THEOREM 12.64 *If two lines are both parallel to a third in the same sense, there is a line meeting all three.*

Proof. We have to show that, if lines p and s are both parallel to r in the same sense, then the three lines p, r, s have a transversal. In affine geometry this is obvious, so let us assume the geometry to be hyperbolic. Of the two lines parallel to r through a point A on p, one is p itself. Let q be the other, and let r be in the angular region $\mathrm{p_1q_1}$, so that the rays $\mathrm{p_1}$ and $\mathrm{q_1}$ (from A) are parallel to r in opposite senses and s is parallel to r in the same sense as $\mathrm{p_1}$. Let B and D be arbitrary points on r and s, respectively.

If D is in the region $\mathrm{p_1q_1}$, the line AD is a transversal. If D is in $\mathrm{p_1q_2}$, BD is a transversal. If D is in $\mathrm{p_2q_2}$, both AD and BD are transversals. Finally, if D is in $\mathrm{p_2q_1}$, AB is a transversal.

Hyperbolic geometry will be considered further in Chapters 15, 16, and 20.

EXERCISES

1. If p is parallel to s and $[prs]$, then p is parallel to r. (See Figure 15.2c with s for q.)
2. Consider all the points strictly inside a given circle in the Euclidean plane. Regard all other points as nonexistent. Let chords of the circle be called lines. Then all the axioms 12.21–12.27, 12.41, and 12.51 are satisfied. Locate the two rays through a given point parallel to a given line. Note that they form an angle (as in Figure 16.2b).

13 Affine geometry

The first three sections of this chapter contain a systematic development of the foundations of affine geometry. In particular, we shall see how length may be measured along a line, though independent units are required for lines in different directions. In §§ 13.4–7 we shall investigate such topics as area, affine transformations, lattices, vectors, barycentric coordinates, and the theorems of Ceva and Menelaus. Finally, in § 13.8 and § 13.9, we shall extend these ideas from two dimensions to three.

According to Blaschke [**1,** p. 31; **2,** p. 12], the word "affine" (German *affin*) was coined by Euler. But it was only after the launching of Klein's Erlangen program (see Chapter 5) that this geometry became recognized as a self-contained discipline. Many of the propositions may seem familiar; in fact, most readers will discover that they have often been working in the affine plane without realizing that it could be so designated.

Our treatment is somewhat more geometric and less algebraic than that of Artin's *Geometric Algebra* [Artin **1;** see especially pp. 58, 63, 71]. Incidentally, we shall find that our Axiom 13.12 (which he calls DP) implies Theorem 13.122 (his $\mathrm{D_a}$): this presumably means that his Axiom 4b implies 4a.

13.1 THE AXIOM OF PARALLELISM AND THE "DESARGUES" AXIOM

Mathematical language is difficult but imperishable. I do not believe that any Greek scholar of to-day can understand the idiomatic undertones of Plato's dialogues, or the jokes of Aristophanes, as thoroughly as mathematicians can understand every shade of meaning in Archimedes' works.

M. H. A. Newman

(*Mathematical Gazette* **43,** 1959, p. 167)

In this axiomatic treatment, we regard the real affine plane as a special case of the ordered plane. Accordingly, the primitive concepts are *point* and *intermediacy,* satisfying Axioms 12.21–12.27, 12.41 and 12.51. Affine geometry is derived from ordered geometry by adding the following two extra axioms:

AXIOM 13.11 *For any point A and any line r, not through A, there is at most one line through A, in the plane Ar, which does not meet r.*

AXIOM 13.12 *If A, A', B, B', C, C', O are seven distinct points, such that AA', BB', CC' are three distinct lines through O, and if the line AB is parallel to $A'B'$, and BC to $B'C'$, then also CA is parallel to $C'A'$.*

The affine axiom of parallelism (13.11) combines with 12.62 to tell us that, for any point A and any line r, there is exactly one line through A, in the plane Ar, which does not meet r. Hence the two rays from A parallel to r are always collinear, *any two lines in a plane that do not meet are parallel,* and parallelism is an *equivalence relation.* The last remark comprises three properties:

Parallelism is *reflexive.* (Each line is parallel to itself.)
Parallelism is *symmetric.* (If p is parallel to r, then r is parallel to p.)
Parallelism is *transitive.* (If p and q are parallel to r, then p is parallel to q. Euclid I. 30.)

In the manner that is characteristic of equivalence relations, every line belongs to a *pencil* of parallels whose members are all parallel to one another.

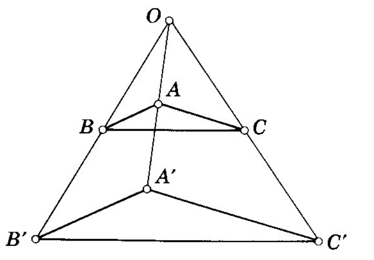

Figure 13.1a

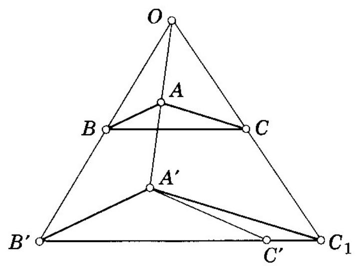

Figure 13.1b

Axiom 13.12 (see Figure 13.1*a*) is probably familiar to most readers either as a corollary of Euclid VI.2 or as an affine form of Desargues's theorem. We shall see that it implies

THEOREM 13.121 *If ABC and $A'B'C'$ are two triangles with distinct vertices, so placed that the line BC is parallel to $B'C'$, CA to $C'A'$, and AB to $A'B'$, then the three lines AA', BB', CC' are either concurrent or parallel.*

Proof. If the three lines AA', BB', CC' are not all parallel, some two of them must meet. The notation being symmetrical, we may suppose that these two are AA' and BB', meeting in O, as in Figure 13.1*b*. Let OC meet $B'C'$ in C_1. By Axiom 13.12, applied to AA', BB', CC_1, the line AC is parallel to $A'C_1$ as well as to $A'C'$. By Axiom 13.11, C_1 lies on $A'C'$ as well as on $B'C'$. Since $A'B'C'$ is a triangle, C_1 coincides with C'. Thus, if AA', BB', CC' are not parallel, they are concurrent [Forder **1**, p. 158].

Roughly speaking, Axiom 13.12 is the converse of one half of Theorem 13.121. The converse of the other half is

THEOREM 13.122 *If A, A', B, B', C, C' are six distinct points on three distinct parallel lines AA', BB', CC', so placed that the line AB is parallel to $A'B'$, and BC to $B'C'$, then also CA is parallel to $C'A'$.*

Proof. Through A' draw $A'C_1$ parallel to AC, to meet $B'C'$ in C_1, as in Figure 13.1*c*. By 13.121, applied to the triangles ABC and $A'B'C_1$, since AA' and BB' are parallel, CC_1 is parallel to both of them, and therefore also to CC'. Hence C_1 lies on CC' as well as on $B'C'$. Since the parallel lines BB' and CC' are distinct, B' cannot lie on CC'. Therefore C_1 coincides with C', and $A'C'$ is parallel to AC.

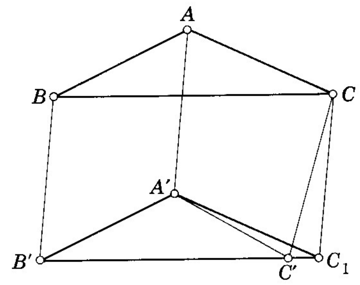

Figure 13.1c

EXERCISES

1. If a line in the plane of two parallel lines meets one of them, it meets the other also.

2. Can we always say, of three distinct parallel lines, that one lies between the other two?

13.2 DILATATIONS

Dilatations . . . are one-to-one maps of the plane onto itself which move all points of a line into points of a parallel line.

E. Artin [**1**, p. 51]

Four non-collinear points A, B, C, D are said to form a *parallelogram* $ABCD$ if the line AB is parallel to DC, and BC to AD. Its *vertices* are the four points; its *sides* are the four segments AB, BC, CD, DA, and its *diagonals* are the two segments AC, BD. Since B and D are on opposite sides of AC, the diagonals meet in a point called the *center* [Forder **1**, p. 140].

As in § 5.1, we define a *dilatation* to be a transformation which transforms each line into a parallel line. But now we must discuss more thoroughly the important theorem 5.12, which says that *two given segments, AB and $A'B'$, on parallel lines, determine a unique dilatation $AB \to A'B'$.*

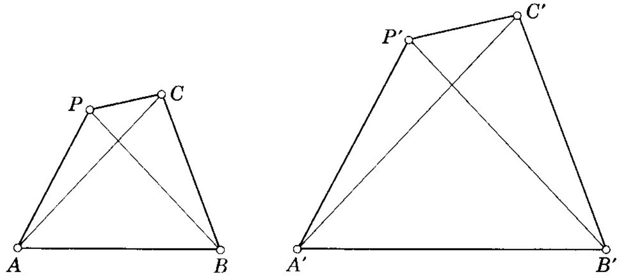

Figure 13.2a

For any point P, not on AB, we can find a corresponding point P' by drawing $A'P'$ parallel to AP, and $B'P'$ parallel to BP, as in Figure 5.1*a*. (The lines thus drawn through A' and B' cannot be parallel, for, if they were, AP and BP would be parallel.) Similarly, another point C yields C', as in Figure 13.2*a*. By 13.121, the three lines AA', BB', CC' are either concurrent or parallel. So likewise are AA', BB', PP'.

If the two parallel lines AB and $A'B'$ do not coincide, it follows that the four lines AA', BB', CC', PP' are all either concurrent or parallel. Then, by 13.12 or 13.122 (respectively), CP and $C'P'$ are parallel, so that the transformation is indeed a dilatation. If the lines AB and $A'B'$ do coincide, we can reach the same conclusion by regarding the transformation as $AC \to A'C'$ instead of $AB \to A'B'$.

We see now that a given dilatation may be specified by its effect on any given segment. The *inverse* of the dilatation $AB \to A'B'$ is the dilatation $A'B' \to AB$. The *product* of two dilatations, $AB \to A'B'$ and $A'B' \to A''B''$, is the dilatation $AB \to A''B''$. In particular, the product of a dilatation with its inverse is the *identity*, $AB \to AB$. Thus all the dilatations together form a (continuous) *group*.

The argument used in proving 5.13 shows that, for a given dilatation, the lines PP' which join pairs of corresponding points are *invariant* lines. The discussion of 5.12 shows that all these lines are either concurrent or parallel.

If the lines PP' are concurrent, their intersection O is an invariant point, and we have a *central* dilatation

$$OA \to OA'$$

(where A' lies on the line OA). The invariant point O is unique; for, if O and O_1 were two such, the dilatation would be $OO_1 \to OO_1$, which is the identity.

