \cos \frac{A}{2}$$

$$= \frac{2 \{\sin s \sin (s - a) \sin (s - b) \sin (s - c)\}^{1/2}}{\sin b \sin c}$$

which yields 8.13, since the expression for $(\sin A/\sin a)$ is symmetrical in each of a, b, c.

It is important to observe that we can have a spherical triangle in which $A = B = C = \pi/2$, so that *the sum of the angles is in this case* $>\pi$. We shall see shortly that this striking difference from plane geometry holds for *every* spherical triangle.

8.2 THE POLAR TRIANGLE

So far, we have not drawn attention to the fact that two points B, C define not one but two (complementary) great circular arcs on Σ. Associated with the great circle BC in Figure 8.2 are the two *poles A', A''* which are the

*The basic formulas of elementary trigonometry which we require are as follows:

$$\sin (\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta; \qquad \sin (\alpha - \beta) = \sin \alpha \cos \beta - \cos \alpha \sin \beta$$

$$\cos (\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta; \qquad \cos (\alpha - \beta) = \cos \alpha \cos \beta + \sin \alpha \sin \beta$$

from which we deduce that

$$\sin 2\alpha = 2 \sin \alpha \cos \alpha$$

$$\cos 2\alpha = \cos^2 \alpha - \sin^2 \alpha = 2 \cos^2 \alpha - 1 = 1 - 2 \sin^2 \alpha$$

and also that

$$\cos \alpha - \cos \beta = -2 \sin \frac{\alpha + \beta}{2} \sin \frac{\alpha - \beta}{2}$$

intersections with Σ of the diameter through O such that $A'OA''$ is normal to the plane OBC. Corresponding to the other two sides of the spherical triangle we have the poles B', B'' and C', C''

FɪG. 8.2
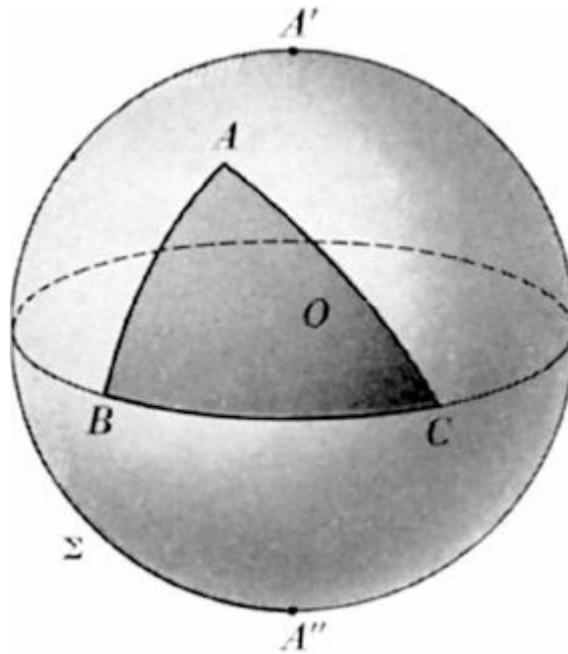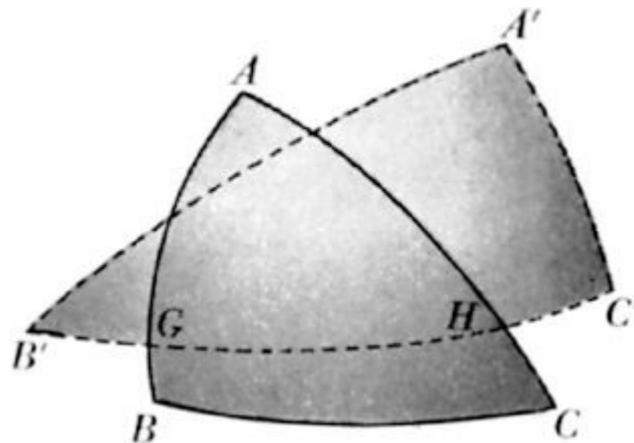
FɪG. 8.3

It is clear that either A' or A'' will lie in the hemisphere which contains A, defined by the great circle BC; let us suppose it to be A'. Similarly, we

distinguish B' and C' and we call $A'B'C'$ the *polar triangle* of ABC. If we examine Figure 8.3 it appears that AB', AC'; BA', BC'; $'CA'$, CB' are all quadrants of great circles, so that *ABC is the polar triangle of A'B'C'*. Moreover, GH is precisely the measure of the angle between the planes OAB and OAC which we have denoted by A, while $B'H = C'G = \pi/2$. Thus

$$B'C' = a' = \pi - A, \quad C'A' = b' = \pi - B, \quad A'B' = c' = \pi - C$$

From the similar properties of ABC we conclude that

$$A' = \pi - a, \quad B' = \pi - b, \quad C' = \pi - c$$

If now we substitute these expressions in the analogue of 8.12 for the polar triangle we have

$$\cos (\pi - A) = \cos (\pi - B) \cos (\pi - C) \\ + \sin (\pi - B) \sin (\pi - C) \cos (\pi - a)$$

or

8.21 $$\cos A = -\cos B \cos C + \sin B \sin C \cos a$$

which we can think of as the *dual* of 8.12.

There is another approach to spherical trigonometry which depends on properties of the vector product. If we write $\overrightarrow{OA} = A$, $\overrightarrow{OB} = B$, $\overrightarrow{OC} = C$, then the Lagrange identity (cf. Exercise 5, Section 5.3) leads to the conclusion

8.22 $$(A \times B) \cdot (A \times C) = \begin{vmatrix} A \cdot A & A \cdot C \\ B \cdot A & B \cdot C \end{vmatrix} = \begin{vmatrix} 1 & \cos b \\ \cos c & \cos a \end{vmatrix}$$
$$= \cos a - \cos b \cos c$$

since $|A| = |B| = |C| = 1$. But we also have

8.23 $$(A \times B) \cdot (A \times C) = |A \times B||A \times C| \cos A$$
$$= \sin c \sin b \cos A$$

by 1.45 and 5.372. Taken together, 8.22 and 8.23 yield 8.12. In order to derive 8.13, we use 5.372 again to yield

8.24 $$|(A \times B) \times (A \times C)| = |A \times B||A \times C| \sin A$$
$$= \sin c \sin b \sin A$$

By evaluating $|(A \times B) \times (A \times C)|$ differently, we obtain a quantity σ which is unchanged by permutation of the vectors A, B, C in any manner. Thus we

conclude that

$$8.25 \qquad \frac{\sin A}{\sin a} = \frac{\sin B}{\sin b} = \frac{\sin C}{\sin c} = \frac{\sigma}{\sin a \sin b \sin c}$$

where σ can be shown to have the value indicated in 8.13.

EXERCISES

1. Derive the special cases of 8.12, 8.13, and 8.21 for a spherical triangle in which $A = \pi/2$.

2. Prove that

$$\sin^2 \frac{a}{2} = \frac{-\cos \frac{1}{2}(A + B + C) \cos \frac{1}{2}(B + C - A)}{\sin B \sin C} = \frac{-\cos S \cos (S - A)}{\sin B \sin C}$$

where $2S = A + B + C$.

3. Prove that

$$\cos \frac{a}{2} = \left\{ \frac{\cos (S - B) \cos (S - C)}{\sin B \sin C} \right\}^{1/2}$$

and thence that

$$\sin a = \frac{2\{-\cos S \cos (S - A) \cos (S - B) \cos (S - C)\}^{1/2}}{\sin B \sin C}$$

4. Show that

$$\tan \frac{A}{2} \tan \frac{B}{2} = \frac{\sin (s - c)}{\sin s}$$

5. Verify the correctness of Napier's *Analogies*

$$\tan \tfrac{1}{2}(A + B) \tan \tfrac{1}{2}C = \frac{\cos \frac{1}{2}(a - b)}{\cos \frac{1}{2}(a + b)}$$

$$\tan \tfrac{1}{2}(A - B) \tan \tfrac{1}{2}C = \frac{\sin \frac{1}{2}(a - b)}{\sin \frac{1}{2}(a + b)}$$

by expanding and substituting from Exercise 4. By applying these formulas to the polar triangle, obtain their duals.

6. Describe a procedure for solving a spherical triangle in which the given

elements are (i) a, b, c; (ii) A, B, C; (iii) a, b, C; (iv) A, B, a. Can you arrange that each problem be solved conveniently by using logarithms?

8.3 AREA OF A SPHERICAL TRIANGLE

Consider now Figure 8.4, in which is represented the hemisphere containing A and defined by the great circle BC. If we suppose the sphere Σ to have radius r and area $4\pi r2$, then the area of the *lune* made up of the parts ABC and $AB1C1$ is $(A/\pi)2\pi r2$. Similarly, the area of $ABC + AB1C$ is $(B/\pi)2\pi r2$, and the area of $ABC + ABC1$ is $(C/\pi)2\pi r2$.

Thus,

8.31

$$\text{Area of triangle } ABC = \frac{1}{2}\left(\frac{A}{\pi} + \frac{B}{\pi} + \frac{C}{\pi} - 1\right)2\pi r^2$$

$$= (A + B + C - \pi)r^2$$

which is a famous result that leads to the designation of the quantity $A + B + C - \pi$ as the *spherical excess*.

Here we have a sharp distinction between geometry on the surface of a sphere and in the Euclidean plane. In the first place the area of a hemisphere is finite ($2\pi r2$), while that of the Euclidean plane is not. If we think of r as tending to infinity and the area of the triangle ABC remaining constant, then it is clear that $A + B + C - \pi$ must tend to zero, which yields the Euclidean theorem concerning the sum of the angles of a triangle.
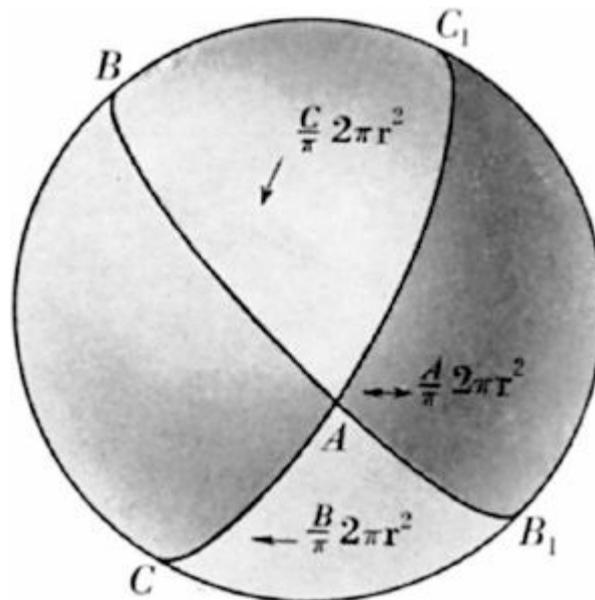
FIG. 8.4

In the second place, any two great circles will intersect (once in any given hemisphere) so that there is no such thing as *parallelism* on the sphere.

Finally, given any two great circles on the sphere there exists a uniquely determined great circle through their poles which is orthogonal to each great circle. The analogous statement concerning lines in the plane is true only if the lines are parallel, and then there is an infinite number of common perpendiculars (see the end of Section 8.8).

EXERCISES

1. Prove that no two triangles on a sphere Σ have equal angles, i.e., are "similar," unless their sides are equal also.

2. From the properly that the medians of a triangle are concurrent in Euclidean geometry deduce the corresponding property of a triangle on the surface of a sphere.

3. If $A = B = C = \pi/2$, express the area of the spherical triangle ABC as a fraction of the total area of the sphere Σ.

4. Show that the two restrictions on a spherical triangle ABC, (i) that every angle be less than π, (ii) that every side be less than πr are equivalent.

8.4 THE INVERSION TRANSFORMATION

Though we have seen how geometry on the sphere can approximate that in the Euclidean plane for r sufficiently large, yet we would like to establish a closer correspondence between the two. To this end we investigate the quadratic transformation known as *inversion*.

In order to define the transformation we consider it first with reference to the circle $c0$,

$$x_1^2 + x_2^2 = a^2$$
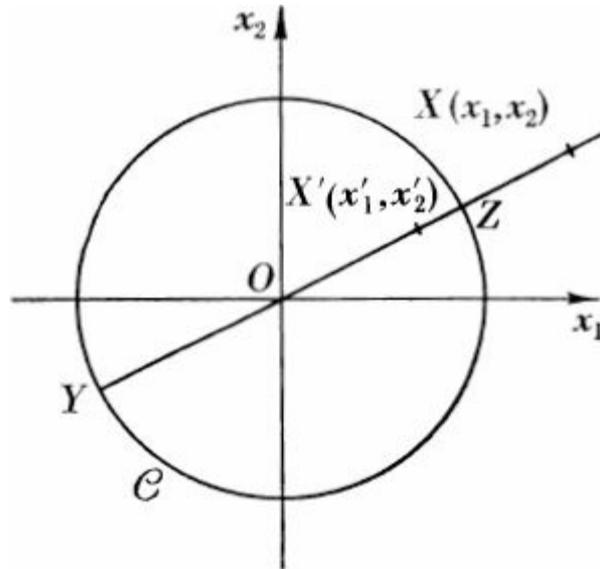
FIG. 8.5

If $X(x_1, x_2)$ is any point different from O, then OX will meet $\mathcal{C}0$ in Y, Z, and the *inverse point* X' of X with regard to $\mathcal{C}0$ is defined to be the harmonic conjugate of X with regard to Y and Z. Thus

8.41
$$\{YZ, XX'\} = \frac{YX \cdot ZX'}{YX' \cdot ZX} = \frac{(YO + OX)(ZO + OX')}{(YO + OX')(ZO + OX)} = -1$$

which on expanding and simplifying leads to the alternative definition

8.42
$$OX \cdot OX' = \sqrt{x_1^2 + x_2^2}\ \sqrt{x_1'^2 + x_2'^2} = a^2$$

Rewriting the equation 8.41 in the form

$$\frac{XY \cdot X'Z}{XZ \cdot X'Y} = \frac{XY(XZ - XX')}{XZ(XY - XX')} = -1$$

we have

8.43
$$\frac{2}{XX'} = \frac{1}{XY} + \frac{1}{XZ}, \qquad \frac{2}{X'X} = \frac{1}{X'Y} + \frac{1}{X'Z}$$

If we substitute from the relation

$$\frac{x_1}{x_2} = \frac{x_1'}{x_2'}$$

in 8.42, we obtain the equations

8.441

$$x_1' = \frac{x_1 a^2}{x_1^2 + x_2^2}, \qquad x_2' = \frac{x_2 a^2}{x_1^2 + x_2^2}$$

of the *quadratic* transformation known as *inversion,* with the inverse transformation given by the equations

8.442

$$x_1 = \frac{x_1' a^2}{x_1'^2 + x_2'^2}, \qquad x_2 = \frac{x_2' a^2}{x_1'^2 + x_2'^2}$$

Leaving the geometrical properties of the transformation to be developed in the next section, let us consider the following

Problem. Find the locus of the harmonic conjugate X of a fixed point X' with regard to the circle $\mathcal{C}$, $x_1^2 + x_2^2 = a2$.

Of course the significance of the problem is that we take an *arbitrary* line through X' meeting $\mathcal{C}$ in Y and Z and look for the locus of the harmonic conjugate X of X'. Such a line could be written in parametric form:

8.45

$$x_1 = x_1' + l_1 t, \qquad x_2 = x_2' + l_2 t$$

where t is the length of the segment $X'X$ and $l1$, $l2$ are the direction cosines of the line. Substituting the expressions 8.45 in the equation of $\mathcal{C}$ we have

8.46

$$(x_1'^2 + x_2'^2 - a^2)\frac{1}{t^2} + \frac{2}{t}(x_1'l_1 + x_2'l_2) + 1 = 0$$

If now we require that the roots $1/t$ shall satisfy 8.43, we have

$$\frac{1}{t} = -\frac{x_1'l_1 + x_2'l_2}{x_1'^2 + x_2'^2 - a^2}$$

so that after cross-multiplying and substituting from 8.45 we have the required locus,

8.47

$$x_1 x_1' + x_2 x_2' = a^2$$

This line is perpendicular to OX' and is called the *polar (line)* of X' with regard to the circle $\mathcal{C}$; X' is the *pole* of 8.47.