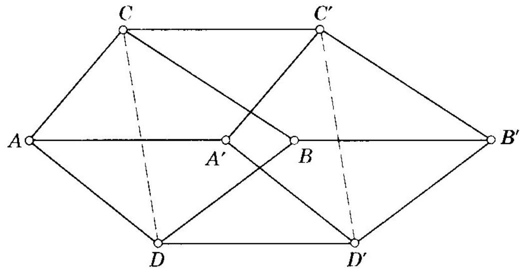

Figure 13.2b

If, on the other hand, the lines PP' are parallel, there is no invariant point, and we have a *translation* $AB \to A'B'$, where not only is AB parallel to $A'B'$ but also AA' is parallel to BB'. If these two parallel lines are distinct, $AA'B'B$ is a parallelogram. If not, we can use auxiliary parallelograms $AA'C'C$ and $C'CBB'$ (or $AA'D'D$ and $D'DBB'$) as in Figure 13.2*b*. Two applications of 13.122 suffice to prove that, when A, B, A' are given, B' is independent of the choice of C (or D). Hence

13.21 *Any two points A and A' determine a unique translation $A \to A'$.*

We naturally include, as a degenerate case, the identity, $A \to A$. It follows that a dilatation, other than the identity, is a translation if and only if it has no invariant point. Moreover, a given translation may be specified by its effect on any given point; in fact, the translation $A \to A'$ is the same as $B \to B'$ if $AA'B'B$ is a parallelogram, or if, for any parallelogram $AA'C'C$ based on AA', there is another parallelogram $C'CBB'$.

We next prove that dilatations are "ordered transformations:"

13.22 *The dilatation $AB \to A'B'$ transforms every point between A and B into a point between A' and B'.*

Proof. If the lines AB and $A'B'$ are distinct, the fact that $[ACB]$ implies $[A'C'B']$ follows at once from 12.401 (for a translation) or 12.402 (for a central dilatation). To obtain the analogous result for two corresponding triads on an invariant line CC', we draw six parallel lines through the six points, as in Figure 13.2*c*, and use the fact that $[acb]$ implies $[a'c'b']$.

To prove Theorem 3.21, which says that *the product of two translations is a translation,* we can argue thus: since translations are dilatations, the product is certainly a dilatation. If it is not a translation it has a unique invariant point O. If the first of the two given translations takes O to O', the second must take O' back to O. But the translation $O' \to O$ is the inverse of $O \to O'$. Thus the only case in which the product of two translations has an invariant point is when one of the translations is the inverse of the other. (By our convention, the product is still a translation even then.) Hence

13.23 *The product of two translations $A \to B$ and $B \to C$ is the translation $A \to C$.*

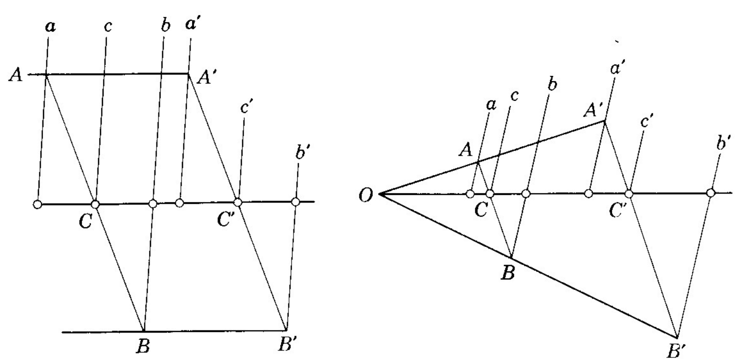

Figure 13.2c

To prove that this is a *commutative* product (as in 3.23), we consider first the easy case in which the two translations are along nonparallel lines. Completing the parallelogram $ABCD$, we observe that the translations $A \to B$ and $B \to C$ are the same as $D \to C$ and $A \to D$, respectively. Hence their product in either order is the translation $A \to C$:

$$\begin{aligned}(A \to B)(B \to C) &= (A \to D)(D \to C)\\ &= (B \to C)(A \to B).\end{aligned}$$

To deal with the product of two translations T and X along the same line, let Y be any translation along a nonparallel line, so that X commutes with both Y and TY. Then

$$\mathrm{TXY} = \mathrm{TYX} = \mathrm{XTY}$$

and therefore

$$\mathrm{TX} = \mathrm{XT}$$

[cf. Veblen and Young **2**, p. 76].

As a special case of 5.12, we see that any two distinct points, A and B, are interchanged by a unique dilatation $AB \to BA$, or, more concisely,

$$A \leftrightarrow B,$$

which we call a *half-turn*. (Of course, $A \leftrightarrow B$ is the same as $B \leftrightarrow A$.) If C is any point outside the line AB, the half-turn transforms C into the point D in which the line through B parallel to AC meets the line through A parallel to BC (Figure 13.2*d*). Therefore $ADBC$ is a parallelogram, and the same half-turn can be expressed as $C \leftrightarrow D$. The invariant lines AB and CD, being the diagonals of the parallelogram, intersect in a point O, which is the invariant point of the half-turn. It follows that any segment AB has a *midpoint* which can be defined to be the invariant point of the half-turn $A \leftrightarrow B$, and we have proved that the center of a parallelogram is the midpoint of each diagonal, that is, that the two diagonals "bisect" each other. To see how the

half-turn transforms an arbitrary point on AB, we merely have to join this point to C (or D) and then draw a parallel line through D (or C).

By considering their effect on an arbitrary point B, we may express any two half-turns as $A \leftrightarrow B$ and $B \leftrightarrow C$. If their product has an invariant point O, each of them must be expressible in the form $O \leftrightarrow O'$, that is, they must coincide. In every other case, there is no invariant point. Hence

13.24 *The product of two half-turns $A \leftrightarrow B$ and $B \leftrightarrow C$ is the translation $A \to C$.*

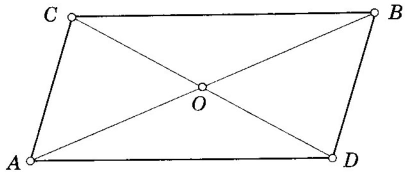

Figure 13.2d

We have seen (Figure 13.2*d*) that, if $ADBC$ is a parallelogram, the half-turn $A \leftrightarrow B$ is the same as $C \leftrightarrow D$, and the translation $A \to D$ is the same as $C \to B$. This connection between half-turns and translations remains valid when the parallelogram collapses to form a symmetrical arrangement of four collinear points, as in Figure 13.2*e*:

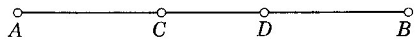

Figure 13.2e

13.25 *The half-turns $A \leftrightarrow B$ and $C \leftrightarrow D$ are equal if and only if the translations $A \to D$ and $C \to B$ are equal.*

In fact, the relation $(A \leftrightarrow B) = (C \leftrightarrow D)$ implies

$$\begin{aligned}(A \to D) &= (A \leftrightarrow B)(B \leftrightarrow D)\\ &= (C \leftrightarrow D)(D \leftrightarrow B) = (C \to B)\end{aligned}$$

and, conversely, the relation $(A \to D) = (C \to B)$ implies

$$\begin{aligned}(A \leftrightarrow B) &= (A \to D)(D \leftrightarrow B)\\ &= (C \to B)(B \leftrightarrow D) = (C \leftrightarrow D).\end{aligned}$$

In the special case when C and D coincide, we call them C' and deduce that C' is the midpoint of AB if and only if the translations $A \to C'$ and $C' \to B$

are equal. This involves the existence of parallelograms $AC'A'B'$ and $A'B'C'B$, as in Figure 13.2*f*. Completing the parallelogram $B'C'A'C$, we obtain a triangle ABC with A', B', C' at the midpoints of its sides. Hence

13.26 *The line joining the midpoints of two sides of a triangle is parallel to the third side, and the line through the midpoint of one side parallel to another passes through the midpoint of the third.*

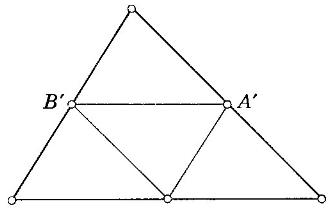

Figure 13.2f

Two figures are said to be *homothetic* if they are related by a dilatation, *congruent* if they are related by a translation or a half-turn. In particular, a directed segment AB is congruent to its "opposite" segment BA by the half-turn $A \leftrightarrow B$. Thus, in Figure 13.2*f*, the four small triangles

$$AC'B', \quad C'BA', \quad B'A'C, \quad A'B'C'$$

are all congruent, and each of them is homothetic to the large triangle ABC.

EXERCISES

1. Such equations as those used in proving 13.25 are easily written down if we remember that each must involve an even number of double-headed arrows (indicating half-turns). Explain this rule.

2. The translations $A \to C$ and $D \to B$ are equal if the translations $A \to D$ and $C \to B$ are equal. (This is obvious when $ADBC$ is a parallelogram, but remarkable when all the points are collinear.)

3. Setting $A = C$ in the equation

$$(A \leftrightarrow B)(B \to C) = (A \leftrightarrow C),$$

deduce that any given point C is the invariant point of a half-turn $(C \leftrightarrow B)(B \to C)$ which, by a natural extension of the symbolism, may be written as

$$C \leftrightarrow C.$$

4. If the three diagonals of a hexagon (not necessarily convex) all have the same midpoint, any two opposite sides are parallel (as in Figure 4.1*e*).

5. From any point A_1 on the side BC of a triangle ABC, draw A_1B_1 parallel to BA to meet CA in B_1, then B_1C_1 parallel to CB to meet AB in C_1, and then C_1A_2 parallel to AC to meet BC in A_2. If A_1 is the midpoint of BC, A_2 coincides with it. If not, continue the process, drawing A_2B_2 parallel to BA, B_2C_2 parallel to CB, and C_2A_3 parallel to AC. The path is now closed: A_3 coincides with A_1. (This is called Thomsen's figure. See Geometrical Magic, by Nev R. Mind, *Scripta Mathematica*, **19** (1953), pp. 198–200.)

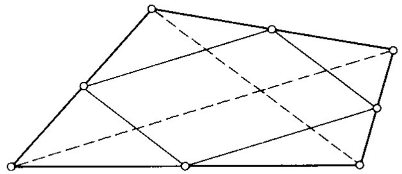

Figure 13.2g

6. The midpoints of the four sides of any simple quadrangle are the vertices of a parallelogram (Figure 13.2*g*; cf. Figure 4.2*c*). This theorem was discovered by Pierre Varignon (1654–1722). It shows that the *bimedians*, which join the midpoints of opposite sides of the quadrangle, bisect each other. Thus the corollary to Hjelmslev's theorem (§ 3.6) becomes an affine theorem when we replace the hypotheses 3.61 by

$$AB = BC, \quad A'B' = B'C'.$$

7. The midpoints of the six sides of any complete quadrangle are the vertices of a centrally symmetrical hexagon (of the kind considered in Ex. 4, above).

13.3 AFFINITIES

"Yes, indeed," said the Unicorn, . . . "What can we measure? . . . We are experts in the theory of measurement, not its practice."

J. L. Synge [**2**, p. 51]

The results of § 13.2 may be summarized in the statement that all the translations of the affine plane form a continuous Abelian group, which is a subgroup of index 2 in the group of translations and half-turns; and the latter is a subgroup (of infinite index) in the group of dilatations [Veblen and Young **2**, pp. 79, 93].

Moreover, the group of translations is a *normal* subgroup (or "self-conjugate" subgroup)* in the group of dilatations, that is, if T is a translation while S is a dilatation, then $S^{-1}TS$ is a translation [Artin **1**, p. 57]. To prove this, suppose if possible that the dilatation $S^{-1}TS$ has an invariant point. Since this invariant point could have been derived from a suitable point O by applying S, we may denote it by O^S. Thus $S^{-1}TS$ leaves O^S invariant. But $S^{-1}TS$ transforms O^S into O^{TS}. Hence $O^{TS} = O^S$. Applying S^{-1}, we deduce $O^T = O$, which is absurd (since T has no invariant point).

If T is $A \to B$ and S is $AB \to A^SB^S$, then $S^{-1}TS$ is $A^S \to B^S$. Accordingly, it is sometimes convenient to write T^S for $S^{-1}TS$ [see, e.g., Coxeter **1**, p. 39] and to say that the dilatation S *transforms* the translation T into the translation T^S. (Since A^SB^S is parallel to AB, T^S has the same direction as T.) In other words, a dilatation transforms the group of translations into

* Birkhoff and MacLane **1**, p. 141; Coxeter **1**, p. 42.

itself in the manner of an *automorphism:* if it transforms T into T^S and another translation U into U^S, it transforms the product TU into $(TU)^S = T^S U^S$ and any power of T into the same power of T^S.

It is convenient to use the italic letter T for the point into which the translation T transforms an arbitrarily chosen initial point (or origin) 1. Then, if a central dilatation S has 1 as its invariant point, it not only transforms T into T^S but also transforms T into T^S.

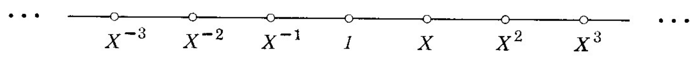

Figure 13.3a

Applying to the arbitrary point 1 all the integral powers of a given translation X, we obtain a *one-dimensional lattice* consisting of infinitely many points "evenly spaced" along a line, as in Figure 13.3*a*. We may regard every such point X^μ as being derived from the point X by a dilatation $1X \to 1X^\mu$ (which leaves the point 1 invariant). At first we take μ to be an integer; but since the same dilatation transforms each X^n into

$$(X^\mu)^n = X^{\mu n},$$

we can consistently extend the meaning of X^μ so as to allow μ to have any rational value, and finally any real value. In other words, we can interpolate new points between the points of the one-dimensional lattice and then define X^μ, for any real μ, to mean the translation $1 \to X^\mu$. The details are as follows.

For each rational number $\mu = a/b$ (where a is an integer and b is a positive integer) we derive from the point X a new point X^μ by means of the dilatation $1X^b \to 1X^a$. A convenient way to construct this point X^μ is to use the lattice of powers of an arbitrary translation Y along another line through the initial point 1, drawing a line through the point Y parallel to the join of the points Y^b and X^a, as in Figure 13.3*b* (cf. Figure 9.1*c*).

To verify that the order of such points X^μ agrees with the order of the rational numbers μ, we take three of them and reduce their μ's to a common denominator so as to express them as $X^{a_1/b}$, $X^{a_2/b}$, $X^{a_3/b}$. If $a_1 < a_2 < a_3$, so that $[X^{a_1}\ X^{a_2}\ X^{a_3}]$, we can apply 13.22 to the dilatation $1X^b \to 1X$, with the conclusion that

$$[X^{a_1/b}\ X^{a_2/b}\ X^{a_3/b}].$$

If μ is irrational, we define X^μ to be the Dedekind section between all the rational points $X^{a/b}$ for which $a/b < \mu$ and all those for which $a/b > \mu$. More precisely, supposing for definiteness that μ is positive, we apply the "ray" version of 12.51 to two sets of points, one consisting of all the points

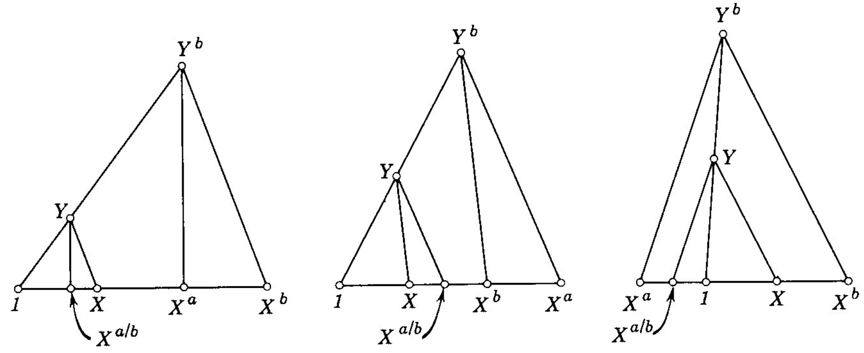

Figure 13.3b

whose exponents are positive rational numbers less than μ, and all the points between pairs of these, whereas the other set consists of the rest of the "positive" ray $1X$. (If μ is negative, we make the same kind of partition of the "negative" ray $1/X$.) Finally X^μ is, by definition, the translation $1 \to X^\mu$.

We have now interpreted the symbol X^μ for all real values of μ (including 0 and 1, which yield $X^0 = 1$ and $X^1 = X$). Conversely, *every point on the line $1X$ can be expressed in the form X^μ.*

This is obvious for any point of the interval from X^{-1} to X. Any other point T satisfies either $[1\ X\ T]$ or $[1\ X^{-1}\ T]$. If $[1\ X\ T]$, the dilatation $1T \to 1X$ transforms X into a point between 1 and X, say X^λ. The inverse dilatation $1X^\lambda \to 1X$ transforms X into $X^{1/\lambda}$; therefore $T = X^{1/\lambda}$. If, on the other hand, $[1\ X^{-1}\ T]$, we make the analogous use of $1T \to 1X^{-1}$. In either case we obtain an expression for T as a power of X.

Thus, assuming Dedekind's axiom, we have proved the "axiom of Archimedes":

13.31 *For any point T (except 1) on the line of a translation* X, *there is an integer n such that T lies between the points 1 and X^n.*

The exponent μ provides a measure of distance along the line $1X$. In fact, the segment $X^\nu X^\mu$ $(\nu < \mu)$ is said to have *length* $\mu - \nu$ in terms of the segment $1X$ as unit:

$$\frac{X^\nu X^\mu}{1X} = \mu - \nu.$$

Along another line $1Y$ (Figure 13.3*c*) we have an independent unit. Since the dilatation $1X \to 1X^\mu$ transforms the point Y into Y^μ, where the line $X^\mu Y^\mu$ is parallel to XY, we have

$$\frac{1X^\mu}{1X} = \frac{1Y^\mu}{1Y}$$

in agreement with Euclid VI.2 (see § 1.3). Thus we can define ratios of the lengths on one line, or on parallel lines, and we can compare such ratios on different lines. But affine geometry contains no machinery for comparing lengths in different directions: it is a meaningless question whether the translation Y is longer or shorter than X.

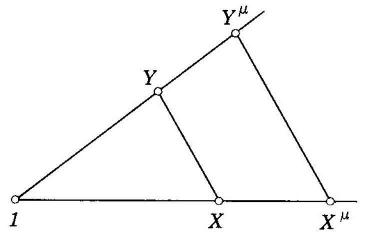

Figure 13.3c

The above definition for the length of the segment $X^\nu X^\mu$ ($\nu < \mu$) suggests the propriety of allowing the oppositely directed segment $X^\mu X^\nu$ to have the negative length $\nu - \mu$. This convention enables us to write $\mu = 1X^\mu/1X$ for negative as well as positive values of μ, and to add lengths of collinear segments according to such formulas as

$$AB + BC = AC, \qquad BC + CA + AB = 0,$$

regardless of the order of their end points A, B, C.

Now, to set up a system of *affine coordinates* in the plane, we let (x, y) denote the point into which the origin 1 is transformed by the translation $\mathrm{X}^x\mathrm{Y}^y$. This simple device establishes a one-to-one correspondence between points in the plane and ordered pairs of real numbers. In particular, the point X^x is $(x, 0)$, Y^y is $(0, y)$, and the origin itself is $(0, 0)$. When x and y are integers, the points (x, y) form a *two-dimensional lattice,* as in Figure 4.1*b*. The remaining points (x, y) are distributed between the lattice points in the obvious manner.

In affine coordinates (as in Cartesian coordinates) a line has a linear equation. The powers of the translation $\mathrm{X}^{-b}\,\mathrm{Y}^a$ transform the origin into the points $(-\mu b, \mu a)$ whose locus is the line $ax + by = 0$. The same powers transform (x_1, y_1) into the points

$$(x_1 - \mu b,\ y_1 + \mu a)$$

whose locus is

$$a(x - x_1) + b(y - y_1) = 0.$$

We can thus express a line in any of the standard forms 8.11, 8.12, 8.13.

A dilatation is a special case of an *affinity,* which is any transformation (of the whole affine plane onto itself) preserving collinearity. Thus, an affinity transforms parallel lines into parallel lines, and preserves ratios of distances along parallel lines. It also preserves intermediacy (compare 13.22).

13.32 *An affinity is uniquely determined by its effect on any one triangle.*

For, if it transforms a triangle $1XY$ into $1'X'Y'$, it transforms the point (x, y) referred to the former triangle into the point having the *same* coordinates referred to the latter. Here $1XY$ and $1'X'Y'$ may be *any* two triangles [Veblen and Young **2,** p. 72], and we naturally speak of "the affinity $1XY \to 1'X'Y'$." In particular, if ABC and ABC' are two triangles with a common side, $ABC \to ABC'$ is called a *shear* or a *strain* according as the line CC' is or is not parallel to AB. One kind of strain is sufficiently important to deserve a special name and a special symbol: the *affine reflection* $A(CC')$ or $B(CC')$, which arises when the midpoint of CC' lies on AB. In other words, any triangle ACC' determines an affine reflection $A(CC')$ whose *mirror* (or "axis") is the median through A and whose *direction* is the direction of all lines parallel to CC'.

In the language of the Erlangen program (see page 67), the principal group for affine geometry is the group of all affinities.

EXERCISES

1. The shear or strain $ABC \to ABC'$ leaves invariant every point on the line AB. What is its effect on a point P of general position?
2. Every affinity of period 2 is either a half-turn or an affine reflection.
3. If, for a given affinity, every noninvariant point lies on at least one invariant line, then the affinity is either a dilatation or a shear or a strain.
4. In terms of affine coordinates, affinities are "affine transformations"

13.33
$$x' = ax + by + l, \qquad y' = cx + dy + m, \qquad ad \neq bc.$$

5. Describe the transformations

(i) $x' = x + 1,\ y' = y$; (ii) $x' = ax,\ y' = ay$;
(iii) $x' = x + by,\ y' = y$; (iv) $x' = ax,\ y' = y$.

13.4 EQUIAFFINITIES

For he, by Geometrick scale,
Could take the size of Pots of Ale.