EXERCISES

1. Prove that if the polar of X' with regard to a circle $\mathcal{C}$ passes through Y' then the polar of Y' passes through X'

2. By a method similar to that used above, find the polar of X' with regard to each of the three conics

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1, \quad \frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = 1, \quad x_2^2 = 4px_1$$

3. Prove that the polar line of the focus of a conic is the corresponding directrix.

4. If the polar of L with regard to a given conic $\mathfrak{c}$ is l and the polar of M with regard to $\mathfrak{c}$ is m, prove that the pole of LM is the point of intersection of l and m.

5. Using homogeneous coordinates, find the polar lines of the points at infinity $(1,0,0)$, $(0,1,0)$ with regard to the ellipse

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = x_0^2$$

In what point do these polar lines intersect?

6. Use the result of Exercise 5 to define the center of a conic in affine geometry.

8.5 GEOMETRICAL PROPERTIES OF INVERSION

Since the equation of an arbitrary circle $\mathfrak{c}$ can be written in the form

8.51 $$x_1^2 + x_2^2 + 2gx_1 + 2fx_2 + c = 0$$

we obtain the *inverse curve* $\mathfrak{c}'$ with regard to the circle $\mathfrak{c}0$ by substituting from 8.442 and multiplying by $x_1'^2 + x_2'^2 \neq 0$:

8.52 $$a^4 + 2ga^2x_1' + 2fa^2x_2' + c(x_1'^2 + x_2'^2) = 0$$

Thus,

.53 The inverse of a circle $\mathfrak{c}$ is a circle $\mathfrak{c}'$, unless $\mathfrak{c}$ passes through the origin, in which case the inverse of $\mathfrak{c}$ is a line. Conversely, the inverse of a line l is a circle $\mathfrak{c}$ through O, unless l passes through O when l is mapped upon itself by the transformation.

More generally, we can invert with regard to any circle of radius a, the center of which is called the *center of inversion* and a the *radius of inversion*. Clearly, any point on the circle of inversion remains fixed under the transformation.

The three-dimensional analogue of inversion in a circle is inversion in a *sphere* $\Sigma 0$,

$$x_1^2 + x_2^2 + x_3^2 = a^2$$

and the analogues of 8.441 and 8.442 are easily seen to be

8.541
$$x_1' = \frac{x_1 a^2}{x_1^2 + x_2^2 + x_3^2}, \qquad x_2' = \frac{x_2 a^2}{x_1^2 + x_2^2 + x_3^2}$$

$$x_3' = \frac{x_3 a^2}{x_1^2 + x_2^2 + x_3^2}$$

and

8.542
$$x_1 = \frac{x_1' a^2}{x_1'^2 + x_2'^2 + x_3'^2}, \qquad x_2 = \frac{x_2' a^2}{x_1'^2 + x_2'^2 + x_3'^2}$$

$$x_3 = \frac{x_3' a^2}{x_1'^2 + x_2'^2 + x_3'^2}$$

Note that points on the circle or sphere of inversion remain *fixed*, and these are the only fixed points of the transformation.

As in the case of a circle, the sphere with equation

8.55 $$x_1^2 + x_2^2 + x_3^2 + 2gx_1 + 2fx_2 + 2hx_3 + c = 0$$

inverts into the sphere with equation

8.56 $$a^4 + 2ga^2 x_1' + 2fa^2 x_2' + 2ha^2 x_3' + c(x_1'^2 + x_2'^2 + x_3'^2) = 0$$

under the transformations 8.542. Thus,

.57 The inverse of a sphere Σ is a sphere Σ', unless Σ passes through the center of inversion O, in which case the inverse is a plane. Conversely, the inverse of a plane π is a sphere Σ passing through O unless π itself passes through O, in which case π is mapped upon itself by the transformation.

The following special case of 8.53 and 8.57 is of some interest.

.58 Any circle (sphere) orthogonal to the circle (sphere) of inversion is its own inverse.
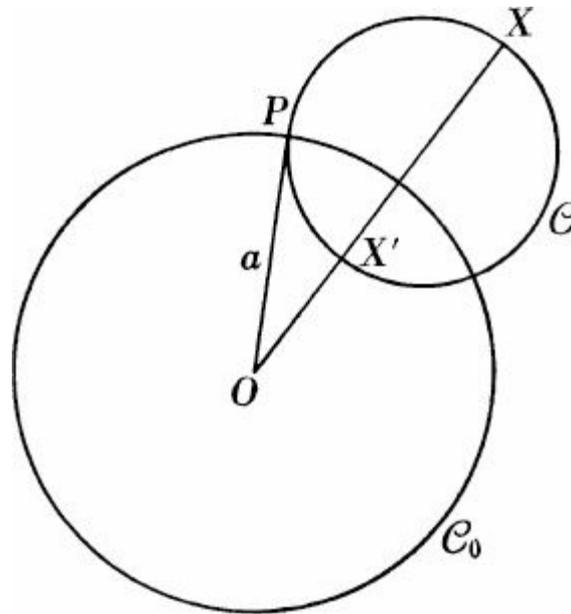
FIG. 8.6
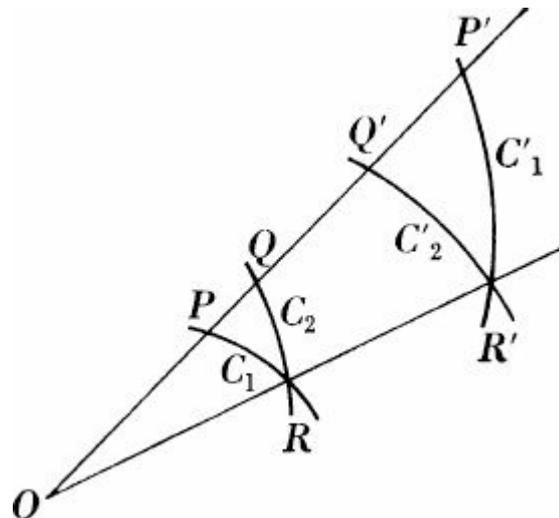
FIG. 8.7

Proof. The proof is an immediate consequence of the fact that if $\mathcal{C}$ is orthogonal to $\mathcal{C}0$, then OP is tangent to $\mathcal{C}$ and

$$OP^2 = OX \cdot OX' = a^2$$

so that the points X, X' on $\mathcal{C}$, collinear with O, are inverse points with regard to $\mathcal{C}0$. The argument for spheres is the same.

There is one further property which is important:

.59 Inversion is a conformal transformation.
This means that if two curves $C1$ and **C**2 intersect in R at an angle α, then their inverses C_1' and C_2' will also intersect in R', the inverse of R, and at the same angle α.

Proof. By definition,

$$OP \cdot OP' = OQ \cdot OQ' = OR \cdot OR'$$

so that

$$\frac{OP}{OR} = \frac{OR'}{OP'} \quad \text{and} \quad \frac{OQ}{OR} = \frac{OR'}{OQ'}$$

From similar triangles, it follows that

$$\angle OPR = \angle OR'P', \qquad \angle OQR = \angle OR'Q'$$

so that

$$\angle PRQ = \angle OPR - \angle OQR = \angle OR'P' - \angle OR'Q' = \angle P'R'Q'$$

Since the angle between the curves $C1$, $C2$ is defined to be the limit α of $\angle PRQ$, as $P \rightarrow R$, it follows that the limit of $\angle P'R'Q'$ is also α, proving the theorem.

EXERCISES

1. Invert the property of circles which states that the angle in a semicircle is a right angle.

2. Show that:

(a) Inverse points with regard to a circle $\mathfrak{e}$ invert into inverse points with regard to $\mathfrak{e}$s.

(b) The limit points of a coaxal system are inverse points with regard to every circle of the system.

(c) A system of nonintersecting coaxal circles may be inverted into a system of concentric circles.

(d) A system of intersecting coaxal circles can be inverted into a system of concurrent lines.

8.6 STEREOGRAPHIC PROJECTION

If we designate by N and S the "north" and "south" poles of a sphere Σ, the process known as *stereographic projection* consists of

(i) projecting Σ onto the tangent plane at N from the point S, or

(ii) projecting Σ onto the equatorial plane from the point N.

We consider first case (i), and show that the effect of projection is identical with that of inverting Σ with respect to a sphere $\Sigma 0$ with center S, passing through N. Clearly the plane π tangent to Σ at N is also the plane tangent to $\Sigma 0$ at N, so that the inverse of Σ is π. Moreover, any plane through NS meets Σ in a line of "longitude" so that such lines invert into lines through N in π. Finally, any small circle on Σ is the intersection of a plane ω with Σ, so that such small circles invert into the inter section of a sphere with π, i.e., into a circle in π. Thus lines of latitude on Σ project into concentric circles in π. Since lines of latitude and longitude intersect orthogonally on Σ, the inverse curves will also have this property, and we can describe the position of a point P' on the map
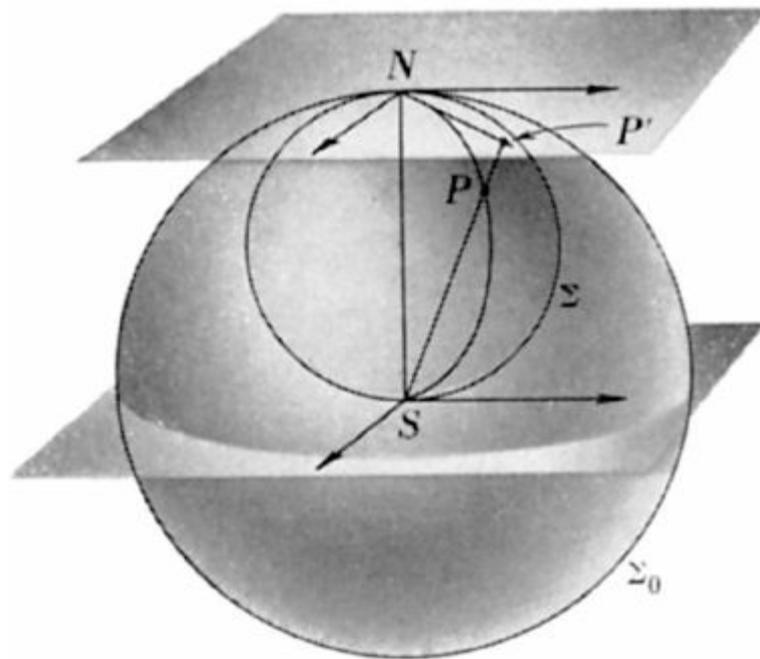
FIG. 8.8
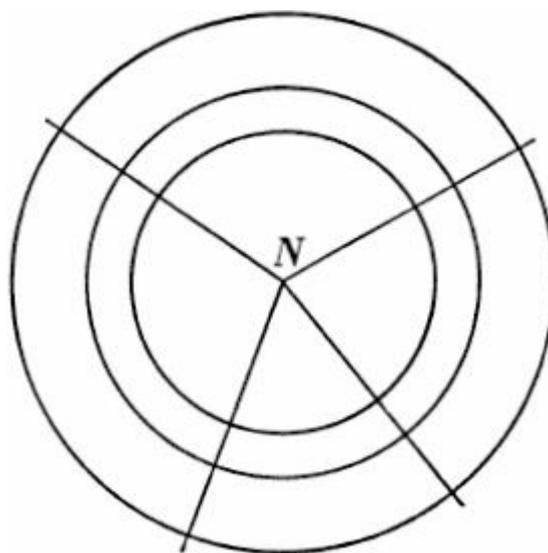
FIG. 8.9
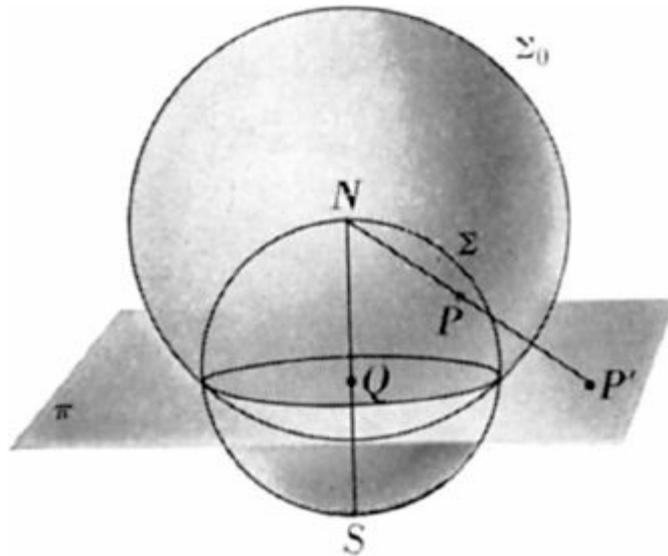
in the same way as P was described on the surface of Σ. There is an increasing amount of distortion, however, as we approach the circumference of the map, since the Euclidean metric in π is different from that on the surface of Σ.

In case (ii) a similar relation between projection and inversion holds, except that the sphere of inversion $\Sigma 0$ has its center at N and passes through the equator of Σ. Great circles on Σ through S project into lines through Q, the center of Σ in π, while small circles on Σ project into circles in π. Conformality is preserved in both cases by the appropriate generalization of 8.59, and distortion again increases as we move away from the point Q.

8.7 ELLIPTIC GEOMETRY

That only certain great circles on the sphere project into lines in π is an undesirable feature of stereographic projection which must be tolerated to obtain conformality. But from an abstract geometrical point of view conformality is unnecessary, and by projection from the center of the sphere onto a tangent plane every great circle will project into a line.

Let us proceed analytically, taking the equation of Σ to be

$$x_1^2 + x_2^2 + x_3^2 = r^2$$

The equation of the plane π tangent to Σ at $\overline{O}$, in Figure 8.11, may be taken to be

$$x_3 = r$$

and we choose $\overline{Ox}\,1$, $\overline{Ox}\,2$ parallel to $Ox1$, $0x2$. We are in exactly the position we considered in such detail in the preceding chapter: the coordinates
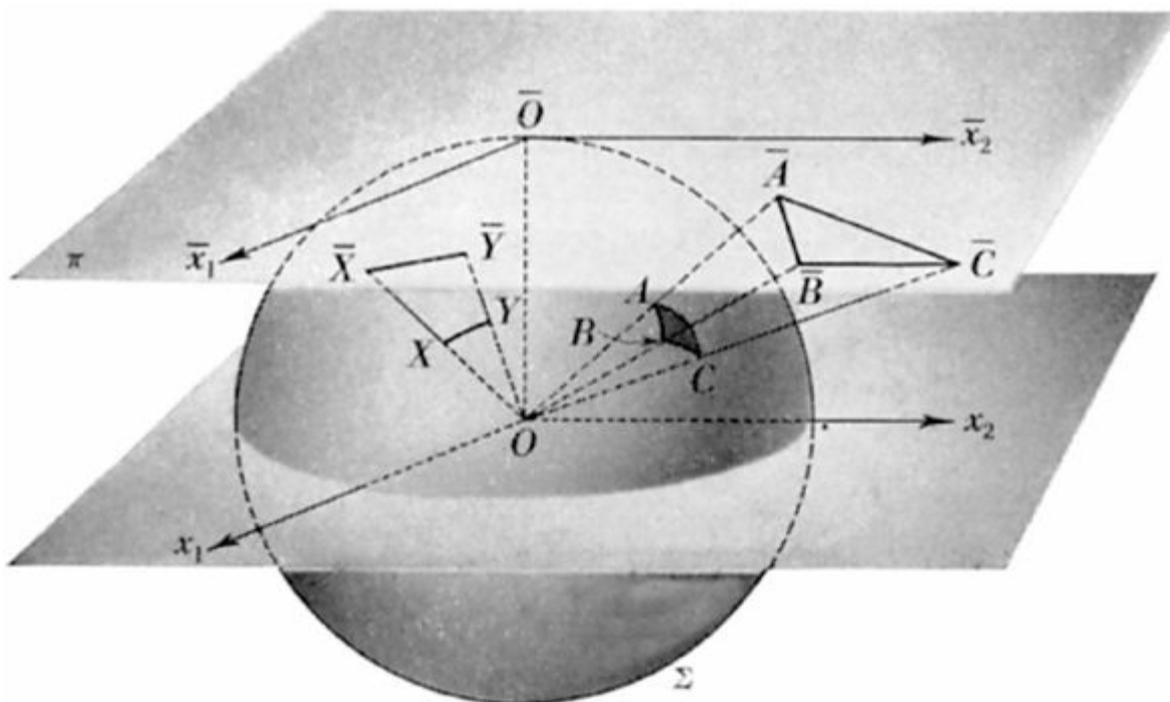
FIG. 8.11

($x1$, $x2$, $x3$) of any point X on Σ are related to those of its projection $\overline{X}$ in the tangent plane π by the relations

8.71
$$\frac{\overline{x}_1}{x_1} = \frac{\overline{x}_2}{x_2} = \frac{r}{x_3}$$

But here we have a *metric* which derives from that on the surface of the sphere Σ and which we can transfer in a nonambiguous manner to the tangent plane.

Using the notion of the inner product in 1.45, let us define the distance between two points X, Y on Σ to be $d(XY)$ so that

8.72
$$d(XY) = r\,\cos^{-1}\left\{\frac{x_1y_1 + x_2y_2 + x_3y_3}{\sqrt{x_1^2 + x_2^2 + x_3^2}\,\sqrt{y_1^2 + y_2^2 + y_3^2}}\right\}$$

8.73
$$= r\,\cos^{-1}\left\{\frac{\overline{x}_1\overline{y}_1 + \overline{x}_2\overline{y}_2 + r^2}{\sqrt{\overline{x}_1^2 + \overline{x}_2^2 + r^2}\,\sqrt{\overline{y}_1^2 + \overline{y}_2^2 + r^2}}\right\}$$

taken positive. If, now, we denote the *distance* between $\overline{X}$ and $\overline{Y}$ in π by $d(\overline{X}\,\overline{Y})$, this is determined by setting

8.74
$$d(\overline{X}\overline{Y}) = d(XY)$$

It follows immediately from 8.73 that $d(\overline{X}\overline{Y})$ *is always finite* for $\overline{X}$ fixed and any $\overline{Y}$, and *every pair of lines intersects*. We complete our definition of the metric in the tangent plane by defining the measure of an angle to be the same as

that on the sphere Σ, so that

8.75
$$\Delta \overline{ABC} = \Delta ABC = (A + B + C - \pi)r^2 > 0$$

Thus all the formulas of Sections 8.1–8.3 applicable to the spherical triangle *ABC* apply also to $\overline{ABC}$, and the resulting geometry in π is called *elliptic*.

Actually, our mapping of the sphere Σ on the tangent plane is 2:1, since two diametrically opposite points on the sphere are mapped on the same point of the plane. Thus *the area of the elliptic plane is finite and equal to $2\pi r2$*. It follows immediately from 8.75 that

.76 In any triangle ABC in elliptic geometry,

$$A + B + C > \pi$$

In order to study analytically the effect of allowing *r* to increase, we write 8.73 in the form

8.77
$$d(\overline{X}\,\overline{Y}) = r \sin^{-1} \frac{1}{r} \left\{ \frac{r^{-2}(\overline{x}_1\overline{y}_2 - \overline{x}_2\overline{y}_1)^2 + (\overline{x}_1 - \overline{y}_1)^2 + (\overline{x}_2 - \overline{y}_2)^2}{[r^{-2}(\overline{x}_1^2 + \overline{x}_2^2) + 1][r^{-2}(\overline{y}_1^2 + \overline{y}_2^2) + 1]} \right\}^{1/2}$$

taken positive.

Since $\lim\limits_{r \to 0} \frac{1}{x} \sin^{-1}x = 1$, we conclude that

$$\lim_{r \to \infty} d(\overline{X}\,\overline{Y}) = \sqrt{(\overline{x}_1 - \overline{y}_1)^2 + (\overline{x}_2 - \overline{y}_2)^2}$$

which is Pythagoras' theorem, and *Euclidean geometry is a limiting case of elliptic or spherical geometry.*

8.8 HYPERBOLIC GEOMETRY

Suppose we replace *r* in 8.75 by *ir,* where *i2* = $-$ 1; then the condition 8.76 becomes

.81 In any triangle ABC in hyperbolic geometry

$$A + B + C < \pi$$

Moreover, if we wish our projection to be real we must also replace *x3* by *ix3* in 8.71 so that the sphere becomes an hyperboloid of two sheets with equation

8.821
$$x_1^2 + x_2^2 - x_3^2 = -r^2$$

The enveloping or asymptotic cone of this hyperboloid with equation

8.822 $x_1^2 + x_2^2 - x_3^2 = 0$

cuts the tangent plane π: $x3 = r$ in the circle

8.823 $\mathfrak{a}$: $\bar{x}_1^2 + \bar{x}_2^2 = r^2$

called the *absolute* in π. By such a substitution the distance function 8.77 becomes

8.83 $d(XY) = ir \sin^{-1} \dfrac{1}{ir} \left\{ \dfrac{-r^{-2}(\bar{x}_1\bar{y}_2 - \bar{x}_2\bar{y}_1) + (\bar{x}_1 - \bar{y}_1)^2 + (\bar{x}_2 - \bar{y}_2)^2}{[-r^{-2}(\bar{x}_1^2 + \bar{x}_2^2) + 1][-r^{-2}(\bar{y}_1^2 + \bar{y}_2^2) + 1]} \right\}^{1/2}$
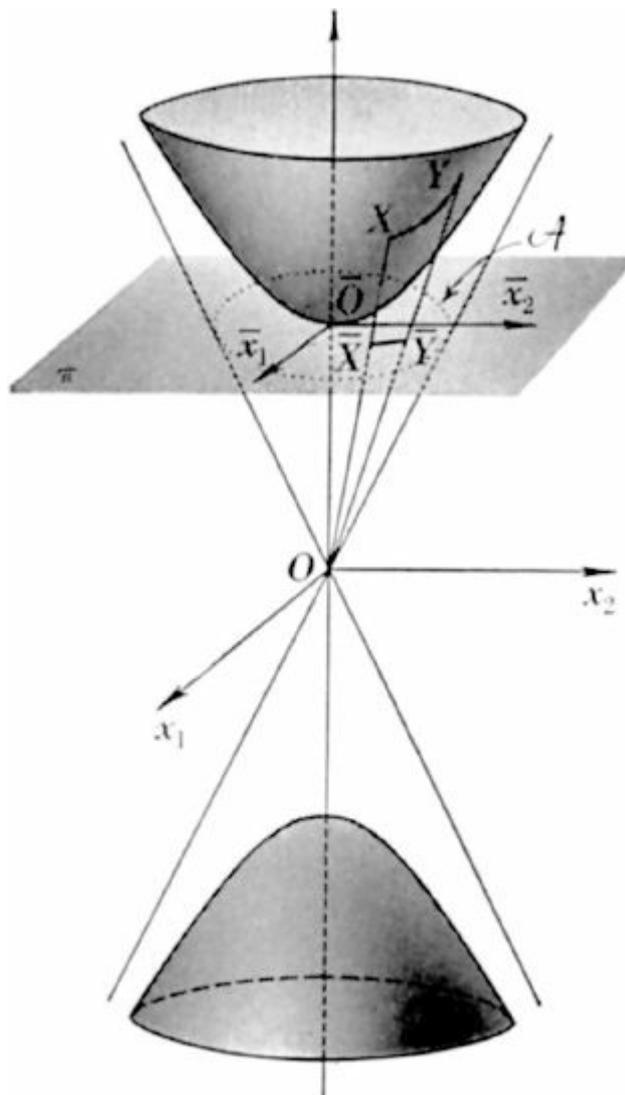
FIG. 8.12

taken positive.

In order to see the further consequences of the change of r into ir we must examine how it affects the *spherical* functions

$$\cos\theta = \frac{1}{2}\left(e^{i\theta} + e^{-i\theta}\right), \qquad \sin\theta = \frac{1}{2i}\left(e^{i\theta} - e^{-i\theta}\right)$$

If we define the corresponding *hyperbolic* functions by the formulas

$$\cosh\theta = \tfrac{1}{2}(e^{\theta} + e^{-\theta}), \qquad \sinh\theta = \tfrac{1}{2}(e^{\theta} - e^{-\theta})$$

then it is easy to see that $i\sin\theta = \sinh i\theta$ so that

$$i\sin^{-1}\frac{x}{i} = \sinh^{-1}x$$

and we can rewrite 8.83 in the real form

$$8.84 \qquad d(\overline{XY}) = r\sinh^{-1}\frac{1}{r}\left\{\frac{-r^{-2}(\bar{x}_1\bar{y}_2 - \bar{x}_2\bar{y}_1) + (\bar{x}_1 - \bar{y}_1)^2 + (\bar{x}_2 - \bar{y}_2)^2}{[-r^{-2}(\bar{x}_1^2 + \bar{x}_2^2) + 1][-r^{-2}(\bar{y}_1^2 + \bar{y}_2^2) + 1]}\right\}^{1/2}$$

taken positive.

There is one notable property of 8.84, namely, that $d(\overline{XY}) \to \infty$ as $\overline{Y}$ approaches the absolute α, for any point $\overline{X}$ within α. Thus points outside α are *inaccessible*. As before, the Euclidean metric is also a limiting case of the hyperbolic metric as $r \to \infty$.

Elliptic case   Hyperbolic case
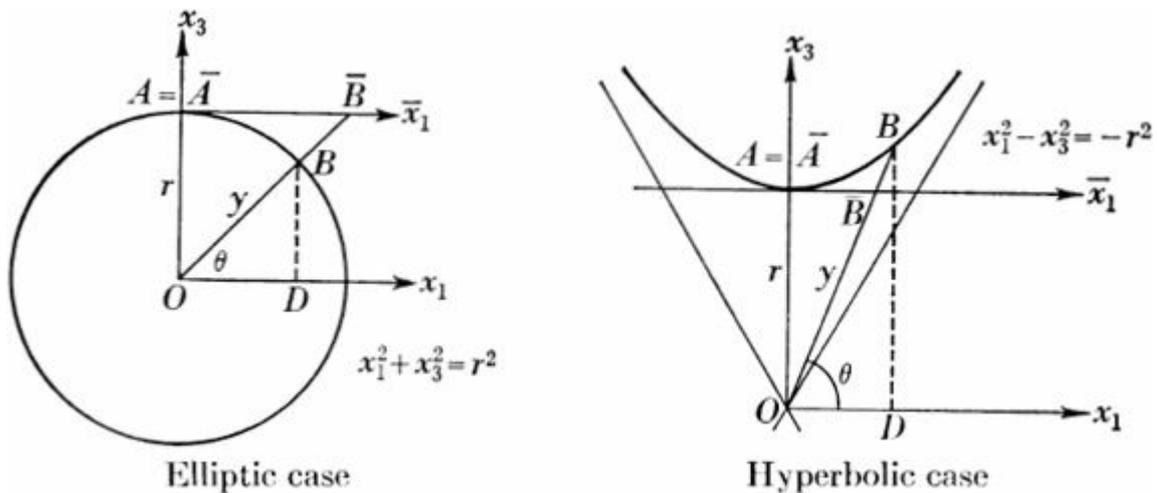
FIG. 8.13

Though the interpretation of the distance function 8.84 as the length of a geodesic on the hyperboloid is no longer valid, by a slight change of emphasis in the elliptic case we can obtain an interpretation which is consistent here. It will be sufficient to consider the case of a plane through the origin intersecting the sphere and the hyperboloid as indicated in Figure 8.13. Instead of defining the length $\overline{AB}$ in terms of the length AB we define it in terms of the *area* of the sector OAB. In the elliptic case the parametric equations of the circle are $x1 = x = r\cos\theta$, $x3 = y = r\sin\theta$, so that the area

8.85
$$OABD = \int y \, dx = r^2 \int_{\pi/2-\varphi}^{\pi/2} \sin^2\theta \, d\theta = \tfrac{1}{2}r^2 \int_{\pi/2-\varphi}^{\pi/2} (1 - \cos 2\theta) \, d\theta$$
$$= \tfrac{1}{2}r^2\varphi + \Delta OBD$$

and in the hyperbolic case $x1 = x = r \cosh \theta$ and $x3 = y = r \sinh \theta$, so that the corresponding area

8.86
$$OABD = \int y \, dx = -r^2 \int_{\pi/2-\varphi}^{\pi/2} \sinh^2\theta \, d\theta$$
$$= \tfrac{1}{2}r^2 \int_{\pi/2-\varphi}^{\pi/2} (1 - \cosh 2\theta) \, d\theta = \tfrac{1}{2}r^2\varphi + \Delta OBD$$

where $d \cosh \theta = \sinh \theta \, d\theta$ and $\cosh 2\theta = 1 + 2 \sinh 2\theta$, as may be easily verified.

Thus in both the elliptic and the hyperbolic case the area of the segment $\overset{\frown}{OAB}$ is $\tfrac{1}{2} r2\varphi$, and

$$d(\overline{AB}) = \frac{2}{r} \cdot \overset{\frown}{OAB}$$

interpreted in the elliptic case by 8.77 and in the hyperbolic case by 8.84. The factor of proportionality $2/r$ is fixed in a given geometry.

In order to obtain the formulas of *hyperbolic trigonometry* we must make the appropriate changes in 8.11, 8.12, and 8.21, which yield

8.871
$$\cosh a = \cosh b \cosh c - \sinh b \sinh c \cos A$$

8.872
$$\frac{\sin A}{\sinh a} = \frac{\sin B}{\sinh b} = \frac{\sin C}{\sinh c}$$

8.873
$$\cos B = -\cos B \cos C + \sin B \sin C \cosh a$$

Thus we are led to a model of the *hyperbolic* plane in which we limit attention to the interior of the absolute circle ɑ. As we have seen, any point on ɑ is to be considered "*at infinity.*" Through X can be drawn *two* parallels $XY\infty$, $XZ\infty$ to any line l not passing through X, as well as any number of *ultraparallels* which intersect l in *inaccessible* points P outside ɑ. Thus two lines may intersect, be parallel, or they may not intersect at all in hyperbolic geometry.
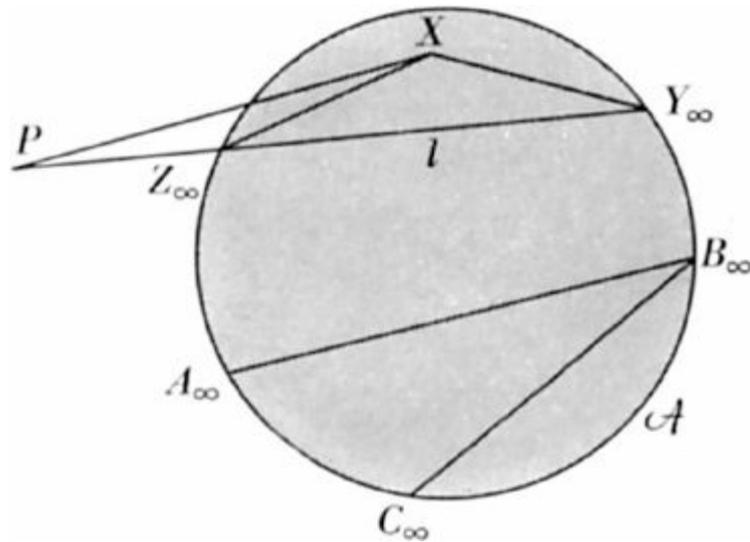
FIG. 8.14

It is beyond the scope of this brief treatment to give a proper definition of "angle" in hyperbolic geometry, but we can say that *perpendicularity* is definable in terms of the harmonic property:

.88 Definition If l is the polar line of an inaccessible point L with regard to the absolute α, then any line "through" L is perpendicular to l It follows after a little argument, which we omit, that:

(i) Intersecting lines in hyperbolic geometry have no common perpendicular. In elliptic geometry every pair of lines *has* a common perpendicular.