Samuel Butler (1600-1680)
(*Hudibras,* I.1)

We are now ready to show how the comparison of lengths on parallel lines can be extended to yield a comparison of areas in any position [cf. Forder **1,** pp. 259–265; Coxeter **2,** pp. 125–128]. For simplicity, we restrict consideration to *polygonal* regions. (Other shapes may be included by a suitable limiting process of the kind used in integral calculus.) Clearly, any polygonal region can be dissected into a finite number of triangles.* Following H. Hadwiger and P. Glur [*Elemente der Mathematik,* **6** (1951), pp. 97–120], we declare two such regions to be *equivalent* if they can be dissected into a finite number of pieces that are congruent in pairs (by translations or by half-turns). In other words, two polygonal regions are equivalent if they can be derived from each other by dissection and rearrangement. Superposing two different dissections, we see that this kind of equivalence, which is obviously reflexive and symmetric, is also transitive; two polygons that are equivalent to the same polygon are equivalent to each other.

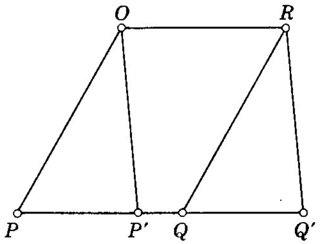

Figure 13.4a

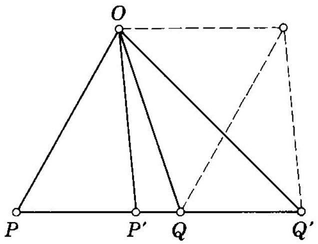

Figure 13.4b

The parallelograms $OPQR$ and $OP'Q'R$ of Figure 13.4*a* are equivalent, since each of them consists of the trapezoid $OP'QR$ plus one of the two congruent triangles OPP', RQQ'. In some such cases, more than two pieces may be needed, but we find eventually:

13.41 *Two parallelograms are equivalent if they have one pair of opposite sides of the same length lying on the same pair of parallel lines.*

Since a parallelogram can be dissected along a diagonal to make two triangles that are congruent by a half-turn, it follows that two triangles (such as OPQ and $OP'Q'$ in Figure 13.4*b*) are equivalent if they have a common vertex while their sides opposite to this vertex are congruent segments on one line. In particular, if points $P_0, P_1, \ldots, P_n$ are evenly spaced along a line (not through O), so that the segments P_0P_1, $P_1P_2, \ldots$ are all congruent, as in Figure 13.4*c*, then the triangles OP_0P_1, $OP_1P_2, \ldots$ are all equivalent, and we naturally say that the *area* of OP_0P_n is n times the area of OP_0P_1. By interpolation of further points on the same line, we can extend this idea to all real values of n, with the conclusion that, if Q is on the side PQ' of a triangle OPQ', as in Figure 13.4*d*, the *Cevian* OQ divides the area of the triangle in the same ratio that the point Q divides the side:

13.42
$$\frac{OPQ}{OPQ'} = \frac{PQ}{PQ'}.$$

* N. J. Lennes, *American Journal of Mathematics,* **33** (1911), p. 46.

We naturally regard this ratio as being negative if P lies between Q and Q', that is, if the two triangles are oppositely oriented.

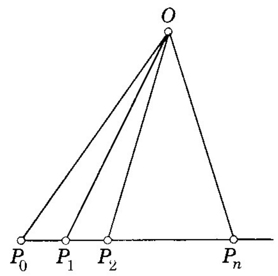

Figure 13.4c

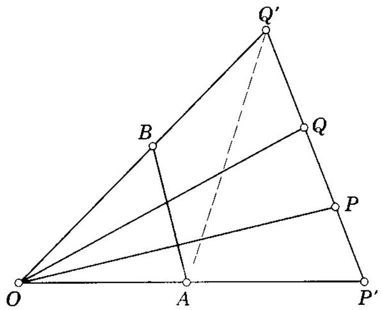

Figure 13.4d

These ideas enable us to define the area of any polygon in such a way that *equivalent polygons have the same area,* and when two polygons are stuck together to make a larger polygon, the areas are added. To compute the area of a given polygon in terms of a standard triangle OAB as unit of measurement, we dissect the polygon into triangles and add the areas of the pieces, each computed as follows.

By applying a suitable translation, any given triangle can be shifted so that one vertex coincides with the vertex O of the standard triangle OAB. Accordingly, we consider a triangle OPQ. Let the line PQ meet OA in P', and OB in Q', as in Figure 13.4*d*. Multiplying together the three ratios

$$\frac{OPQ}{OP'Q'} = \frac{PQ}{P'Q'}, \quad \frac{OP'Q'}{OAQ'} = \frac{OP'}{OA}, \quad \frac{OAQ'}{OAB} = \frac{OQ'}{OB}$$

we obtain the desired ratio

13.43
$$\frac{OPQ}{OAB} = \frac{PQ}{P'Q'}\frac{OP'}{OA}\frac{OQ'}{OB}.$$

To obtain an analytic expression for the area of a triangle OPQ, referred to axes through the vertex O, we take the coordinates of the points

$$O, \quad A, \quad B, \quad P, \quad Q, \quad P', \quad Q'$$

to be

$$(0, 0), (1, 0), (0, 1), (x_1, y_1), (x_2, y_2), (p, 0), (0, q),$$

respectively. Since the equation

$$\frac{x}{p} + \frac{y}{q} = 1$$

for the line PQ is satisfied by (x_1, y_1) and (x_2, y_2), we have

$$\frac{1 - y_1/q}{x_1} = \frac{1}{p} = \frac{1 - y_2/q}{x_2},$$

whence $$q = \frac{x_1y_2 - x_2y_1}{x_1 - x_2}.$$

Taking the product of

$$\frac{PQ}{P'Q'} = \frac{x_1 - x_2}{p}, \quad \frac{OP'}{OA} = p, \quad \frac{OQ'}{OB} = \frac{x_1y_2 - x_2y_1}{x_1 - x_2},$$

we obtain

13.44 $$\frac{OPQ}{OAB} = x_1y_2 - x_2y_1 = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}.$$

We deduce, as in § 8.2, that a triangle

$$(x_1, y_1)\,(x_2, y_2)\,(x_3, y_3),$$

of general position, has area PQR, where

13.45 $$\frac{PQR}{OAB} = \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}.$$

Since the homogeneous linear transformation

$$x' = ax + by, \qquad y' = cx + dy$$

takes the triangle OAB to

$$(0, 0)(a, c)(b, d),$$

we conclude that the affinity 13.33 *preserves area* if and only if

$$ad - bc = 1.$$

An area-preserving affinity is called an *equiaffinity* (or "equiaffine collineation" [Veblen and Young **2,** pp. 105–113]). The group of all equiaffinities, like the group of all dilatations, includes the group of all translations and half-turns as a normal subgroup, and is itself a normal subgroup in the group of all affinities. Equiaffinities are of many kinds. Here are some examples:

The *hyperbolic rotation* ("Lorentz transformation" or "Procrustean stretch")

13.46 $$x' = \mu^{-1}x, \quad y' = \mu y \qquad (\mu > 0, \quad \mu \neq 1),$$

for which $x'y' = xy$, leaves invariant each branch of the hyperbola $xy = 1$. The crossed hyperbolic rotation

13.47 $$x' = -\mu^{-1}x, \quad y' = -\mu y \qquad (\mu > 0, \quad \mu \neq 1)$$

interchanges the two branches. The *parabolic rotation*

13.48 $$x' = x + 1, \quad y' = 2x + y + 1,$$

for which $x'^2 - y' = x^2 - y$, leaves invariant the parabola $y = x^2$. The *elliptic rotation*

13.49 $$x' = x\cos\theta - y\sin\theta, \quad y' = x\sin\theta + y\cos\theta$$

leaves invariant the ellipse $x^2 + y^2 = 1$, and is periodic if θ is commensurable with π.

In §2.8 (page 36) we derived a regular polygon $P_0P_1P_2\ldots$ from a point P_0 (other than the center) by repeated application of a rotation through $2\pi/n$. (The rotation takes P_0 to P_1, P_1 to P_2, and so on.) Although measurement of angles has no meaning in affine geometry, we can define an *affinely regular polygon* whose vertices P_j are derived from suitable point P_0 by repeated application of an equiaffinity. The polygon is said to be of type $\{n\}$ if the equiaffinity is an elliptic rotation 13.49, where $\theta = 2\pi/n$ and n is rational, so that P_j has affine coordinates

$$(\cos j\theta, \sin j\theta) \qquad (\theta = 2\pi/n).$$

Figure 13.4*e* shows an affinely regular pentagram ($n = 5/2$) and pentagon ($n = 5$).

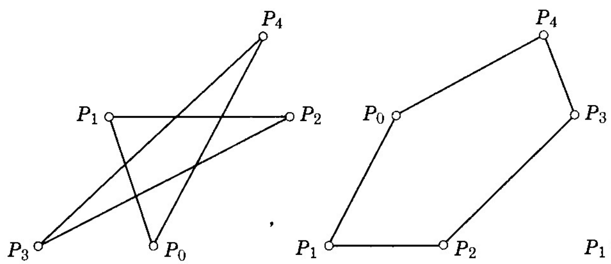

Figure 13.4e

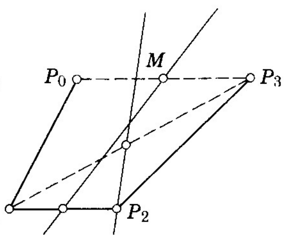

Figure 13.4f

EXERCISES

1. Two triangles with a common side (such as ABC and BCD in Figure 13.2*d*) have the same area if and only if the line joining their remaining vertices is parallel to the common side (that is, AD parallel to BC).
2. If a pentagon has four of its diagonals parallel to four of its sides, the remaining diagonal is parallel to the remaining side.
3. When is a dilatation an equiaffinity?
4. When is a shear an equiaffinity?
5. When is a strain an equiaffinity?
6. The product of any even number of affine reflections is an equiaffinity.

7. Any translation or half-turn or shear can be expressed as the product of two affine reflections.

8. If an equiaffinity is neither a translation nor a half-turn nor a shear, it can be expressed as $P_0P_1P_2 \to P_1P_2P_3$ where P_0P_3 is parallel to P_1P_2. (See Figure 13.4*f*.)

9. Every equiaffinity can be expressed as the product of two affine reflections. (Veblen.)

10. In an affinely regular polygon $P_0P_1P_2 \ldots$, the lines P_iP_j and P_hP_k are parallel whenever $i + j = h + k$.

11. Why did we call $x^2 + y^2 = 1$ an ellipse rather than a circle (just below 13.49)?

12. What triangles and quadrangles are affinely regular?

13. Construct an affinely regular hexagon.

14. Compute the ratio P_0P_3/P_1P_2 for an affinely regular polygon of type $\{n\}$.

15. For which values of n can an affinely regular polygon of type $\{n\}$ be constructed with a parallel-ruler?

13.5 TWO-DIMENSIONAL LATTICES

Farey has a notice of twenty lines in the Dictionary of National Biography. . . . His biographer does not mention the one thing in his life which survives.