(ii) Parallel lines have no common perpendicular in hyperbolic geometry.

(iii) Nonintersecting lines have one common perpendicular, namely, the polar line of their inaccessible point of intersection.

Consider now the triangle $PQA\infty$ in which $PA\infty$ is parallel to l and $\angle PQA\infty = \pi/2$ as in Figure 8.15. If we set $\angle QPA\infty = \theta$ and $d(PQ) = p$ in 8.73, we conclude that

8.89 $$\sin \theta \cosh p = 1$$
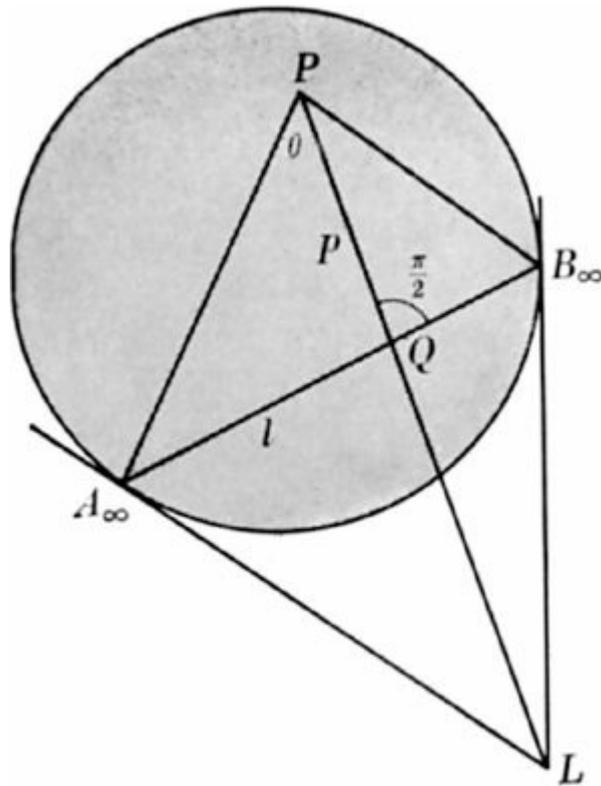
FIG. 8.15

Thus the perpendicular distance p from P to l is determined by and determines the angle θ, called the *angle of parallelism*. It can be shown that the locus of P, such that p remains fixed, is a conic having double contact with @ at $A\infty$ and $B\infty$. Such a curve is called an *equidistant curve*. Since all perpendiculars PQ pass through L, and equidistant curves become parallel lines as $r \to \infty$, we conclude that there is an infinite number of common perpendiculars to two parallel lines in Euclidean geometry, and these are themselves all parallel.

Thus *Euclidean geometry is a limiting case of both elliptic and hyperbolic geometry, and shares its properties in part with one and in part with the other*.

EXERCISES

1. Construct a triangle in the hyperbolic plane whose angle sum is zero.

2. How many such triangles exist having a given line as side?

3. Prove that $\angle QPA\infty = \angle QPB\infty$ in Figure 15.

4. Prove that $\theta \to \pi/2$ as $P \to Q$ along QL in Figure 15, by considering (a) the area of the triangle $PQA\infty$ (b) the relation 8.89.

REDUCTION OF REAL MATRICES TO DIAGONAL FORM

9.1 INTRODUCTION

In this our final chapter, we emphasize two important ideas. In the first place, the linear transformation $X' = PX$ is capable of two interpretations. Either it can, when applied to a vector space υn, be interpreted as a *mapping* of the space on itself in which a vector X is "moved" to a new position X' referred to the same basis vectors; or it can be interpreted as a *renaming* of the fixed vector X relative to a new system of basis vectors $E_i' = PEi$. These two points of view are complementary, and sometimes it is convenient to adopt one, sometimes the other.

The second important idea is the significance of the *characteristic vectors* of a linear transformation $Y = AX$. In Sections 9.3 and 9.4 we study the problem in the vector space υn and in the corresponding projective space $\wp n{-}1$; in Section 9.6 we see when a knowledge of these characteristic vectors of A can determine a matrix P such that $PAP{-}1$ is diagonal.

If A is *symmetric*, P can be chosen to be orthogonal and this leads in Sections 9.7 and 9.8 to the reduction of a general quadratic polynomial equation to normal form. We explain this reduction process against the background of Klein's Erlanger Programm in an attempt to show the significance of Section 9.5 in which the reduction by elementary transformation given in Chapter 3 is applied to symmetric matrices.

Though these ideas are not easy it is the author's hope that, by confining attention to real matrices and transformations and by continually emphasizing their geometrical importance in an algebraic context, the student may see their underlying meaning. To this end, numerous illustrative examples are worked out in the text. If this meaning can once be grasped, much mathematics studied later and most applications to modern physics and chemistry will become clear.

9.2 CHANGE OF BASIS

In order to combine the two interpretations which can be put upon a linear transformation, let us assume that $|A| \neq 0$ and that the vector

9.21
$$Y = AX$$

is referred to the same basis vectors Ei $(i = 1, 2, \dots n)$ of υn, while for $|p| \neq 0$,

9.22
$$X' = PX$$

maps the whole space υn upon itself, in particular the vectors Ei upon the vectors $PEi = E_i'$ As we saw in 3.78, E_i' is the ith column vector of the matrix P.

If we multiply 9.21 by the $n \times n$ matrix P, we have

9.231
$$PY = PAX = PAP^{-1}PX$$

and substituting from 9.22 we obtain

9.232
$$Y' = PAP^{-1}X'$$

referred to the new basis vectors E_i' $(i = 1, 2, \dots n)$. Note that the *matrix* of the transformation is now $PAP{-}1$.

The requirement that $|P| \neq 0$ is important, since we want the E_i' to be linearly independent vectors in the transformed space which we shall denote by υ_n'. But if, in addition, P is *orthogonal*, then

9.241
$$P^t P = P^{-1}P = I$$

which implies that

9.242
$$E_i' E_j' = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

and the basis E_i' is *orthonormal* υ_n'. Conversely, the condition 9.241 is a consequence of 9.242, as in 6.17. We illustrate these simple but important ideas in the following examples.

.25 Example. Let us take basis vectors $E1$ $E2$ as in Figure 9.1, and also

9.251
$$A = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

With

$$P = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

so that

9.252 $Y' = PAP^{-1}X' = AX'$

Since a rotation about the origin is still a rotation through the same angle whether referred to $E1$ $E2$, or E'_1, E'_2, P and A *commute* and $PAP-l = A$, as indicated in 9.252.

$E'_2(-\sin\theta, \cos\theta)$ $E_2(0,1)$

$E'_1(\cos\theta, \sin\theta)$

θ

$E_1(1,0)$
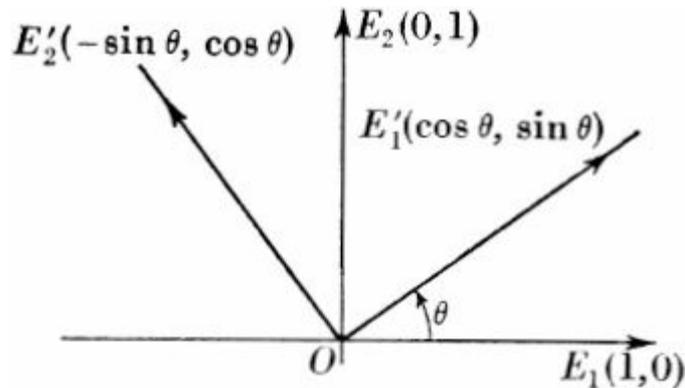
FIG. 9.1

.26 *Example*. The situation is quite different in space. Consider first the simple reflection 4.5

9.261
$$Y = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} X$$

with

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

so that

9.262
$$Y' = PAP^{-1}X' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} X'$$

and $E'_1 = E2$, $E'_2 = E1$, $E'_3 = E3$. Clearly the geometric transformation itself remains unchanged; it is only described as a reflection in the plane $x'_2 = 0$ instead of in the plane $x1 = 0$.

.27 *Example*. To take a more complicated case, let us suppose that

9.271
$$Y = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} X$$

is transformed by

$$X' = PX = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 3 \\ 1 & 0 & 2 \end{pmatrix} X$$

$$X = P^{-1}X' = \begin{pmatrix} 2 & 0 & -1 \\ -3 & -1 & 3 \\ -1 & 0 & 1 \end{pmatrix} X'$$

so that (cf. the example of Section 4.5)

$$Y' = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 3 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & -1 \\ -3 & -1 & 3 \\ -1 & 0 & 1 \end{pmatrix} X'$$

9.272
$$= \begin{pmatrix} -4 & -1 & 4 \\ -5 & 0 & 4 \\ -5 & -1 & 5 \end{pmatrix} X'$$

Again, the geometrical property of reflection is unchanged, but on substitution from $X = P^{-1}X'$ the equation of the plane $x1 = x2$ becomes

$$2x_1' - x_3' = -3x_1' - x_2' + 3x_3'$$

or

9.273
$$5x_1' + x_2' - 4x_3' = 0$$

with $E_1' = (1,0,1)$, $E_2' = (0,-1,0)$, $E_3' = (1,3,2)$. In order to verify the fact that we still have a reflection in the plane 9.273, we note that

9.274
$$\begin{pmatrix} -4 & -1 & 4 \\ -5 & 0 & 4 \\ -5 & -1 & 5 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

if and only if the vector $(a1, a2, a3)$ lies in the plane 9.273.
    Finally, we draw attention to the fact that besides defining $E_i' = PE_i$ as the

*i*th column of the matrix P, we could write:

$$9.28 \qquad E_i' = (p_{1i}, p_{2i}, \ldots, p_{ni}) = p_{1i}E_1 + p_{2i}E_2 + \ldots + p_{ni}E_n$$

These relations express the linear dependence of the new basis on the old, and provide another way of describing the matrix P. One should be careful, however, to note that the matrix of coefficients of the Ei in 9.28 is Pt and *not P*.

EXERCISES

1. Transform the matrix A of 9.272 by the matrix

$$P = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

What is the equation of the locus of the fixed points of the transformation $Y = PAP{-}1X$? Verify your conclusion as in 9.274 above.

2. If the new basis elements are defined by the equations

$$\begin{aligned} E_1' &= E_1 - E_2 + E_3 \\ E_2' &= E_1 + E_2 - E_3 \\ E_3' &= -E_1 + E_2 + E_3 \end{aligned}$$

write the equations of the corresponding transformation $X' = PX$ and express X' linearly in terms of the E_i'.

9.3 CHARACTERISTIC VECTORS

First, we observe that in the case of 9.261 we are reflecting in the plane $x1 = 0$ and this becomes $x_1' = x_2'$, by applying the transformations

$$X' = \begin{pmatrix} \tfrac{1}{2}\sqrt{2} & \tfrac{1}{2}\sqrt{2} & 0 \\ -\tfrac{1}{2}\sqrt{2} & \tfrac{1}{2}\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} X, \qquad X = \begin{pmatrix} \tfrac{1}{2}\sqrt{2} & -\tfrac{1}{2}\sqrt{2} & 0 \\ \tfrac{1}{2}\sqrt{2} & \tfrac{1}{2}\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} X'$$

Moreover,

$$\begin{pmatrix} \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} & 0 \\ -\frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} & 0 \\ \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

so that 9.261 and 9.271 are *geometrically* the same. We have seen that 9.271 and 9.272 are *geometrically* the same, so that all three operations are *geometrically* the same, though they are described with respect to different basis vectors. These basis vectors are orthogonal in the case of 9.261 and 9.271 but not in the case of 9.272.

The important question which we ask now is, *how could we deduce the geometrical identity of the three transformations 9.261, 9.271, 9.272 from their matrices alone?* The answer is provided by considering the nature of their *fixed elements*. Such a line of thought was significant for the inversion transformation in Section 8.5, and we follow it again here. The only difference is that in this case we are dealing with *vectors* and it will be sufficient to ask that the *direction* remain fixed, though the *magnitude* may vary. With this in mind we look for the solutions of the three equations

9.31
$$\lambda X = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} X, \qquad \lambda X = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} X,$$
$$\text{(a)} \hspace{5cm} \text{(b)}$$
$$\lambda X = \begin{pmatrix} -4 & -1 & 4 \\ -5 & 0 & 4 \\ -5 & -1 & 5 \end{pmatrix} X$$
$$\text{(c)}$$

Since the corresponding scalar equations are homogeneous, they have nontrivial solutions if and only if

9.32
$$\begin{vmatrix} \lambda+1 & 0 & 0 \\ 0 & \lambda-1 & 0 \\ 0 & 0 & \lambda-1 \end{vmatrix} = 0, \qquad \begin{vmatrix} \lambda & -1 & 0 \\ -1 & \lambda & 0 \\ 0 & 0 & \lambda+1 \end{vmatrix} = 0,$$
$$\begin{vmatrix} \lambda+4 & 1 & -4 \\ 5 & \lambda & -4 \\ 5 & 1 & \lambda-5 \end{vmatrix} = 0$$

and *each of these three equations reduces to*

9.33
$$(\lambda+1)(\lambda-1)(\lambda-1) = 0$$

In order to treat the problem in general, we note that the matrix equation λX

$= AX$ is always equivalent to a set of homogeneous scalar equations, and for consistency we must have

9.34 $$|\lambda I - A| = 0$$

This is called the *characteristic equation* of A. But we also have the matrix equation

$$P(\lambda I - A)P^{-1} = \lambda I - PAP^{-1}$$

and taking determinants,

$$|P||\lambda I - A||P^{-1}| = |\lambda I - PAP^{-1}| = |\lambda I - A| = 0$$

From the identity of the characteristic equations of A and $PAP{-1}$ we conclude that

.35 *The roots of the characteristic equation of the matrix A are the same as those of the characteristic equation of the matrix PAP–1.*
We call these roots the *characteristic roots* of A, or the *eigenvalues* of A. The matrices A and PAP–l are said to be *similar.*

In order to interpret this algebraic result geometrically we return to the matrix equation $\lambda X = AX$ and find a *characteristic* or *eigen vector Xi* which corresponds to the characteristic root λi. Let us consider our previous examples and, in particular, the three equations 9.31.

In 9.31 (a), setting $\lambda = +1$ we conclude that any vector satisfying the condition $x1 = -x1$ that is, $x1 = 0$, will remain fixed under 9.261. All such vectors lie in the *plane* $x1 = 0$. On the other hand if we set $\lambda = -1$, then the only solution of 9.31 (a) is of the form $(k,0,0)$ which is a vector normal to $x1 = 0$.

In 9.31 (b), setting $\lambda = +1$ we obtain the condition $x1 = x2$ which is the *form* the equation $x1 = 0$ takes with reference to the basis

$$E_1' = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right), \quad E_2' = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \quad E_3' = (0,0,1)$$

Again, the characteristic root $\lambda = -1$ leads to a solution $(k, -k,0)$ which is the corresponding characteristic vector normal to the plane $x1 = x2$.

Finally, if we set $\lambda = +1$ in 9.31 (c) we obtain the equations

9.36
$$
\begin{aligned}
-4x_1 - x_2 + 4x_3 &= x_1 \\
-5x_1 \quad\quad + 4x_3 &= x_2 \\
-5x_1 - x_2 + 5x_3 &= x_3
\end{aligned}
$$

which all coincide with 9.273. Setting $\lambda = -1$,

9.37
$$\begin{aligned}
-4x_1 - x_2 + 4x_3 &= -x_1 \\
-5x_1 \quad\quad + 4x_3 &= -x_2 \\
-5x_1 - x_2 + 5x_3 &= -x_3
\end{aligned}$$

and these equations have as solution space the vector (k,k,k), which is the way the normal to the plane $x1 = x2$ is described relative to the new basis. This is easily verified by transforming the vector $(1, -1, 0)$ according to

9.38
$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 3 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

Note that perpendicularity is no longer easily recognizable, since the transformation 9.38 is not orthogonal.

We draw the obvious conclusion from these illustrative examples that *it would be highly desirable if we could confine our attention to orthogonal transformations only*. Though this is not always possible, we shall see shortly how much can be accomplished with such a restriction.

EXERCISES

1. Determine the characteristic roots and the characteristic vectors of the linea transformation

$$Y = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 2 \end{pmatrix} X$$

and explain their geometrical significance.

2. Make the change of variables in Exercise 1 indicated by the transformation

$$X' = PX = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 3 \\ 1 & 0 & 2 \end{pmatrix} X$$

(cf. the example of Section 3.5) and verify that the characteristic roots remain unchanged. What are the characteristic vectors?

3. Euler's transformation for a rotation through an angle θ about an axis with direction cosines $l1, l2, l3$ can be written

$$X' = \begin{pmatrix} \cos\theta + & -l_3\sin\theta + & l_2\sin\theta + \\ l_1^2(1-\cos\theta) \, ' & l_1l_2(1-\cos\theta) \, ' & l_1l_3(1-\cos\theta) \\ l_3\sin\theta + & \cos\theta + & -l_1\sin\theta + \\ l_2l_1(1-\cos\theta) \, ' & l_2^2(1-\cos\theta) \, ' & l_2l_3(1-\cos\theta) \\ -l_2\sin\theta + & l_1\sin\theta + & \cos\theta + \\ l_3l_1(1-\cos\theta) \, ' & l_3l_2(1-\cos\theta) \, ' & l_3^2(1-\cos\theta) \end{pmatrix} X$$

Verify that any point on the axis remains fixed.

4. Obtain the various rotations of the cube in Section 4.5 as special cases of Euler's transformation in Exercise 3.

5. Obtain the various rotations of the cube in Section 4.5 as special cases of Rodrigues' transformation in Exercise 8 of Section 6.1

9.4 COLLINEATIONS

As we have done so often before, we can interpret the components of a 3-vector as the homogeneous coordinates of a point in the plane. Since the equation of a line l is homogeneous in $x1, x2, x3 = x0$, for $n = 3$ the linear transformation $Y = AX$ can be thought of as transforming the line l into a line l' described with reference to the same basis vectors. With this interpretation, we ask what points remain fixed under such a *collineation*? Since $X = \lambda X$ in homogeneous coordinates, as in Section 7.2, this is just the same question we asked before and leads to the characteristic equation

$$|\lambda I - A| = 0 = \begin{vmatrix} \lambda - a_{11} & -a_{12} & -a_{13} \\ -a_{21} & \lambda - a_{22} & -a_{23} \\ -a_{31} & -a_{32} & \lambda - a_{33} \end{vmatrix}$$

which, when expanded, yields the cubic

9.41 $$\lambda^3 + c_1\lambda^2 + c_2\lambda + c_3 = 0$$

with real coefficients. Since 9.41 has three roots $\lambda1, \lambda2, \lambda3$, there are the following possibilities:

.42 $\lambda1 = \lambda2 = \lambda3$. In this case every point in the plane remains fixed and the collineation is the identity transformation in the plane.

.43 $\lambda1 = \lambda2 \neq \lambda3$. In the Examples 9.25, 9.26, 9.27 the cubic 9.41 takes the form 9.33 and the transformations are called *homologies*. In the case of 9.261 the

center of the homology is the point (1,0,0) while the *axis* of the homology is the line $x1 = 0$. Every point $(0,x2,x3)$ of the axis remains fixed under 9.261, as also does every line through the center. To see this last statement it is only necessary to observe that such a line *PS* has equation

$$\lambda x_2 + \mu x_3 = 0$$
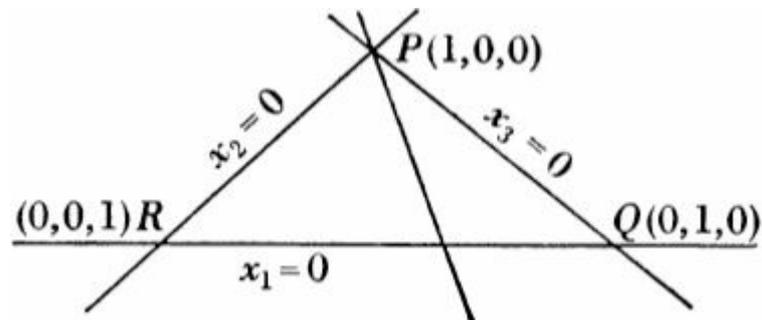
FIG. 9.2

which remains unchanged by 9.261, though a point $(1,\mu, -\lambda)$ will be transformed into the different point $(-1,\mu, -\lambda)$, also on *PS*.

9.44 $\lambda 1 \neq \lambda 2 \neq \lambda 3$. The Collineation

$$Y = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} X$$

with *a*, *b*, *c* real and different, leaves fixed each of $P(1,0,0)$, $Q(0,1,0)$, and $R(0,0,1)$, and the lines joining these points, but no other points remain fixed. For example, the point $(\lambda,\mu,0)$ on *PQ* is transformed into the *different* point $(a\lambda,b\mu,0)$ which also lies on *PQ*.

.45 $\lambda_1 = \bar{\lambda}_2 \neq \lambda_3$. In order to illustrate the case of only one real root, consider the transformation (cf. Example 9.25)

$$Y = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} X$$

which, in 3-space, is a rotation about the axis $x1 = x2 = 0$. The characteristic equation is

$$\begin{vmatrix} \lambda - \cos\theta & \sin\theta & 0 \\ -\sin\theta & \lambda - \cos\theta & 0 \\ 0 & 0 & \lambda - 1 \end{vmatrix} = 0$$

or

$$(\lambda^2 - 2 \cos \theta + 1)(\lambda - 1) = (\lambda - e^{i\theta})(\lambda - e^{-i\theta})(\lambda - 1) = 0$$

and $(0,0,1)$ is the only fixed point of the transformation.

EXERCISES

Describe the fixed elements of the collineations

(a) $Y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} X,$ (b) $Y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} X$

(c) $Y = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} X,$ (d) $Y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & a \end{pmatrix} X$

(e) $Y = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & a \end{pmatrix} X$

9.5 REDUCTION OF A SYMMETRIC MATRIX

In we studied the reduction of an arbitrary matrix A to canonical form PAQ by multiplying on the left and on the right with products of elementary matrices. The particular case in which A is *symmetric*, i.e., in which $a_{ij} = a_{ji}$ $(i,j = 1,2, \ldots, n)$, is important, since then we may perform the *same* operation on rows and columns by multiplying on the left by, say, P_i, and on the right by Q_i; $= Q_i = P_i^t$. In such a case

9.51 $\qquad P = P_s \ldots P_2 P_1, \qquad Q = P^t = P_1^t P_2^t \ldots P_s^t$

and *we can choose P so that PAPt is diagonal.* Moreover,

9.52 $\qquad (PAP^t)^t = (P^t)^t A^t P^t = PAP^t$

so that *PAPt* remains symmetric at every stage of the process.

Example. Consider the symmetric matrix

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

Evidently

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

and similarly

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

so that

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

In this case:

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

The chief reason we are so interested in symmetric matrices A is that every homogeneous quadratic form can be written $X^t A X$, so that the equation of any conic, quadric, etc., can be written

9.53 $$X^t A X + \text{linear terms} + \text{constant} = 0$$

in nonhomogeneous coordinates, or in the form $X^t A X = 0$ in homogeneous coordinates (cf. 7.25 and 7.26). If we make the change of basis represented by the linear transformation $X = P^t X'$, the quadratic form $X^t A X$ becomes

9.54 $$X'^t (P A P^t) X'$$

and we have seen how to choose P so that $PAPt$ is diagonal. Thus we can calculate the rank of A, which is also called the *rank* of the quadratic form $X^t A X$, and so classify conics, quadrics, etc., relative to this invariant.

EXERCISES

1. Write the equations of the conics 6.21, 6.22, 6.24 in homogeneous form $XtAX = 0$, and calculate the *rank* of A in each case.

2. Consider the conics in Exercise 1 as cylinders in space—how does this affect the rank of A?

3. Calculate the rank of A for each of the quadrics 6.31–6.36 when these are written in homogeneous form.

4. Calculate the rank of A if $XtAX = 0$ represents a pair of coplanar lines which (a) intersect, (b) are parallel, (c) coincide.

9.6 SIMILAR MATRICES

As we saw in 9.35, $PAP{-}1$ and A have the same characteristic equations and the same characteristic roots but *different* characteristic vectors. In fact, if X is characteristic for A with

9.61 $$\lambda X = AX$$

then

9.62 $$\lambda PX = PAP^{-1} \cdot PX$$

and PX is characteristic for $PAP{-}1$. If we assume that $PAP{-}1$ is diagonal then its characteristic vectors must be the basis vectors Ei, each associated with a particular characteristic root λi. It follows that in this case the characteristic vectors of A are Xi, where $= Ei = PXi$ or

9.63 $$X_i = P^{-1}E_i$$

Thus we conclude that

.64 *If $|P| \neq 0$ and if $PAP{-}1$ is diagonal, then the characteristic vectors of A are the column vectors of $P{-}1$.*

To state a converse theorem, we must know more than the characteristic roots since if some of these are repeated, various possibilities arise such as are referred to in the exercises at the end of Section 9.4. However, if the characteristic roots are all *distinct* we can make the following statement:

.65 *If the characteristic roots of A are all distinct, then A is similar to a diagonal matrix.*
Proof. It is only necessary to prove that the n characteristic vectors Xi,

corresponding to the n characteristic roots λ_i ($i = 1,2, \ldots,n$) are linearly independent, so that, taking them as the column vectors of $P{-}1$, we have a nonsingular matrix.

To this end we have from 9.61 that

9.66 $$\lambda_i X_i = AX_i \qquad i = 1,2, \ldots, n$$

If $c1, c2, \ldots cn$ are constants, not all zero, such that $\sum c_i X_i = 0$, then also

9.67 $$\Sigma\, c_i \lambda_i X_i = A\, \Sigma\, c_i X_i = 0$$

But the λ_i are all different real numbers so that 9.67 implies that

9.68 $$\Sigma'\, c_i(\lambda_i - \lambda_j)X_i = 0$$

where $\sum'$ omits the term with $i = j$, and j can take any value from 1 *to* n. But this yields a contradiction, since such a relation can be assumed to involve a minimum number of c's which are all different from zero and 9.68 then implies that the number could be reduced further by one. Thus no such relation exists and the X_i are linearly independent. It follows that the matrix $P{-}1$ is nonsingular and $PAP{-}1$ has characteristic vectors $E_i = PX_i$, from which we conclude that $PAP{-}1$ is diagonal with characteristic roots $\lambda 1, \lambda 2, \ldots \lambda n$.

Example. In order to illustrate these ideas, consider the matrix

$$A = \begin{pmatrix} -1, & 0, & 2 \\ -3, & 2, & 3 \\ -4, & 0, & 5 \end{pmatrix}$$

whose characteristic roots are $\lambda 1 = 1$, $\lambda 2 = 2$, $\lambda 3 = 3$. The corresponding characteristic vectors are obtained by solving the sets of equations

$$\begin{aligned} -\ x_1 \qquad\quad + 2x_3 &= x_1 \\ -3x_1 + 2x_2 + 3x_3 &= x_2 \\ -4x_1 \qquad\quad + 5x_3 &= x_3 \end{aligned} \qquad\qquad \begin{aligned} -\ x_1 \qquad\quad + 2x_3 &= 2x_1 \\ -3x_1 + 2x_2 + 3x_3 &= 2x_2 \\ -4x_1 \qquad\quad + 5x_3 &= 2x_3 \end{aligned}$$

$$\begin{aligned} -\ x_1 \qquad\quad + 2x_3 &= 3x_1 \\ -3x_1 + 2x_2 + 3x_3 &= 3x_2 \\ -4x_1 \qquad\quad + 5x_3 &= 3x_3 \end{aligned}$$

Thus $X1 = (\alpha, 0, \alpha)$, $X2 = (0,\beta,0)$, $X3 = (\gamma,3\gamma,2\gamma)$, where $\alpha \neq 0$, $\beta \neq 0$, $\gamma \neq 0$, so that

$$P^{-1} = \begin{pmatrix} \alpha & 0 & \gamma \\ 0 & \beta & 3\gamma \\ \alpha & 0 & 2\gamma \end{pmatrix}, \qquad P = \begin{pmatrix} \dfrac{2}{\alpha} & 0 & -\dfrac{1}{\alpha} \\[2mm] \dfrac{3}{\beta} & \dfrac{1}{\beta} & -\dfrac{3}{\beta} \\[2mm] -\dfrac{1}{\gamma} & 0 & \dfrac{1}{\gamma} \end{pmatrix}$$

It may be verified that

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

and the determination of the characteristic vectors has enabled us to make a change of basis, such that $PAP{-}1$ is diagonal. Note that P is determined up to arbitrary nonzero constants α, β, γ; in other words, the *magnitudes* of the characteristic vectors of A, which are also the basis vectors of $PAP{-}1$, are not important.

EXERCISES

1. Using the characteristic vectors of the transformation $Y = AX$ in Exercise 1 of Section 9.3, construct a matrix P such that $PAP{-}1$ is diagonal. Is the matrix P unique?

2. Construct the most general matrix which commutes with the matrix of collineation (b) in the exercises of Section 9.4 (set $AM = MA$ and equate coefficients) .

3. Interpret the most general matrix which commutes with the matrix of collineation (d) in the exercises of Section 9.4 with a $\neq$ 1, with reference to the three examples of Section 9.2 (cf. Section 9.3).

9.7 ORTHOGONAL REDUCTION OF A SYMMETRIC MATRIX

In Section 9.5 we saw that the *symmetry* of a matrix about its principal diagonal has certain consequences; a further consequence is the following :

.71 *The characteristic roots of a real symmetric matrix are all real.*
Proof. Let us assume that $A = A$t and if the equation 9.61 has a characteristic

root λ and characteristic vector X, then we may assume for the moment that the components $(x1, x2, \dots xn)$ of X are complex, and denote the conjugate vector by $\overline{X}$ with components $(\overline{x}1, \overline{x}2, \dots \overline{x}n)$. Thus

$$\lambda X'\overline{X} = X'A\overline{X}, \qquad \lambda \overline{X}'X = \overline{X}'AX$$

but since

$$X'\overline{X} = \overline{X}'X = \overline{X'X}$$

we conclude that

$$X'A\overline{X} = \overline{X}'AX = \overline{X'AX}$$

and both $Xt\overline{X}$ and $XtA\overline{X}$ are real. It follows that

$$\lambda = \overline{X}'AX/\overline{X}'X$$

since $\overline{X}tX \neq 0$, and λ must be real.

The question now arises, does this *reality* of the characteristic roots of a real symmetric matrix A have significance for the process of bringing A to diagonal form? We prove the following important result:

.72 *If A is a real symmetric matrix, there exists an orthogonal matrix P such that PAP−1 is diagonal.*

Proof. Our problem is to utilize the fact that A is symmetric in the argument of the preceding section. Since it has real characteristic roots λi, we conclude that the equations

$$\lambda_i X = AX$$

must define *real* characteristic vectors Xi. These vectors need not all be distinct, but we can assume that

$$|X_i| = \{X_iX_i\}^{1/2} = 1$$

as in Section 5.2.