G. H. Hardy

[Hardy and Wright **1**, p. 37]

Our treatment of lattices in § 4.1 (as far as the description of Figure 4.1*d*) is purely affine. In fact, a lattice is the set of points whose affine coordinates are integers. Any one of the points will serve as the origin O.

Let A' be any lattice point, and A the first lattice point along the ray OA'. Following Hardy and Wright [**1**, p. 29], we call A a *visible* point, because there is no lattice point between O and A to hide A from an observer at O. In terms of affine coordinates, a necessary and sufficient condition for (x, y) to be visible is that the integers x and y be coprime, that is, that they have no common divisor greater than 1. The three visible points

$$(1, 0), \quad (1, 1), \quad (0, 1)$$

form with the origin a parallelogram. This is called a *unit cell* (or "typical parallelogram") of the lattice, because the translations transform it into infinitely many such cells filling the plane without overlapping and without interstices: it is a fundamental region for the group of translations. Thus it serves as a convenient unit for computing the area of a region.

According to Steinhaus [**2**, pp. 76–77, 260] it was G. Pick, in 1899, who discovered the following theorem:*

* For an extension to three dimensions, see J. E. Reeve, On the volume of lattice polyhedra, *Proceedings of the London Mathematical Society* (3), **7** (1957), pp. 378–395.

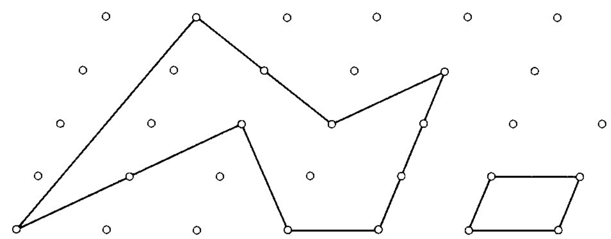

Figure 13.5a

13.51 *The area of any simple polygon whose vertices are lattice points is given by the formula*

$$\tfrac{1}{2}b + c - 1,$$

where b is the number of lattice points on the boundary while c is the number of lattice points inside.

(By a "simple" polygon we mean one whose sides do not cross one another. Figure 13.5*a* shows an example in which $b = 11$, $c = 3$.)

Proof. We first observe that the expression $\frac{1}{2}b + c - 1$ is additive when two polygons are juxtaposed. In fact, if two polygons, involving $b_1 + c_1$ and $b_2 + c_2$ lattice points respectively, have a common side containing n ($\geqslant 0$) lattice points in addition to the two vertices at its ends, then the values of b and c for the combined polygon are

$$b = b_1 + b_2 - 2n - 2, \qquad c = c_1 + c_2 + n,$$

so that

$$\tfrac{1}{2}b + c - 1 = (\tfrac{1}{2}b_1 + c_1 - 1) + (\tfrac{1}{2}b_2 + c_2 - 1).$$

Next, the formula holds for a parallelogram having no lattice points on its sides (so that $b = 4$ and the expression reduces to $c + 1$). For, when N such parallelograms are fitted together, four at each vertex, to fill a large region, the number of lattice points involved (apart from a negligible peripheral error) is $N(c + 1)$, and this must be the same as the number of unit cells needed to fill the same region.

Splitting the parallelogram into two congruent triangles by means of a diagonal, we see that the formula holds also for a triangle having no lattice points on its sides. A triangle that does have lattice points on a side can be dealt with by joining such points to the opposite vertex so as to split the triangle into smaller triangles. This procedure may have to be repeated, but obviously only a finite number of times. Finally, as we remarked on page 204, any given polygon can be dissected into triangles; then the expressions for those pieces can be added to give the desired result.

In particular, any parallelogram for which $b = 4$ and $c = 0$ has area 1 and can serve as a unit cell. If the vertices of such a parallelogram (in counterclockwise order) are

$$(0, 0),\quad (x, y),\quad (x + x_1, y + y_1),\quad (x_1, y_1),$$

we see from 13.44 that

13.52 $$xy_1 - yx_1 = 1.$$

In other words, this is the condition for the points

13.53 $$(0, 0),\quad (x, y),\quad (x_1, y_1)$$

to form a positively oriented "empty" triangle of area $\frac{1}{2}$, which could be used just as well as (0, 0) (1, 0) (0, 1) to generate the lattice. Thus a lattice is completely determined, apart from its position, by the area of its unit cell. Moreover, although there are infinitely many visible points in a given lattice, they all play the same role. (These properties of affine geometry are in marked contrast to Euclidean geometry, where the shape of a lattice admits unlimited variation and each lattice contains visible points at infinitely many different distances.)

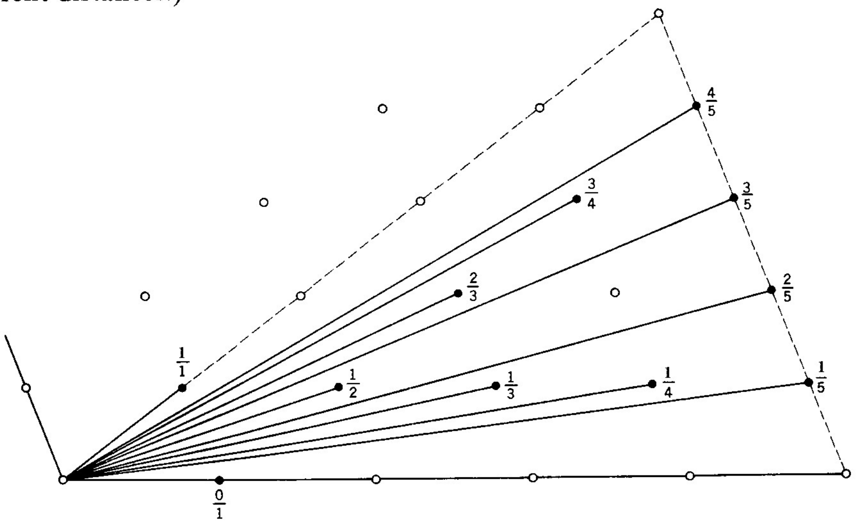

Figure 13.5b

George Pólya* has applied 13.52 to a useful lemma in the theory of numbers. The *Farey series* F_n of order n is the ascending sequence of fractions from 0 to 1 whose denominators do not exceed n. Thus y/x belongs to F_n if x and y are coprime and

13.54 $$0 \leqslant y \leqslant x \leqslant n.$$

For instance, F_5 is

$$\tfrac{0}{1}, \tfrac{1}{5}, \tfrac{1}{4}, \tfrac{1}{3}, \tfrac{2}{5}, \tfrac{1}{2}, \tfrac{3}{5}, \tfrac{2}{3}, \tfrac{3}{4}, \tfrac{4}{5}, \tfrac{1}{1}.$$

* *Acta Litterarum ac Scientiarum Regiae Universitatis Hungaricae Francisco-Josephinae*, Sectio

The essential property of such a sequence, from which many other properties follow by simple algebra, is that 13.52 holds for any two adjacent fractions

$$\frac{y}{x} \text{ and } \frac{y_1}{x_1}.$$

To prove this, we represent each term y/x of the sequence by the point (x, y) of a lattice. For example, the terms of F_5 are the lattice points emphasized in Figure 13.5*b* (where, for convenience, the angle between the axes is obtuse). Since the fractions are in their "lowest terms," the points are visible. By 13.54, they belong to the triangle (0, 0) $(n, 0)$ (n, n). A ray from the origin, rotated counterclockwise, passes through the representative points in their proper order. If y/x and y_1/x_1 are consecutive terms of the sequence, then (x, y) and (x_1, y_1) are visible points such that the triangle joining them to the origin contains no lattice point in its interior. Hence this triangle is one half of a unit cell, and 13.52 holds, as required.

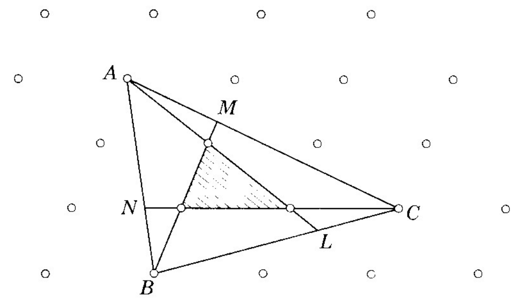

Figure 13.5c

Another result belonging to affine geometry is

13.55 *If the sides BC, CA, AB of a triangle ABC are divided at L, M, N in the respective ratios* $\lambda : 1$, $\mu : 1$, $\nu : 1$, *the Cevians AL, BM, CN form a triangle whose area is*

$$\frac{(\lambda\mu\nu - 1)^2}{(\lambda\mu + \lambda + 1)(\mu\nu + \mu + 1)(\nu\lambda + \nu + 1)}$$

times that of ABC.

This was discovered by Routh [**1**, p. 82; see also Dörrie **1**, pp. 41–42]. We shall give a general proof in § 13.7, but it is interesting to observe that, when $\lambda = \mu = \nu$, so that the ratio of areas is $(\lambda - 1)^3/(\lambda^3 - 1)$, the result can be deduced from 13.51. For instance, when $\lambda = \mu = \nu = 2$, so that each side is trisected [Steinhaus **2**, p. 8], the central triangle is one-seventh of the whole, and we can see this immediately by embedding the figure in a lattice, as in Figure 13.5*c*. Since the central triangle has $b = 3$, $c = 0$ while ABC has $b = 3$, $c = 3$, the ratio of areas is $\frac{1}{2}/\frac{7}{2} = \frac{1}{7}$.

EXERCISES

1. If y/x and y_1/x_1 are two consecutive terms of a Farey series, x and x_1 are coprime.

2. If y_0/x_0, y/x, y_1/x_1 are three consecutive terms of a Farey series,

$$\frac{y_0 + y_1}{x_0 + x_1} = \frac{y}{x}.$$

(C. Haros, 1802.)

3. The points A, B, C in Figure 13.5c belong to a lattice whose unit cell has seven times the area of that of the basic lattice. (For the Euclidean theory of such compound lattices, see Coxeter, Configurations and maps, *Reports of a Mathematical Colloquium* (2), **8** (1948), pp. 18–38, especially Figs. i, v, vii.)

4. Use lattices to verify 13.55 when (a) $\lambda = \mu = \nu = 3$, (b) $\lambda = \mu = \nu = \frac{3}{2}$.

5. Join the vertices A, B, C, D of a parallelogram to the midpoints of the respective sides BC, CD, DA, AB so as to form a smaller parallelogram in the middle. Its area is one-fifth that of $ABCD$. Another such parallelogram is obtained by joining A, B, C, D to the midpoints of CD, DA, AB, BC. The common part of these two small parallelograms is a centrally symmetrical octagon whose area is one-sixth that of $ABCD$ [Dörrie **1**, p. 40].

6. In the notation of 13.55, the area of the triangle LMN is

$$\frac{\lambda\mu\nu + 1}{(\lambda + 1)(\mu + 1)(\nu + 1)}$$

times that of ABC. (*Hint*: Use 13.42 to compute the relative area of CLM, etc.)