Consider now one of these real characteristic vectors, say $X1$, whose components we may suppose to be

$$p_{11}, p_{12}, \dots p_{1n}$$

Since $\lambda 1 X1 = AX1$ we conclude that the subspace $\upsilon n{-}1$ of υn orthogonal to $X1$ is mapped on itself by A. Thus if we choose an orthonormal basis

$$q_{i1}, q_{i2}, \ldots q_{in} \qquad\qquad i = 2, \ldots n$$

in $\mho n-1$, the matrix P_1^{-1} with these as column vectors is *orthogonal:*

$$
P_1 = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \cdot \\ \cdot \\ \cdot \\ q_{n1} & q_{n2} & \cdots & q_{nn} \end{pmatrix}, \qquad
P_1^{-1} = \begin{pmatrix} p_{11} & q_{21} & \cdots & q_{n1} \\ p_{12} & q_{22} & \cdots & q_{n2} \\ \cdot \\ \cdot \\ \cdot \\ p_{1n} & q_{2n} & \cdots & q_{nn} \end{pmatrix} = P_1'
$$

It follows from 9.63 that $(1, 0, 0, \ldots 0)$ is a characteristic vector of $PAP{-}1$, and since $PAP{-}1. \mathbf{P}X1 = \lambda 1 \mathbf{P}X1$

$$P_1 A P_1^{-1} = \begin{pmatrix} \lambda_1 & 0 \\ B_1 & A_1 \end{pmatrix}$$

Since

$$(P_1 A P_1^{-1})^t = (P_1 A P_1')^t = P_1 A^t P_1' = P_1 A P_1^{-1}$$

$P1Ap_1^{-1}$ is symmetric, so that $B1 = 0$ and

$$P_1 A P_1^{-1} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{pmatrix}$$

where $A1$ is symmetric.

We have thus established the basis for an induction, since we can similarly construct an orthogonal matrix $P2$ such that

$$
P_2 P_1 A P_1^{-1} P_2^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & P_2' \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & P_2'^{-1} \end{pmatrix}
$$

$$
= \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & A_2 \end{pmatrix}
$$

where $A2$ is symmetric and $P2P1$ is again orthogonal. Proceeding thus, we reach the desired conclusion after at most n steps. Moreover, the matrix of transformation

$$P = P_n P_{n-1} \ldots P_2 P_1 = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & & p_{2n} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ p_{n1} & p_{n2} & & p_{nn} \end{pmatrix}$$

is orthogonal, and the columns of $P{-}1 = P\mathrm{t}$ (i.e., the rows of P) are *normal orthogonal characteristic vectors of A.*

EXERCISES

1. Find the characteristic vectors of the symmetric matrix

$$A = \begin{pmatrix} \frac{7}{3} & -\frac{2}{3} & 0 \\ -\frac{2}{3} & 2 & -\frac{2}{3} \\ 0 & -\frac{2}{3} & \frac{5}{3} \end{pmatrix}$$

and construct the orthogonal matrix P such that $\boldsymbol{PAP}{-}1$ is diagonal (cf. Exercise 9 of Section 6.1).

2. Given that the characteristic roots of the matrix

$$A = \begin{pmatrix} \frac{13}{9} & -\frac{4}{9} & \frac{2}{9} \\ -\frac{4}{9} & \frac{13}{9} & -\frac{2}{9} \\ \frac{2}{9} & -\frac{2}{9} & 1 \end{pmatrix}$$

are 1, 1, 2, carry out the construction of Section 9.7 to obtain an orthogonal matrix P such that $\boldsymbol{PAP}{-}1$ is diagonal.

9.8 THE REAL CLASSICAL GROUPS

The significance of the notion of a group of linear transformations in geometry goes back to Klein whose Erlanger Programm of 1872 laid the foundations for many of the ideas we have been studying in this book. The following results can easily be verified and are of great importance.

.81 The totality of all nonsingular real linear transformations on n variables forms a group called the *full linear group* $\mathfrak{gl}(n)$. Every such transformation represents a mapping of the vector space υn on itself.

.82 If we interpret the n variables $(x1, x2, \ldots xn)$ as homogeneous coordinates

of a point P in a projective space $\wp n{-}1$, then

$$(kx_1,\ kx_2,\ \ldots\ kx_n)$$

represents the same point P for all $k \neq 0$. This freedom in the choice of k enables us to restrict our attention to *unimodular* transformations, i.e., all transformations $\boldsymbol{Y} = \boldsymbol{AX}$ having $|A| = 1$ (cf. Section 7.2). Clearly, all such collineations form a subgroup of $\mathfrak{gl}(n)$ called the *unimodular group*.

.83 If we interpret $xn = x0 = 0$ as the equation of the *space at infinity (line at infinity* for $n = 3$), we have *affine geometry*. If we insist that our unimodular transformation $\boldsymbol{Y} = \boldsymbol{AX}$ leave $x0 = 0$ fixed, it follows that A must have the form

$$\begin{pmatrix} A_0 & T \\ 0 & 1 \end{pmatrix}$$

where $A0$ is a unimodular $(n-1) \times (n-1)$ matrix, and T is a column vector having $(n-1)$ components. Since

$$\begin{pmatrix} A_1 & T_1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} A_2 & T_2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} A_3 & T_3 \\ 0 & 1 \end{pmatrix}$$

all such affine transformations form a group called the *affine group* $\mathfrak{a}(n-1)$. We can distinguish two important subgroups:

.84 those transformations in which $T = 0$, and

.85 those transformations in which $A0 = I$.

.86 In the particular case where all the linear transformations $\boldsymbol{Y} = \boldsymbol{AX}$ are orthogonal, if AAt $= 1$, BBt $= 1$, then $(AB)(AB)$t $= \boldsymbol{AB}B$tAt $= 1$, so that we have a subgroup $\mathfrak{o}(n)$ of $\mathfrak{gl}(n)$, called the *orthogonal group*.

    The condition that A and so $|A|$ be real implies that the coefficients of the characteristic equation $|\lambda I - A| = 0$ are real, and further, since At$A = I$, that $|A| = \pm 1$. From this we conclude that if λ is a complex characteristic root, then its conjugate $\bar{\lambda}$ is also a characteristic root, and *the corresponding characteristic vectors are complex*. We have an example of this in 9.45, where it can be verified that the characteristic vectors corresponding to $\lambda 1 = e i\theta$, $\lambda 2 = e{-}i\theta$ are $X1 = (i,\ 1)$ and $X2 = (-i,\ 1)$. In order to speak of the *magnitude* of a complex vector we must modify slightly our definition, writing $|X| = \bar{X}$tX, which reduces to XtX if X is real; but we still have

$$\overline{\boldsymbol{Y}}{}'\boldsymbol{Y} = \overline{\boldsymbol{X}}{}'A'A\boldsymbol{X} = \overline{\boldsymbol{X}}{}^t\boldsymbol{X}$$

if A is orthogonal. It follows that if $Y = \lambda X$, where X is any characteristic vector of the transformation $Y = AX$, we must have $\lambda\bar{\lambda} = 1$. Thus

.861 *The characteristic roots of an orthogonal matrix are roots of unity and, since $|A| = \pm 1$, they occur in conjugate complex pairs.*

Just as in the case of the transformation of 9.45, these two conjugate complex characteristic vectors span a real plane, so by suitably choosing the basis vectors we have proved that

.862 *Any orthogonal $n \times n$ matrix A can be transformed by an orthogonal matrix P so that*

$$PAP^{-1} = \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 & & & & 0 \\ \sin\theta_1 & \cos\theta_1 & & & & \\ & & \cos\theta_2 & -\sin\theta_2 & & \\ & & \sin\theta_2 & \cos\theta_2 & \cdot & \\ & & & & & \cdot \\ 0 & & & & & \cdot \quad \pm 1 \end{pmatrix}$$

We may think of the orthogonal transformation $Y = AX$ as consisting of a succession of rotations and (or) reflections performed in mutually orthogonal subspaces. By combining two reflections to make a rotation through π we conclude that

.863 *Any orthogonal transformation may be considered as a succession of rotations in mutually orthogonal planes if $|A| = 1$, followed by a single reflection if $|A| = -1$.*

.87 If we assume that $A0$ in 9.83 is orthogonal we have the *Euclidean group*. For $n = 3$ the situation has been studied in Section 7.5, and the "circular points" 7.52 remain fixed under such a transformation (cf. Exercises 3, 4 of Section 7.5). In nonhomogeneous coordinates, any Euclidean transformation of the plane can be written

9.871
$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + l_1 \\ y_2 &= a_{21}x_1 + a_{22}x_2 + l_2 \end{aligned}$$

where the matrix

$$A_0 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

is orthogonal (i.e., a rotation or a reflection), and 9.871 can be written in vector form

9.872
$$Y = A_0X + T$$

Thus the two subgroups 9.84 and 9.85 of the affine group 9.83 turn out to be, in Euclidean geometry, the group of all orthogonal transformations, and the group of translations in the plane.

Klein's contention was that, while a geometry determines its group of collineations, conversely the group of collineations describes the geometry completely. A "theorem" (such as Pythagoras' theorem) is a relation which remains *invariant* under all collineations of the appropriate group. Since *angle, distance*, and so *area, volume*, and *generalized volume* are all defined in Euclidean geometry in terms of the inner product $X . X = XtX$, it follows from 6.17 and the invariance of XtX under translation that all these are invariant under any Euclidean transformation. This completes the proof of 5.57.

9.9 REDUCTION OF THE GENERAL CONIC TO NORMAL FORM

We conclude our study of vector geometry with a brief analysis of the Euclidean transformations necessary to bring to normal form the general conic (7.241):

9.91
$$ax_1^2 + bx_2^2 + 2hx_1x_2 + 2gx_1 + 2fx_2 + c = 0$$

Two methods are available.
First Method. If we write the quadratic terms of 9.91 in the form

9.92
$$X^tA_0X = (x_1, x_2) \begin{pmatrix} a & h \\ h & b \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

and find an orthogonal matrix P such that

$$PA_0P^{-1} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

as in the preceding section, then setting $X = P{-1}X'$, 9.91 becomes

9.93
$$\lambda_1 x_1'^2 + \lambda_2 x_2'^2 + 2g'x_1' + 2f'x_2' + c' = 0$$

Completing the squares we have

9.94
$$\lambda_1 \left(x_1' + \frac{g'}{\lambda_1} \right)^2 + \lambda_2 \left(x_2' + \frac{f'}{\lambda_2} \right)^2 = \frac{g'^2}{\lambda_1} + \frac{f'^2}{\lambda_2} - c'$$

as desired. The method is applicable for any n but we illustrate it in the plane only.

.95 Example. If we take 9.91 to be

9.951
$$x_1^2 + 2x_1x_2 - x_2^2 - 2x_1 + 2x_2 + 3 = 0$$

the quadratic terms can be written $X t A_0 X$ where

9.952
$$A_0 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

with the characteristic roots $\lambda 1 = \sqrt{2}$, $\lambda 2 = -\sqrt{2}$ and characteristic vectors $X1 = (1 + \sqrt{2}, 1)$, $X2 = (-1, 1 + \sqrt{2})$. Following the construction of the preceding section,

9.953 $\quad P = \dfrac{1}{k}\begin{pmatrix} 1 + \sqrt{2} & 1 \\ -1 & 1 + \sqrt{2} \end{pmatrix}, \quad P^{-1} = \dfrac{1}{k}\begin{pmatrix} 1 + \sqrt{2} & -1 \\ 1 & 1 + \sqrt{2} \end{pmatrix}$

where $k = \sqrt{4 + 2\sqrt{2}}$. With the substitution $X = P{-}1X'$, the equation 9.951 takes the form

$$\sqrt{2}x_1'^2 - \sqrt{2}x_2'^2 - \frac{2\sqrt{2}}{k}x_1' + kx_2' + 3 = 0$$

and completing squares we have

$$\sqrt{2}\left[x_1' - \frac{1}{k}\right]^2 - \sqrt{2}\left[x_2' - \frac{1}{k}(1 + \sqrt{2})\right]^2 = -4$$

If we set

9.954
$$x_1' = x_1'' + \frac{1}{k}$$

$$x_2' = x_2'' + \frac{1}{k}(1 + \sqrt{2})$$

we finally arrive at the normal form

9.955
$$-\frac{x_1''^2}{2\sqrt{2}} + \frac{x_2''^2}{2\sqrt{2}} = 1$$

of a rectangular hyperbola. We sum up the changes of variable in the vector equation

9.956 $X = P^{-1}X' = P^{-1}(X'' + C')$

or in the scalar equations

9.957 $x_1 = \dfrac{1}{k}(1 + \sqrt{2})x_1'' - \dfrac{1}{k}x_2''$

$x_2 = \dfrac{1}{k}x_1'' + \dfrac{1}{k}(1 + \sqrt{2})x_2'' + 1$

Second Method. If we examine the equation 9.956 we see that $P{-}\mathrm{l}X''$ represents a *rotation* to bring the coordinate axes parallel to the axes of the conic, while $P{-}1\ C'$ represents a *translation* of the origin to the center $C(0,1)$ of the conic, as is illustrated in Figure 9.3. It is natural to inquire whether we might not profitably reverse this order of procedure and first translate the origin to the center of the conic. If this were done in the case of the above example we would set $X = Y + C$ so that 9.951 would become

9.958 $y_1^2 - y_2^2 + 2y_1y_2 + 4 = 0$

and the subsequent rotation of the axes would involve the quadratic terms only.

The problem resolves itself, then, into finding the center C of the conic in question. This can be accomplished most conveniently with the help of the calculus as we shall explain in the Appendix, but an equivalent procedure is to insist that *the curve be symmetrical with regard to C*, i.e., that its equation remain unchanged when Y is replaced by $-Y$. Again referring to the preceding example, this would mean for 9.951 that

FIG. 9.3

$$(y_1 + c_1)^2 + 2(y_1 + c_1)(y_2 + c_2) - (y_2 + c_2)^2 - 2(y_1 + c_1) + 2(y_2 + c_2) + 3$$
$$\equiv (-y_1 + c_1)^2 + 2(-y_1 + c_1)(-y_2 + c_2) - (-y_2 + c_2)^2 - 2(-y_1 + c_1)$$
$$+ 2(-y_2 + c_2) + 3$$

identically, so that

$$4y_1(c_1 + c_2 - 1) + 4y_2(c_1 - c_2 + 1) \equiv 0$$

from which we conclude that

9.959 $$c_1 + c_2 - 1 = 0 = c_1 - c_2 + 1$$

and $c1 = 0$, $c2 = 1$ as we expected.

In order to clarify these two approaches to the reduction problem, we give the two corresponding factorizations of the Euclidean transformation, written in homogeneous form, first the direct transformation in 9.96 and then its inverse in 9.97:

9.96
$$\begin{pmatrix} P & -C' \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} I & -C' \\ 0 & 1 \end{pmatrix}\begin{pmatrix} P & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} P & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} I & -C \\ 0 & 1 \end{pmatrix}$$

$(\textit{first method})\qquad\qquad (\textit{second method})$

9.97
$$\begin{pmatrix} P^{-1} & C \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} P^{-1} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & C' \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} I & C \\ 0 & 1 \end{pmatrix}\begin{pmatrix} P^{-1} & 0 \\ 0 & 1 \end{pmatrix}$$

where $C' = PC$ and C is the center found above. As we have remarked already, the *second method* of procedure is usually preferable.

There remains the possibility that one of the characteristic roots of the matrix $A0$ in 9.92 should be zero. But this could happen only if $|A0| = 0$, in which case we could take $h = ka$, $b = kh = k2a$ so that not only is the rank of $A0$ equal to 1, but when a is divided out *the quadratic terms of 9.91 form a perfect square*. This is the necessary and sufficient condition that 9.91 represent a parabola, and the appropriate Euclidean transformation will bring it to normal form,

Let us summarize our conclusions with regard to the reduction of the quadratic equation

9.98
$$X'AX = X^t \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix} X = 0$$

in homogeneous coordinates:

.981 *If $|A| = 0$ the equation 9.98 factors and represents a pair of lines (6.53).*

.982 *If $|A| \neq 0$ and $|A0| = \begin{vmatrix} a & h \\ h & b \end{vmatrix} = 0$, then 9.98 represents a parabola.*

.983 *If $|A| \neq 0$ and $|A0| \neq 0$, then 9.98 represents a central conic which may be an ellipse or an hyperbola.*

This brings us back to Section 9.5 and explains why it is so important to

determine the *rank* of A. While this rank of A can be obtained most easily by using elementary transformations, we see now that such transformations are not orthogonal. Consequently they do not preserve the *form* of the curve. We must be clear as to our aims; if it is a case of seeing what a conic looks like, i.e., of plotting it, then we should use only Euclidean transformations so that this form is unchanged.
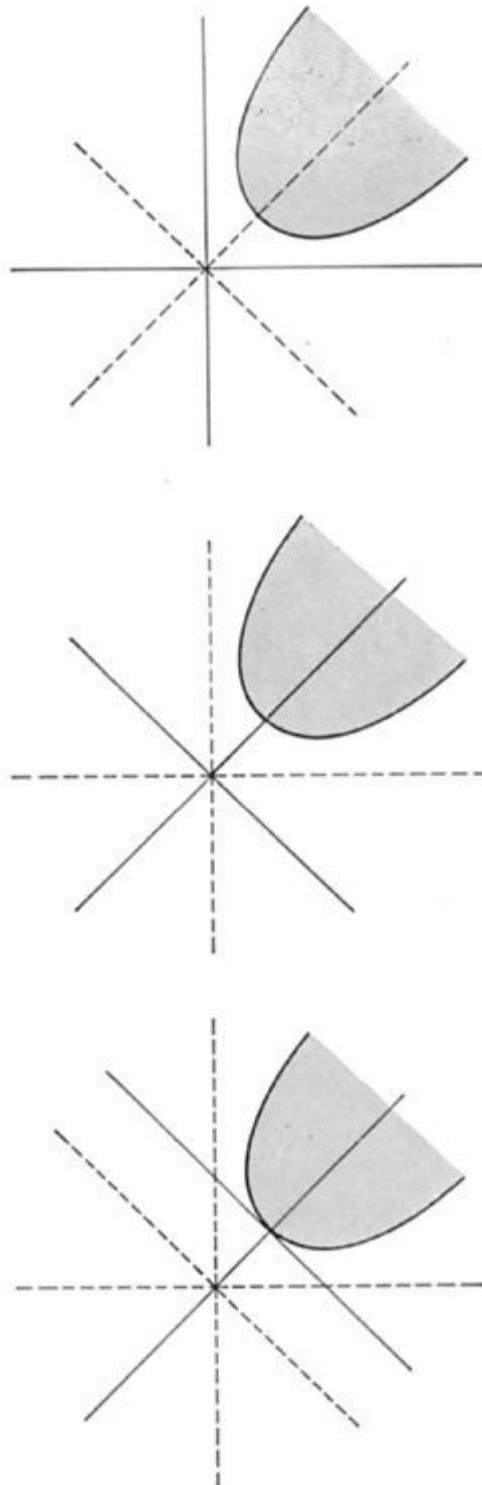
FIG. 9.4

EXERCISES

1. Plot the curve represented by the equation

$$x_1^2 + 2x_1x_2 + x_2^2$$
$$- 2x_1 + 2x_2 + 3 = 0$$

Solution. It is easy to verify that $\lambda 1 = 0$ and $\lambda 2 = 2$ so that the corresponding normal characteristic vectors are

$$\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) \quad \text{and} \quad \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

It follows that

$$P^{-1} = \begin{pmatrix} \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \\ -\dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \end{pmatrix}$$

and

$$P = \begin{pmatrix} \dfrac{1}{\sqrt{2}} & -\dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \end{pmatrix}$$

and substituting $X = P - 1X'$ we have

$$x_2'^2 = \sqrt{2}\left(x_1' - \frac{3}{2\sqrt{2}}\right)$$

Though the parabola has no center, we translate the origin to the *vertex*

$$V' = PV = \left(\frac{3}{2\sqrt{2}}, 0\right)$$

so that the equation takes the normal form

$$x_2''^2 = \sqrt{2}\, x_1''$$

2. Bring each of the following equations to normal form (i) by the first method, (ii) by the second method, and make diagrams to record the changes of coordinates and the shapes of the corresponding loci.

$$(a)\ x_1^2 - 3x_1x_2 + 2x_2^2 + x_2 - 1 = 0$$
$$(b)\ 2x_1^2 + 4x_1x_2 + 2x_2^2 - x_1 + x_2 + 1 = 0$$
$$(c)\ 3x_1^2 - 2x_1x_2 + 3x_2^2 - 4x_1 + 4x_2 + 1 = 0$$
$$(d)\ x_1^2 + 6x_1x_2 + x_2^2 + 2x_1 - 2x_2 = 0$$
$$(e)\ x_1^2 + x_2^2 - x_1 + x_2 + 1 = 0$$

3. Bring the quadric

$$x_1^2 + x_1x_2 + x_1x_3 + x_2x_3 + x_1 + x_3 = 0$$

to the normal form

$$x_1^2 - x_2^2 + x_3^2 = 1$$

(a) by the first method, (b) by the second method.

4. It can be proved that the quadric represented by the equation 7.26 in homogeneous coordinates:

$$X^tAX = 0$$

is (a) a plane counted twice, if A has rank 1; (b) two distinct planes, if A has rank 2; (c) a cone, if A has rank 3. Construct examples to illustrate each of these three cases.

5. Prove that the center of the quadric represented by the equation

$$f(x_1,x_2,x_3,x_0) \equiv X^tAX = 0$$

is given by the solution of the equations

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = \frac{\partial f}{\partial x_3} = 0$$

Show that this center is finite if $|a_{ij}| \neq 0$ for $i,j = 1,2,3$.

6. What are the conditions that 7.26 should represent an ellipsoid or hyperboloid?

7. What is the significance of the condition $|a_{ij}| = 0$ for $i,j = 1,2,3$? Give an example.

APPENDIX:
CALCULUS AND GEOMETRY

Like nearly all our mathematical ideas, the calculus had its origin in geometry. In particular, the study of tangents to curves by Fermat and the study of motion by Newton led to the development of the calculus in the 17th Century, though the notation we use today is largely due to Leibniz. Conversely, the expansion of a function in a Taylor's series has geometrical applications which are so significant that it seems unfair not to draw attention to them.

Perhaps the simplest of these applications is to a plane curve $\mathcal{e}$n given in the form

$$(1) \qquad\qquad x_2 = f(x_1)$$

where $f(x1)$ is a polynomial in $x1$ of degree n. If we expand the function $f(x1)$ about $x1 = a$, we have

$$(2) \qquad x_2 = f(a) + \frac{f'(a)}{1!}(x_1 - a) + \frac{f''(a)}{2!}(x_2 - a)^2 + \ldots + \frac{f^{(n)}(a)(x_1 - a)^n}{n!}$$

since the $(n + 1)th$ and all higher derivatives vanish. The advantage of writing (1) in the form (2) is that it enables us to *approximate* to $\mathcal{e}$n, in the neighborhood of the point $A(a, f(a))$. For example, we could ignore the terms in (2) of order higher than the first to obtain

$$(3) \qquad\qquad x_2 - f(a) = f'(a)(x_1 - a)$$

which is the equation of the *tangent* to the curve at A. If we substitute in (2) we find that the tangent (3) meets the curve in *two* coincident points if $f''(a) \neq 0$ and in *three* coincident points when $f''(a) = 0$ if $f'''(a) \neq 0$; in the latter case A is called a *point of inflection* of $\mathcal{e}$n.

Example 1. Consider the cubic curve $x2 = f(x1) = x_1^3$ so that

$$x_2 = a^3 + \frac{3a^2}{1!}(x_1 - a) + \frac{6a}{2!}(x_1 - a)^2 + \frac{6}{3!}(x_1 - a)^3$$

At the point $(a, a3)$ the equation of the tangent is

$$(4) \qquad\qquad x_2 - a^3 = 3a^2(x_1 - a)$$

which meets the curve in *three* coincident points if $a = 0$. The origin is thus a point of inflection with $x2 = 0$ as the inflectional tangent. The graph of the curve is shown in the figure.

We immediately remark, however, that the condition that $f(x)$ be a polynomial is undesirably restrictive. The best way out of this difficulty is to use homogeneous coordinates and write the equation of the curve ℓn in the form

$$f(x_1, x_2, x_0) = 0$$
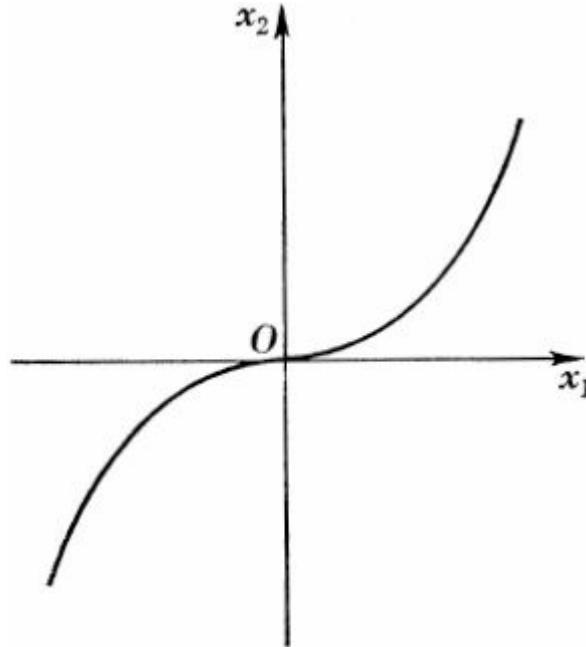
GRAPH

where $f(x1, x2, x0)$ is a homogeneous polynomial of degree n in $x1$, $x2$, $x0$. Fortunately, Taylor's theorem generalizes to any number of variables, though we must use partial differentiation to obtain the desired expansion. Let us suppose that the point $X(x1, x2, x0)$ lies on the curve ℓn and that $Y(y1, y2, y0)$ does *not* lie on ℓn. Then any intersection of XY with ℓn. has homogeneous coordinates

(5)
$$(\lambda x_1 + \mu y_1, \quad \lambda x_2 + \mu y_2, \quad \lambda x_0 + \mu y_0)$$

where $\lambda + \mu = 1$ as in 1.22, and we may write

$$f(\lambda x_1 + \mu y_1, \lambda x_2 + \mu y_2, \lambda x_0 + \mu y_0)$$

$$= \lambda^n f(x_1, x_2, x_0) + \frac{\lambda^{n-1}\mu}{1!}\left(y_1 \frac{\partial f}{\partial x_1} + y_2 \frac{\partial f}{\partial x_2} + y_0 \frac{\partial f}{\partial x_0}\right)$$

(6)
$$+ \frac{\lambda^{n-2}\mu^2}{2!}\left(y_1^2 \frac{\partial^2 f}{\partial x_1^2} + y_2^2 \frac{\partial^2 f}{\partial x_2^2} + y_0^2 \frac{\partial^2 f}{\partial x_0^2} + 2y_1 y_2 \frac{\partial^2 f}{\partial x_1 \partial x_2}\right.$$

$$\left. + 2y_1 y_0 \frac{\partial^2 f}{\partial x_1 \partial x_0} + 2y_2 y_0 \frac{\partial^2 f}{\partial x_2 \partial x_0}\right) + \ldots = 0$$

since we are assuming that the point (5) lies on ℓn.

The equation (6) can be written more compactly if we use the *differential*

operator

$$\Delta = y_1 \frac{\partial}{\partial x_1} + y_2 \frac{\partial}{\partial x_2} + y_0 \frac{\partial}{\partial x_0}$$

and the fact that $\Delta(n) \, f(x1, x2, x0) = n! f(y1, y2, y0)$, so that (6) becomes

$$(7) \qquad \lambda^n f(x_1, x_2, x_0) + \frac{\lambda^{n-1}\mu}{1!} \Delta f + \frac{\lambda^{n-2}\mu^2}{2!} \Delta^{(2)} f + \ldots + \mu^n f(y_1, y_2, y_0) = 0$$

The equation (7) is symmetrical in λ, μ and $\Delta(2)$, $\Delta(3)$, etc., denote the *symbolical* square, cube, etc., of the operator Δ; since $\Delta(n + 1)$ and all higher powers of Δ annihilate f, we may proceed as before.

By assumption, X lies on ℓn so that $f(x1, x2, x0) = 0$. If we look for the locus of Y such that XY meets ℓn. in *two* coincident points $\mu = 0$, we must have $\Delta f = 0$ so that the equation of the *tangent* at X is

$$(8) \qquad y_1 \frac{\partial f}{\partial x_1} + y_2 \frac{\partial f}{\partial x_2} + y_0 \frac{\partial f}{\partial x_0} = 0$$

where $y(y1, y2, y0)$ is the variable point and $X(x1, x2, x0)$ a fixed point on ℓn.

Example 2 Let $f(x1, x2, x0) = x2 x_0^2 - x_1^3$ as in Example 1; then

$$y_1 \frac{\partial f}{\partial x_1} + y_2 \frac{\partial f}{\partial x_2} + y_0 \frac{\partial f}{\partial x_0} \equiv -3y_1 x_1^2 + y_2 x_0^2 + 2y_0 x_2 x_0 = 0$$

which reduces to (4) if we set $x1 = a$, $x2 = a3$ and $x0 = y0 = 1$.

Let us now suppose that X does *not* lie on ℓn and that the n points of intersection of XY with ℓn are $X1 \ X2, \ldots Xn$. The relation

$$(9) \qquad \frac{n}{XY} = \frac{1}{XX_1} + \frac{1}{XX_2} + \ldots + \frac{1}{XX_n}$$

which generalizes 8.43 and defines the point Y for fixed X, can be rewritten in the form

$$\sum_{i=1}^{n} \left(\frac{1}{XY} - \frac{1}{XX_i} \right) = \sum_{i=1}^{n} \frac{YX_i}{XY \cdot XX_i} = 0$$

Since $XY \neq 0$ and $YXi{:}XXi = -\lambda i,{:}\mu i$, we must have

$$\sum_{i=1}^{n} \frac{\lambda_i}{\mu_i} = 0$$

which implies that

$$(10) \qquad \Delta f = y_1 \frac{\partial f}{\partial x_1} + y_2 \frac{\partial f}{\partial x_2} + y_0 \frac{\partial f}{\partial x_0} = 0$$

Thus $\Delta f = 0$ *represents the tangent at X to* ℭn, *if X lies on* ℭn, *and otherwise it represents the polar line of X.*

We have carried out this same analysis for $n = 2$ in Section 8.4 without using the calculus, but now we see that the method may be generalized and is applicable to any curve ℭn, if the polar line of X is defined as in (9) above.

Exercise Prove that if the polar line of X with regard to ℭn, passes through Y, then the polar line of Y passes through X.

It is interesting to apply these ideas to determine the condition that a general quadratic polynomial factors into the product of two linear factors. If we take X to be the point of intersection of the two lines represented by $f(x1, x2, x0) = 0$, then every line XY will meet these lines in *two* coincident points at X so that $\mu2 = 0$; for this to be so we must have $\Delta f = 0$ identically. Thus, taking f as in 6.51 (made homogeneous as in 7.243):

$$(11) \qquad \begin{aligned} \frac{1}{2}\frac{\partial f}{\partial x_1} &\equiv ax_1 + hx_2 + gx_0 = 0 \\[1em] \frac{1}{2}\frac{\partial f}{\partial x_2} &\equiv hx_1 + bx_2 + fx_0 = 0 \\[1em] \frac{1}{2}\frac{\partial f}{\partial x_0} &\equiv gx_1 + fx_2 + cx_0 = 0 \end{aligned}$$

Since these equations must be consistent, the condition 6.53 follows immediately. More generally, the condition that X be a double point of ℭn, is again just that

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = \frac{\partial f}{\partial x_3} = 0$$

and the Taylor expansion yields the machinery whereby such double points on ℭn, may be studied.