7. Of the four triangles ANM, BLN, CML, LMN, the last cannot have the smallest area unless L, M, N are the midpoints of BC, CA, AB. (H. Debrunner.*)

13.6 VECTORS AND CENTROIDS

A vector is really the same thing as a translation, although one uses different phraseologies for vectors and translations. Instead of speaking of the translation $A \to A'$ which carries the point A into A' one speaks of the vector $\overrightarrow{AA'}$. . . . The same vector laid off from B ends in B' if the translation carrying A into A' carries B into B'.

H. Weyl [**1**, p. 45]

As we saw in § 2.5, a *group* is an associative system containing an identity and, for each element, an inverse. Arithmetical instances are provided by the positive rational numbers, the positive real numbers, the complex numbers of modulus 1, and all the complex numbers except 0, combined, in each case, by ordinary multiplication. Such instances make it natural to adopt a multiplicative notation for all groups, so that the combination of S and T is ST, the inverse of S is S^{-1}, and the identity is 1. However, it is often convenient, especially in the case of Abelian (i.e., commutative)

* *Elemente der Mathematik*, **12** (1957), p. 43, Aufgabe 260.

groups, to use instead the additive notation, in which the combination of S and T is S + T, the inverse of S is −S, and the identity is 0. To see that this other notation has equally simple arithmetical instances, we merely have to consider in turn the integers, the rational numbers, the real numbers, and the complex numbers, combined, in each case, by ordinary addition.

The transition from a multiplicative group to the corresponding additive group is the foundation of the theory of logarithms [Infeld **1**, pp. 97–100].

When we go outside the domain of arithmetic, the choice between multiplication and addition is merely a matter of notation. In particular, the Abelian group of translations, which we have expressed as a multiplicative group, becomes the additive group of *vectors*.

In this notation, 13.21 asserts that any two points A and A' determine a unique vector $\overrightarrow{AA'}$ (going from A to A'), Figure 13.2b illustrates a situation in which

$$\overrightarrow{AA'} = \overrightarrow{CC'} = \overrightarrow{BB'},$$

13.23 asserts that

$$\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC},$$

and 3.23 asserts that, for any two vectors $\mathbf{a}$ and $\mathbf{b}$,

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}.$$

In the same spirit the "origin" will henceforth be called O instead of I, and the zero vector will be denoted by $\mathbf{0}$. The integral multiples of any nonzero vector proceed from the origin to the points of a one-dimensional lattice. Two vectors $\mathbf{e}$ and $\mathbf{f}$ are said to be *independent* if neither is a (real) multiple of the other, that is, if the only numbers that satisfy the vector equation

$$x\mathbf{e} + y\mathbf{f} = \mathbf{0}$$

are $x = 0$ and $y = 0$. Two such vectors (corresponding to the translations X and Y in Figure 4.1c) provide a basis for a system of affine coordinates: they enable us to define the coordinates of any point to be the coefficients in the expression

$$x\mathbf{e} + y\mathbf{f}$$

for the *position vector* which goes from the origin to the given point. In other words, with reference to a triangle OAB, the affine coordinates of a point P are the coefficients in the expression

$$\overrightarrow{OP} = x\,\overrightarrow{OA} + y\,\overrightarrow{OB}.$$

We shall find it useful to borrow from statics the notion of the centroid (or "center of gravity") of a set of "weighted" points, that is, of points to each of which a real number is attached in a special way. For convenience, we shall call these numbers masses, although, when some of them are negative, electric charges provide a more appropriate illustration.

Let masses $t_1, \ldots, t_k$ be assigned to k distinct points $A_1, \ldots, A_k$, let O be any point (possibly coincident with one of the A's), and consider the vector

$$t_1\, \overrightarrow{OA_1} + \ldots + t_k\, \overrightarrow{OA_k}.$$

If $t_1 + \ldots + t_k = 0$, this vector is independent of the choice of O. For, if we subtract from it the result of using O' instead, we obtain

$$t_1\,(\overrightarrow{OA_1} - \overrightarrow{O'A_1}) + \ldots + t_k\,(\overrightarrow{OA_k} - \overrightarrow{O'A_k})$$
$$= (t_1 + \ldots + t_k)\,\overrightarrow{OO'} = \mathbf{0}.$$

More interestingly, if

$$t_1 + \ldots + t_k \neq 0,$$

we have

$$t_1\,\overrightarrow{OA_1} + \ldots + t_k\,\overrightarrow{OA_k} = (t_1 + \ldots + t_k)\,\overrightarrow{OP},$$

where the point P is independent of the choice of O. For, if the same procedure with O' instead of O yields P' instead of P, we have, by subtraction,

$$(t_1 + \ldots + t_k)\,\overrightarrow{OO'} = (t_1 + \ldots + t_k)(\overrightarrow{OP} - \overrightarrow{O'P'})$$

whence $$\overrightarrow{OP'} = \overrightarrow{OO'} + \overrightarrow{O'P'} = \overrightarrow{OP},$$

so that P' coincides with P. This point P, given by

$$\Sigma t_i\,\overrightarrow{OP} = \Sigma t_i\,\overrightarrow{OA_i},$$

is called the *centroid* (or "barycenter") of the k masses t_i at A_i.

Since, having found P, we may choose this position for O, we have

$$\Sigma t_i\,\overrightarrow{PA_i} = \mathbf{0}.$$

If there are only two points,

$$t_1\,\overrightarrow{PA_1} = -t_2\,\overrightarrow{PA_2},$$

so that P lies on the line A_1A_2 and divides the segment A_1A_2 in the ratio $t_2 : t_1$. In particular, if $t_1 = t_2$, P is the midpoint of A_1A_2.

For a triangle $A_1A_2A_3$, we have

$$(t_1 + t_2 + t_3)\,\overrightarrow{OP} = t_1\overrightarrow{OA_1} + t_2\overrightarrow{OA_2} + t_3\overrightarrow{OA_3}$$
$$= t_1\overrightarrow{OA_1} + (t_2 + t_3)\,\overrightarrow{OQ},$$

where Q is the centroid of t_2 at A_2 and t_3 at A_3. Thus, in seeking the centroid of three masses, we may replace two of them by their combined mass at their own centroid. (There is an obvious generalization to more than three masses.) In particular, when $t_1 = t_2 = t_3$ (= 1, say), Q is the midpoint of A_2A_3, and P divides A_1Q in the ratio 2 : 1. Thus the "centroid" G of a triangle (§ 1.4) is the centroid of equal masses at its three vertices.

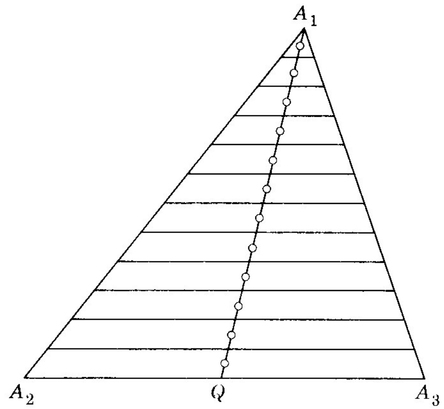

Figure 13.6a

This same point G, where the medians concur, is also the centroid of a triangular *lamina* or "plate" of uniform density. (Strictly speaking, this notion requires integral calculus.) For we may divide the triangle into thin strips parallel to the side A_2A_3, as in Figure 13.6*a*. The centroids of these strips evidently lie on the median A_1Q. Hence the centroid of the whole lamina lies on this median, and similarly on the others. (This argument was used by Archimedes in the third century B.C.)

EXERCISES

1. Verify in detail that
   (i) the positive rational numbers,
   (ii) the positive real numbers,
   (iii) the complex numbers of modulus 1,
   (iv) all the complex numbers except 0

form multiplicative groups; and that
   (v) the integers,
   (vi) the rational numbers,
   (vii) the real numbers,
   (viii) the complex numbers

form additive groups. Explain why the first four sets do not form additive groups, and why the last four do not form multiplicative groups.

2. If A, B, C are on one line and A', B', C' on another with

$$\frac{AB}{A'B'} = \frac{BC}{B'C'},$$

then points dividing all the segments AA', BB', CC' in the same ratio are either collinear or coincident (cf. § 3.6). (*Hint:* Consider the centroid of suitable masses at A, C, A', C'.)

3. The centroid of equal masses at the vertices of a quadrangle is the center of the Varignon parallelogram (Figure 13.2*g*).

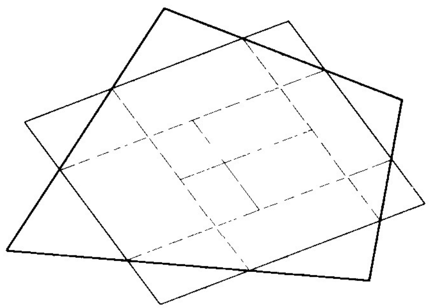

Figure 13.6b

4. The centroid of a quadrangular lamina is the center of the Wittenbauer parallelogram, whose sides join adjacent points of trisection of the sides, as in Figure 13.6*b*. This theorem, due to F. Wittenbauer (1857–1922) [Blaschke **2**, p. 13], was rediscovered by J. J. Welch and V. W. Foss.*

5. For what kind of quadrangle will the centroids described in the two preceding exercises coincide?

13.7 BARYCENTRIC COORDINATES

If $t_1 + t_2 \neq 0$, masses t_1 and t_2 at two fixed points A_1 and A_2 determine a unique centroid P, as in Figure 13.7*a*. This point is A_1 itself if $t_2 = 0$, A_2 if $t_1 = 0$. It is on the segment A_1A_2 if the t's are both positive (or both negative), on the ray A_1/A_2 if

$$t_1 > -t_2 > 0,$$

and on the ray A_2/A_1 if $\quad t_2 > -t_1 > 0.$

* *Mathematical Gazette*, **42** (1958), p. 55; **43** (1959), p. 46.

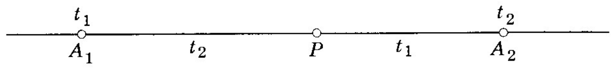

Figure 13.7a

Conversely, given a point P on the line A_1A_2, we can find numbers t_1 and t_2 such that

$$\frac{t_2}{t_1} = \frac{A_1P}{PA_2} \quad \text{or} \quad \frac{t_1}{t_2} = \frac{PA_2}{A_1P};$$

then P will be the centroid of masses t_1 and t_2 at A_1 and A_2. Since masses μt_1 and μt_2 (where $\mu \neq 0$) determine the same point as t_1 and t_2, these *barycentric coordinates* are homogeneous:

$$(t_1, t_2) = (\mu t_1, \mu t_2) \qquad (\mu \neq 0).$$

Similarly, as Möbius observed in 1827, we may set up barycentric coordinates in the plane of a *triangle of reference* $A_1A_2A_3$. If $t_1 + t_2 + t_3 \neq 0$, masses t_1, t_2, t_3 at the three vertices determine a point P (the centroid) whose coordinates are (t_1, t_2, t_3). In particular, $(1, 0, 0)$ is A_1, $(0, 1, 0)$ is A_2, $(0, 0, 1)$ is A_3, and $(0, t_2, t_3)$ is the point on A_2A_3 whose one-dimensional coordinates with respect to A_2 and A_3 are (t_2, t_3). To find coordinates for a given point P of general position, we find t_2 and t_3 from such a point Q on the line A_1P, as in Figure 13.7*b*, and then determine t_1 as the mass at A_1 that will balance a mass $t_2 + t_3$ at Q so as to make P the centroid. Again, as in the one-dimensional case, these coordinates are homogeneous:

$$(t_1, t_2, t_3) = (\mu t_1, \mu t_2, \mu t_3) \qquad (\mu \neq 0).$$

Joining P to A_1, A_2, A_3, we decompose $A_1A_2A_3$ into three triangles having a common vertex P. *The areas of these triangles are proportional to the barycentric coordinates of P*, as in Figure 13.7*c*. This fact follows at once from 13.42, since

$$\frac{t_3}{t_2} = \frac{A_2Q}{QA_3} = \frac{A_1A_2Q}{A_1QA_3} = \frac{PA_2Q}{PQA_3} = \frac{A_1A_2Q - PA_2Q}{A_1QA_3 - PQA_3} = \frac{PA_1A_2}{PA_3A_1},$$

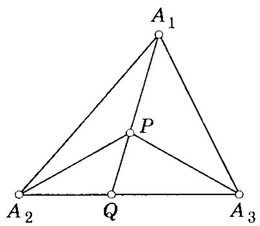

Figure 13.7b

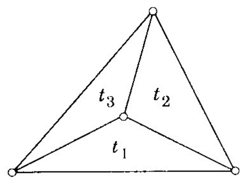

Figure 13.7c

and similarly for t_1/t_3, t_2/t_1. Positions of P outside the triangle are covered by means of our convention for the sign of the area of a directed triangle.

The inequality

$$t_1 + t_2 + t_3 \neq 0$$

enables us to normalize the coordinates so that

13.71 $$t_1 + t_2 + t_3 = 1.$$

(We merely have to divide each coordinate by the sum of all three.) These normalized barycentric coordinates are called *areal* coordinates, because they are just the areas of the triangles PA_2A_3, PA_3A_1, PA_1A_2, expressed in terms of the area of the whole triangle $A_1A_2A_3$ as unit of measurement. Areal coordinates are not homogeneous but "redundant": the position of a point is determined by two of the three, and the third is retained for the sake of symmetry. However, any expression involving them can be made homogeneous by inserting suitable powers of $t_1 + t_2 + t_3$ in appropriate places.

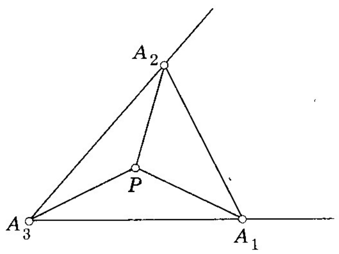

Figure 13.7d

In affine coordinates, as we have seen, a line has a linear equation. In barycentric coordinates, as we shall soon see, *a line has a linear homogeneous equation.* For this purpose we use the segments A_3A_1 and A_3A_2 as axes for affine coordinates, as in Figure 13.7*d*, so that the coordinates of P, A_1, A_2, A_3, which were formerly

$$(t_1, t_2, t_3),\quad (1, 0, 0),\quad (0, 1, 0),\quad (0, 0, 1),$$

are now

$$(x, y),\qquad (1, 0),\qquad (0, 1),\qquad (0, 0).$$

By 13.44, the areas of PA_2A_3 and PA_3A_1, as fractions of the "unit" triangle $A_1A_2A_3$, are just

$$\begin{vmatrix} x & y \\ 0 & 1 \end{vmatrix} = x \quad\text{and}\quad \begin{vmatrix} 1 & 0 \\ x & y \end{vmatrix} = y.$$

By subtraction, the area of PA_1A_2 is $1 - x - y$. Hence the *areal* coordinates of P are related to the affine coordinates by the very simple formulas

$$t_1 = x,\quad t_2 = y,\quad t_3 = 1 - x - y.$$

The general line, having the affine equation 8.11, has the areal equation

$$at_1 + bt_2 + c = 0.$$

Making this homogeneous by the insertion of $t_1 + t_2 + t_3$, we deduce the barycentric equation

$$at_1 + bt_2 + c(t_1 + t_2 + t_3) = 0$$

or $$(a + c)t_1 + (b + c)t_2 + ct_3 = 0$$

or, in a more symmetrical notation,

13.72 $$T_1t_1 + T_2t_2 + T_3t_3 = 0.$$

Thus every line has a linear homogeneous equation. In particular, the lines A_2A_3, A_3A_1, A_1A_2 have the equations

13.73 $$t_1 = 0,\quad t_2 = 0,\quad t_3 = 0.$$

The line joining two given points (r) and (s), meaning

$$(r_1, r_2, r_3) \quad\text{and}\quad (s_1, s_2, s_3),$$

has the equation

13.74 $$\begin{vmatrix} r_1 & r_2 & r_3 \\ s_1 & s_2 & s_3 \\ t_1 & t_2 & t_3 \end{vmatrix} = 0.$$

For, this equation is linear in the t's and is satisfied when the t's are replaced by the r's or the s's. Another way to obtain this result is to ask for the fixed points (r) and (s) to form with the variable point (t) a "triangle" whose area is zero. In terms of areal coordinates, with the triangle of reference as unit, the area of the triangle $(r)(s)(t)$ is, by 13.45 and 13.71,

$$\begin{vmatrix} r_1 & r_2 & 1 \\ s_1 & s_2 & 1 \\ t_1 & t_2 & 1 \end{vmatrix} = \begin{vmatrix} r_1 & r_2 & r_1 + r_2 + r_3 \\ s_1 & s_2 & s_1 + s_2 + s_3 \\ t_1 & t_2 & t_1 + t_2 + t_3 \end{vmatrix} = \begin{vmatrix} r_1 & r_2 & r_3 \\ s_1 & s_2 & s_3 \\ t_1 & t_2 & t_3 \end{vmatrix}.$$

Hence the area in general barycentric coordinates is this last determinant divided by

$$(r_1 + r_2 + r_3)(s_1 + s_2 + s_3)(t_1 + t_2 + t_3).$$

We are now ready to prove Routh's theorem 13.55 in its full generality. Identifying ABC with $A_1A_2A_3$, so that the points L, M, N are

$$(0, 1, \lambda),\quad (\mu, 0, 1),\quad (1, \nu, 0),$$

we can express the lines AL, BM, CN as

$$\lambda t_2 = t_3,\quad \mu t_3 = t_1,\quad \nu t_1 = t_2.$$

They intersect in pairs in the three points

$$(\mu, \mu\nu, 1), \quad (1, \nu, \nu\lambda), \quad (\lambda\mu, 1, \lambda),$$

forming a triangle whose area, in terms of that of the triangle of reference, is the result of dividing the determinant

$$\begin{vmatrix} \mu & \mu\nu & 1 \\ 1 & \nu & \nu\lambda \\ \lambda\mu & 1 & \lambda \end{vmatrix} = (\lambda\mu\nu - 1)^2$$

by $(\mu + \mu\nu + 1)(1 + \nu + \nu\lambda)(\lambda\mu + 1 + \lambda)$, in agreement with the statement of 13.55.

As an important special case we have

CEVA'S THEOREM. *Let the sides of a triangle ABC be divided at L, M, N in the respective ratios $\lambda : 1$, $\mu : 1$, $\nu : 1$. Then the three lines AL, BM, CN are concurrent if and only if $\lambda\mu\nu = 1$.*

The general line 13.72 meets the sides 13.73 of the triangle of reference in the points

$$(0, T_3, -T_2), \quad (-T_3, 0, T_1), \quad (T_2, -T_1, 0),$$

which divide them in the ratios

$$-\frac{T_2}{T_3}, \quad -\frac{T_3}{T_1}, \quad -\frac{T_1}{T_2},$$

whose product is -1. Conversely, any three numbers whose product is -1 can be expressed in this way for suitable values of T_1, T_2, T_3. Hence

MENELAUS'S THEOREM. *Let the sides of a triangle be divided at L, M, N in the respective ratios $\lambda : 1$, $\mu : 1$, $\nu : 1$. Then the three points L, M, N are collinear if and only if $\lambda\mu\nu = -1$.*

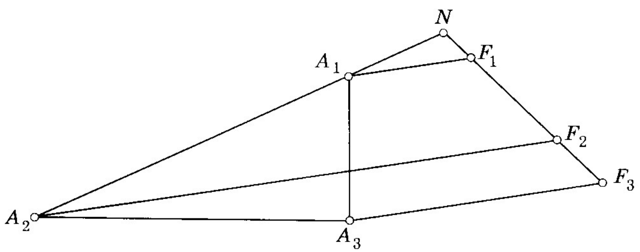

Figure 13.7e

The coefficients T_1, T_2, T_3 in the equation 13.72 for a line are sometimes called the *tangential coordinates* of the line. These homogeneous "coordinates" have a simple geometric interpretation [Salmon **1**, p. 11]: they may be regarded as *the distances from A_1, A_2, A_3 to the line,* measured in any direction (the same for all). To prove this, let A_1F_1, A_2F_2, A_3F_3 be these distances, as in Figure 13.7*e*. Since

$$\frac{A_1N}{NA_2} = -\frac{T_1}{T_2},$$

the homothetic triangles NA_1F_1 and NA_2F_2 yield

$$\frac{A_1F_1}{A_2F_2} = \frac{A_1N}{A_2N} = \frac{T_1}{T_2}.$$

Hence
$$\frac{A_1F_1}{T_1} = \frac{A_2F_2}{T_2},$$

and similarly each of these expressions is equal to $\dfrac{A_3F_3}{T_3}$.

Möbius's invention of homogeneous coordinates was one of the most far-reaching ideas in the history of mathematics: comparable to Leibniz's invention of differentials, which enabled him to express the equation

$$\frac{d}{dx} f(x) = f'(x)$$

in the homogeneous form

$$df(x) = f'(x)\, dx$$

(for instance, $d \sin x = \cos x\, dx$).

EXERCISES

1. Sketch the seven regions into which the lines A_2A_3, A_3A_1, A_1A_2 decompose the plane, marking each according to the signs of the three areal coordinates.