As a final illustration of the power of the method, consider the problem of finding the center of the general conic. In the context of Section 7.5, we saw in Exercises 5 and 6 of Section 8.4 that it is natural to define the center X as the pole of the line at infinity with equation $y0, = 0$. For (*10*) to so reduce, it is necessary and sufficient that the coordinates of the center satisfy the equations

$$(12) \qquad\qquad \frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = 0$$

Example 3 In the Example 9.95 the equation 9.951 becomes, in homogeneous coordinates,

$$f \equiv x_1^2 + 2x_1 x - x_2^2 - 2x_1 x_0 + 2x_2 x_0 + 3x_0^2 = 0$$

so that the equations (12) become

$$\frac{1}{2}\frac{\partial f}{\partial x_1} \equiv x_1 + x_2 - x_0 = 0$$

$$\frac{1}{2}\frac{\partial f}{\partial x_2} \equiv x_1 - x_2 + x_0 = 0$$

Setting $x0 = 1$ we have the nonhomogeneous coordinates of the center to be $(0,1)$, as in 9.959. With this information we can complete the reduction to normal form as previously explained.

It is clear that this approach is applicable to surfaces in space and to more general "varieties," but we shall pursue it no further.

ANSWERS TO EXERCISES

Section 1.2

2. $\dfrac{1}{\sqrt{3}}, \dfrac{1}{\sqrt{3}}, \dfrac{1}{\sqrt{3}}$

3. $x1 = 1, x2 = \tau, x3 = 2\tau$; no

4. $A'D$: $x1 = 1 - 2\tau, x2 = 1, x3 = 1$, etc.

5. AB: $\dfrac{1}{\sqrt{2}}, -\dfrac{1}{\sqrt{2}}, 0$

Section 1.3

1. Components $(2, -3, 1)$, magnitude $\sqrt{14}$

2. $\overrightarrow{AB}(-3, 1, -2)$, $\overrightarrow{BC}(5, -5, -5)$, $\overrightarrow{CA}(-2, 4, 7)$

3. $|\overrightarrow{AB}| = \sqrt{14}$, $\left(-\dfrac{3}{\sqrt{14}}, \dfrac{1}{\sqrt{14}}, -\dfrac{2}{\sqrt{14}}\right)$;
$\left(\mp\dfrac{3}{\sqrt{14}}, \pm\dfrac{1}{\sqrt{14}}, \mp\dfrac{2}{\sqrt{14}}\right)$

Section 1.5

1. $\overrightarrow{AB} = 3i - j - k$

2. $|\overrightarrow{OA} + \overrightarrow{OB}|2 = 11 = |\overrightarrow{OA} - \overrightarrow{OB}|2$

3. $|U|2 = 69, |V|2 = 65$

4. $\pm\dfrac{2}{\sqrt{3}}, \pm\dfrac{2}{\sqrt{3}}, \mp\dfrac{2}{\sqrt{3}}$

7. $(0, 0, 0)$

Section 1.6

2. ABC with equation $x1 + x2 + x3 = 1$ and $A'BC$ with equation $-x1 + x2 + x3 = 1$ intersect at an angle $\cos^{-1}(-\frac{1}{3})$. $AB'C'$ has equation $x1 - x2 - x3 = 1$ and is inclined to ABC at an angle $\cos^{-1}(\frac{1}{2})$. ABC is parallel to $A'B'C'$, etc.

3. $\overrightarrow{AB} = -E1 + E2$, $\overrightarrow{BC}' = -E2 - E3$, etc.

Section 2.1

1. 23, 18, 0

2. $x1 = \dfrac{7}{23}, x2 = -\dfrac{4}{23}, x3 = -\dfrac{7}{23}$

Section 2.3

2. 48, 0, 9

3. $x = y = z = 0$, $w = 2$

Section 2.4

2. $x1 = 1 - 2k$, $x2 = 1 - k$, $x3 = 3k$ with direction numbers $(2,1, -3)$; this direction is perpendicular to the normal to $x1 - 2x2 = 3$ and so parallel to the plane.

4. $x1 = 1 + 3k$, $x2 = 2k$, $x3 = -k$ and
$x1 = 3 + 3k'$, $x2 = 1 + 2k'$, $x3 = -k'$

5. (i) $x1 = 3 + k$, $x2 = 2 + k$, $x3 = k$;
(ii) $(\frac{1}{2}, \frac{5}{2}, \frac{3}{2})$

Section 2.5

2. $\begin{vmatrix} x_1 + 1 & x_2 - 2 & x_3 \\ 2 & -2 & 3 \\ 0 & 0 & -2 \end{vmatrix} = 0 = x1 + x2 - 1$

3. $x1 + x3 = 1$

4. $7x1 + 5x2 - 3x3 = 8$

5. $\begin{vmatrix} x_1 & x_2 & x_3 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \end{vmatrix} = 0$

Section 3.2

2.
$$X^2 = \begin{pmatrix} -5 & 2 & 1 \\ 2 & -2 & 2 \\ 1 & 2 & -5 \end{pmatrix}, \quad XY = \begin{pmatrix} -2 & 4 & -1 \\ -1 & -3 & 1 \\ 4 & 2 & -1 \end{pmatrix}$$

$$YX = \begin{pmatrix} -1 & 1 & -1 \\ 2 & -3 & 4 \\ 4 & -1 & -2 \end{pmatrix}, \quad Y^2 = \begin{pmatrix} -1 & 1 & 1 \\ -2 & -4 & 1 \\ 4 & -2 & -1 \end{pmatrix}$$

4. $X = 0 + X$;
$$Y = \begin{pmatrix} 1 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 1 \end{pmatrix} + \begin{pmatrix} 0 & \frac{3}{2} & 0 \\ -\frac{3}{2} & 0 & \frac{3}{2} \\ 0 & -\frac{3}{2} & 0 \end{pmatrix}$$

Section 3.3

2.
$$Y^{-1} = \frac{1}{4} \begin{pmatrix} 2 & -1 & 1 \\ 2 & 1 & -1 \\ 4 & 2 & 2 \end{pmatrix}$$

3.
$$A^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}, \quad (A')^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Section 3.4

1. 1, 2

Section 3.5

2.
$$A^{-1} = \begin{pmatrix} -2 & -3 & -4 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

$$B{-1} = -\tfrac{1}{16} \begin{pmatrix} -7 & 1 & 3 & -5 \\ -5 & 3 & -7 & 1 \\ 1 & -7 & -5 & 3 \\ 3 & -5 & 1 & -7 \end{pmatrix}$$

Section 3.7

1. $x1 = -1$, $x2 = 0$, $x3 = 2$, $x4 = 1$

2. (2,3,4,6)

3. (a) (9,17,22,34), (b) (4,1,2,2)

Section 4.2

3. The subgroup I, (12)(34), (13)(24), (14)(23) is normal in a_4 and also in s_4. Besides this "four group" the only normal subgroups in s_4 are I and a_4, along with s_4 itself.

Section 4.3

2. (AB) and (AC) lead to the linear transformations

$$Y = \begin{pmatrix} -\dfrac{1}{2} & \dfrac{\sqrt{3}}{2} \\ \dfrac{\sqrt{3}}{2} & \dfrac{1}{2} \end{pmatrix} X \quad \text{and} \quad Y = \begin{pmatrix} -\dfrac{1}{2} & -\dfrac{\sqrt{3}}{2} \\ -\dfrac{\sqrt{3}}{2} & \dfrac{1}{2} \end{pmatrix} X$$

Section 5.1

1. (1,0,1,3)

2.
$$Y = \tfrac{1}{2} \begin{pmatrix} 0 & 2 & 0 & 0 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & 1 \end{pmatrix} X$$

4. $X2 - X3 - X4 = 0$

5. $X2 - 2X1 = 0$, $X4 + 2X3 = 0$

Section 5.2

3. $Y1 = (1,1,0,0)$, $Y2 = (0,0,1,1)$
   $Y3 = (\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2})$, $Y4 (\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2})$

4. $Y1 = \left(\dfrac{1}{\sqrt{2}}, 0, \dfrac{1}{\sqrt{2}}\right)$, $Y2 = \left(\dfrac{1}{\sqrt{6}}, -\dfrac{2}{\sqrt{6}}, -\dfrac{1}{\sqrt{6}}\right)$,
   $Y3 = \left(-\dfrac{1}{\sqrt{3}}, -\dfrac{1}{\sqrt{3}}, \dfrac{1}{\sqrt{3}}\right)$

Section 5.3

1. 4

2. $(1,1,-1)$; $\sqrt{3}$

3. Fourth vertex can be chosen in three different ways; $\sqrt{6}$.

4. 8; $\frac{4}{3}$

Section 5.4

1. $\dfrac{x_1}{35} = \dfrac{x_2}{18} = \dfrac{x_3}{12}$

2. l: $x1 = 1 + s$, $x2 = 3s$, $x3 = 2s$
   n: $x1 = 1 + 2t$, $x2 = t$, $x3 = 6 + t$

3. $\dfrac{30}{\sqrt{35}}$

Section 5.5

1. 1

3. $\sqrt{2}$

5. $\dfrac{1}{6}$; $\dfrac{\sqrt{2}}{6}$

Section 5.8

1. If $Y1 = (1,1,0)$, $Y2 = (-\frac{1}{2}, \frac{1}{2}, 1)$, $Y3 = (1,-1,1)$, then
$$(1,0,1) = \tfrac{1}{2}Y_1 + \tfrac{1}{3}Y_2 + \tfrac{2}{3}Y_3$$

2. $x1 = t1 + t2$, $x3 = 3t2$, $x3 = t1$, $x4 = t2$;
   $(\frac{1}{2}, 3, -\frac{1}{2}, 1)$, $(1,0,1,0)$

Section 6.1

4. $P(\tfrac{1}{2},\tfrac{1}{2})$

5. $\begin{vmatrix} x_1^2 + x_2^2 & x_1 & x_2 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 13 & -2 & 3 & 1 \end{vmatrix} = 0$

7. $\begin{vmatrix} x_1^2 + x_2^2 + x_3^2 & x_1 & x_2 & x_3 \\ -1 & -1 & 0 & 0 \\ -1 & 0 & -1 & 0 \\ -1 & 0 & 0 & -1 \end{vmatrix} = 0; \quad r = \dfrac{\sqrt{3}}{2}$

Section 6.2

2. $x_1^2 - 2x_1x_2 + x_2^2 + 2x_1 + 2x_2 - 1 = 0$

5. $x_1^2 + 4x_1x_2 + x_2^2 - 2x_1 - 2x_2 = 0$

Section 6.3

1. $x_1 + x_2 = x_3,\ x_1 - x_2 = 1$
   $x_1 + x_2 = 3,\ 3(x_1 - x_2) = x_3$

2. $4x_1 - 2x_2 - x_3 - 3 = 0$

Section 6.4

1. $\dfrac{x_1^2}{2} + \dfrac{x_2^2}{3} + \dfrac{x_3^2}{2} = 1,\ \dfrac{x_1^2}{2} + \dfrac{x_2^2}{3} + \dfrac{x_3^2}{3} = 1$

   for the first conic, and similarly for the second.

2. (a) $x_2^2 + x_3^2 = 4x_1^2$, (b) $x_2^2 = 4x_1^2 + 4x_3^2$

4. $4a^2(x_1^2 + x_3^2) = (x_1^2 + x_2^2 + x_3^2 + b^2)^2$

Section 6.5

2. All real k; $x_1 - x_2 + 1 = 0 = 2x_1 + x_2 - k$

3. $(x_1 + 2x_2 - 1)(x_1 - x_2 + 1) = 0$

4.

$$X^t \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix} X = 0$$

$$X^t \begin{pmatrix} 4 & -1 & 1 \\ -1 & -2 & 2 \\ & & \end{pmatrix} X = 0$$

$$X^t \begin{pmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \\ 2 & -2 & 4 \end{pmatrix} X = 0$$

Section 6.6

1. $x1 = a \cos \theta \cosh \varphi$, $x2 = a \sin \theta \cosh \varphi$, $x3 = c \sinh \varphi$

5. If $|M - \lambda I| = 0$, then $XtMX - \lambda XtX = 0$ represents a pair of planes which meet the quadric $XtMX = 1$ in circles lying on the sphere $XtX = \lambda{-}1$.

Section 7.2

1. $x1 = 2t$, $x2 = t$, $x0 = 0$

2. $x1 = x2 = 0$, $x0 = 1$; $x1 = 0$, $x2 = \pm x0 = 1$;
   $x1 = 1$, $x2 = x0 = 0$

Section 7.5

5.
$$x_1' = \frac{a_{11}x_1 + a_{12}x_2 + a_{10}}{a_{01}x_1 + a_{02}x_2 + a_{00}}$$
$$x_2' = \frac{a_{21}x_1 + a_{22}x_2 + a_{20}}{a_{01}x_1 + a_{02}x_2 + a_{00}}$$

Section 7.8

2. The internal bisector of $\angle BAC$ is the right bisector of BC, and the external bisector of $\angle BAC$ is parallel to BC.

Section 8.2

1. For $A = \pi/2$: $\cos a = \cos b \cos c = \cot B \cot C$,

$$\frac{1}{\sin a} = \frac{\sin B}{\sin b} = \frac{\sin C}{\sin c}$$

Section 8.3

3. Area $\triangle ABC = \frac{1}{8}$ the area of Σ.

Section 8.4

2. $\dfrac{x_1 x_1'}{a^2} \pm \dfrac{x_2 x_2'}{b^2} = 1$, $x_2 x_2' = 2p(x_1 + x_1')$

5. Polar line of $(1,0,0)$ is $x1 = 0$ and that of $(0,1,0)$ is $x2 = 0$, and these lines intersect in the center of the ellipse.

Section 9.2

1. $4x1 - 5x2 - x3 = 0$

2.
$$X' = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix} X$$

$X' = x1_{E'_1} + x2_{E'_2} + x3_{E'_3}$

Section 9.3

1. $\lambda = 1, 2, 3$, with associated characteristic vectors $(1,0,1)$, $(0,1,0)$, $(1,0,-1)$

2. $(2,3,3)$, $(0,-1,0)$, $(0,-3,-1)$

Section 9.4

(a) every point fixed; (b) every point $(a,b,0)$ on $x3 = 0$ fixed and every line $x1 + \lambda x3 = 0$ fixed; (c) $(1,0,0)$ fixed and $x3 = 0$ fixed; (d) for $a \neq 1$, every point $(a,b,0)$ on $x3 = 0$ fixed and also $(0,0,1)$ fixed, so that every line through $(0,0,1)$ is fixed; (e) for $a \neq 1$, $(1,0,0)$ and $(0,0,1)$ fixed so that $x2 = 0$ is fixed.

Section 9.5

1. 3 in each case.

2. The rank is unchanged.

3. 4 in each case.

4. (a) 2, (b) 2, (c) 1

Section 9.6

1.
$$P = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix}; \quad \text{no}$$

2. (b)
$$\begin{pmatrix} a & 0 & b \\ c & d & e \\ 0 & 0 & d \end{pmatrix}$$

Section 9.7

1. $\left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)$, $\left(\frac{2}{3}, \frac{1}{3}, -\frac{2}{3}\right)$, $\left(\frac{2}{3}, -\frac{2}{3}, \frac{1}{3}\right)$

Section 9.9

2. (a) $(x1 - x2 - 1)(x1 - 2x2 + 1) = 0$
   (b) $2(x1 + x2)2 = (x1 - x2 - 1)$
   (c) $(x1 + x2)2 + 2(x1 - x2 - 1)2 = 1$
   (d) $2(x1 + x2)2 - (x1 - x2 - 1)2 = 1$

(e) $2(x_1 - \frac{1}{2})^2 + 2(x_2 + \frac{1}{2})^2 = -1$

INDEX

A

B

C

www.doverpublications.com