2. Verify that 13.45 yields $1 - x - y$ as the area of the triangle PA_1A_2 in Figure 13.7*d*.

3. In areal coordinates, the midpoint of $(s_1, s_2, s_3)(t_1, t_2, t_3)$ is

$$\left(\frac{s_1 + t_1}{2}, \frac{s_2 + t_2}{2}, \frac{s_3 + t_3}{2}\right).$$

4. The centroid of masses σ and τ at points whose areal coordinates are (s_1, s_2, s_3) and (t_1, t_2, t_3) is the point whose barycentric coordinates are

$$(\sigma s_1 + \tau t_1, \quad \sigma s_2 + \tau t_2, \quad \sigma s_3 + \tau t_3).$$

5. In barycentric coordinates, any point on the line $(s)(t)$ may be expressed in the form

$$(\sigma s_1 + \tau t_1, \quad \sigma s_2 + \tau t_2, \quad \sigma s_3 + \tau t_3).$$

6. Apply barycentric coordinates to Ex. 6 at the end of § 13.5. What becomes of this result when L, M, N are collinear?

7. In what way do the signs of T_1, T_2, T_3 depend on the position of the line 13.72 in relation to the triangle of reference? When T_2 and T_3 are positive, describe the cases $T_2 < T_3$, $T_2 = T_3$, $T_2 > T_3$.

13.8 AFFINE SPACE

Give me something to construct and I shall become God for the time being, pushing aside all obstacles, winning all the hard knowledge I need for the construction . . . advancing Godlike to my goal!

J. L. Synge [**2**, p. 162]

Affine geometry can be extended from two dimensions to three by using Axioms 12.42 and 12.43 instead of 12.41. The total number of axioms is not really increased, as 13.12 now becomes a provable theorem [Forder **1**, pp. 155–157]. A line and a plane, or two planes, are said to be *parallel* if they have no common point (or if the line lies in the plane, or if the two planes coincide). Thus any plane that meets two parallel planes meets them in parallel lines; if two planes are parallel, any line in either plane is parallel to the other plane; if two lines are parallel, any plane through either line is parallel to the other line.

The existence of parallel planes is ensured by the following theorem (cf. Axiom 13.11):

13.81 *For any point A and any plane γ, not through A, there is just one plane through A parallel to γ.*

Proof. Let q and r be two intersecting lines in γ. Let q' and r' be the respectively parallel lines through A. Then the plane $q'r'$ is parallel to γ. For otherwise, by 12.431, the two planes would meet in a line l. Since q' and r' are parallel to γ, they cannot meet l. Thus q' and r' are two parallels to l through A, contradicting 13.11. This proves that $q'r'$ is parallel to γ. Moreover, $q'r'$ is the only plane through A parallel to γ. For, two such would meet in a line s' through A, and we could obtain a contradiction by considering their section by the plane As, where s is a line in γ not parallel to s'.

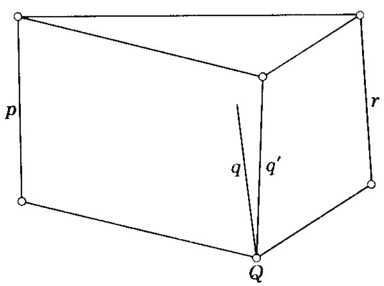

Figure 13.8a

Parallelism for lines is transitive in space as well as in a plane:

13.82 *If p and q are both parallel to r, they are parallel to each other.*

Proof [Forder **1**, p. 140]. When all three lines are in one plane, this follows at once from 13.11, so let us assume that they are not. For any point Q on q, the planes Qp and Qr meet in a line, say q' (Figure 13.8*a*). Any common point of q' and r would lie in both the planes Qp, pr, and therefore on their common line p; this is impossible, since p is parallel to r. Hence q' is parallel to r. But the only line through Q parallel to r is q. Hence q coincides with q', and is coplanar with p. Any common point of p and q would lie also on r. Hence p and q are parallel.

The transitivity of parallelism provides an alternative proof for 13.81. To establish the impossibility of a point O lying on both planes γ and $q'r'$, we imagine two lines through O, parallel to q (and q'), r (and r'). The planes γ and $q'r'$, each containing both these lines, would coincide, contradicting our assumption that A does not lie in γ.

The three face planes OBC, OCA, OAB of a tetrahedron $OABC$ form with the respectively parallel planes through A, B, C a *parallelepiped* whose faces are six parallelograms, as in Figure 13.8*b* [Forder **1**, p. 155].

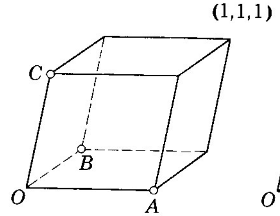

Figure 13.8b

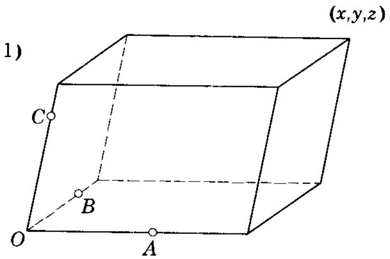

Figure 13.8c

It is now easy to build up a three-dimensional theory of dilatations, translations, and vectors. Three vectors **d**, **e**, **f** are said to be *dependent* if they are coplanar, in which case each is expressible as a linear combination of the other two. Three vectors **e**, **f**, **g** are said to be *independent* if the only solution of the vector equation

$$x\mathbf{e} + y\mathbf{f} + z\mathbf{g} = \mathbf{0}$$

is $x = y = z = 0$. Three such vectors provide a basis for a system of three-dimensional *affine coordinates*. In fact, if

$$\mathbf{e} = \overrightarrow{OA}, \quad \mathbf{f} = \overrightarrow{OB}, \quad \mathbf{g} = \overrightarrow{OC},$$

as in Figure 13.8*c*, the general vector $\overrightarrow{OP}$ may be exhibited as a diagonal of the parallelepiped formed by drawing through P three planes parallel to OBC, OCA, OAB. Then

$$\overrightarrow{OP} = x\mathbf{e} + y\mathbf{f} + z\mathbf{g},$$

where the terms of this sum are vectors along three edges of the parallelepiped.

In space, as in a plane, the centroid P of masses t_i at points A_i is determined by a vector $\overrightarrow{OP}$ such that

$$\Sigma t_i\, \overrightarrow{OP} = \Sigma t_i\, \overrightarrow{OA_i} \qquad (\Sigma t_i \neq 0).$$

If $\overrightarrow{OA_i} = x_i\mathbf{e} + y_i\mathbf{f} + z_i\mathbf{g}$, we deduce

$$\Sigma t_i\, \overrightarrow{OP} = \Sigma t_i x_i \mathbf{e} + \Sigma t_i y_i \mathbf{f} + \Sigma t_i z_i \mathbf{g}.$$

Hence, in terms of affine coordinates,

13.83 *The centroid of k masses t_i $(\Sigma t_i \neq 0)$ at points (x_i, y_i, z_i) $(i = 1, \ldots, k)$ is*

$$\left(\frac{\Sigma t_i x_i}{\Sigma t_i}, \frac{\Sigma t_i y_i}{\Sigma t_i}, \frac{\Sigma t_i z_i}{\Sigma t_i}\right).$$

In particular, if $t_1 + t_2 + t_3 = 1$, the centroid of three masses t_1, t_2, t_3 at the points

$$(1, 0, 0), \quad (0, 1, 0), \quad (0, 0, 1)$$

is (t_1, t_2, t_3). Hence

13.84 *The affine coordinates of any point in the plane $x + y + z = 1$ are the same as its areal coordinates referred to the triangle cut out from this plane by the coordinate planes $x = 0$, $y = 0$, $z = 0$.*

It follows that there is a line

$$\frac{x}{t_1} = \frac{y}{t_2} = \frac{z}{t_3}$$

through the origin (in affine space) for each point with barycentric coordinates (t_1, t_2, t_3). On the other hand, lines lying in the plane $x + y + z = 0$ yield no corresponding points in the parallel plane $x + y + z = 1$, unless we agree to extend the affine plane by postulating a line at infinity

$$t_1 + t_2 + t_3 = 0$$

so as to form the projective plane. This possibility has already been mentioned in § 6.9; we shall explore it more systematically in Chapter 14.

EXERCISES

1. If a line a is parallel to a plane α, and a plane through a meets α in b, then a and b are parallel lines. If another plane through a meets α in c, then b and c are parallel lines.

2. If α, β, γ are planes intersecting in lines $\beta \cdot \gamma = a$, $\gamma \cdot \alpha = b$, $\alpha \cdot \beta = c$, and a is parallel to b, then a, b, c are all parallel.

3. All the lines through A parallel to α are in a plane parallel to α [Forder **1**, p. 155].

4. Each of the six edges of a tetrahedron lies on a plane joining this edge to the midpoint of the opposite edge. The six planes so constructed all pass through one point: the centroid of equal masses at the four vertices.

5. Develop the theory of three-dimensional barycentric coordinates referred to a tetrahedron $A_1A_2A_3A_4$.

13.9 THREE-DIMENSIONAL LATTICES

The small parallelepiped built upon the three translations selected as unit translations . . . is known as the unit cell. . . . The entire crystal structure is generated through the periodic repetition, by the three unit translations, of the matter contained within the volume of the unit cell.

M. J. Buerger (1903 -   )
[Buerger **1**, p. 5]

The theory of volume in affine space is more difficult than that of area in the affine plane, because of the complication introduced by M. Dehn's observation that two polyhedra of equal volume are not necessarily derivable from each other by dissection and rearrangement. A valid treatment, suggested by Mrs. Sally Ruth Struik, may be described very briefly as follows. It is found that any two tetrahedra are related by a unique *affinity* $ABCD \rightarrow A'B'C'D'$, which transforms the whole space into itself in such a way as to preserve collinearity. In particular, a tetrahedron $ABCC'$ is transformed into $ABC'C$ by the *affine reflection*

$$AB(CC'),$$

which interchanges C and C' while leaving invariant every point in the plane that joins AB to the midpoint of CC'. Two tetrahedra are said to have the same *volume* if one can be transformed into the other by an *equiaffinity*: the product of an even number of affine reflections. Such a comparison is easily extended from tetrahedra to parallelepipeds, since a parallelepiped can be dissected into six tetrahedra all having the same volume.

In three dimensions, as in two, a *lattice* may be regarded as the set of points whose affine coordinates are integers. However, as it is independent of the chosen coordinate system, it is more symmetrically described as a discrete set of points whose set of position vectors is *closed under subtraction*, that is, along with any two of the vectors the set includes also their difference. Subtracting any one of the vectors from itself, we obtain the zero vector

$$\mathbf{c} - \mathbf{c} = \mathbf{0}$$

and hence also $\mathbf{0} - \mathbf{b} = -\mathbf{b}$, $\mathbf{a} - (-\mathbf{b}) = \mathbf{a} + \mathbf{b}$, $\mathbf{a} + \mathbf{a} = 2\mathbf{a}$, and so on: the set of vectors, containing the difference of any two, also contains the sum of any two, and all the integral multiples of any one. The lattice is one-, two-, or three-dimensional according to the number of independent vectors. In the three-dimensional case, a set of three independent vectors **e**, **f**, **g** is called a *basis* for the lattice if all the vectors are expressible in the form

13.91 $$x\mathbf{e} + y\mathbf{f} + z\mathbf{g},$$

where x, y, z are integers. If three of these vectors, say $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$ form another basis for the same lattice, there must exist 18 integers

$$a_\alpha, b_\alpha, c_\alpha, A_\alpha, B_\alpha, C_\alpha \qquad (\alpha = 1, 2, 3)$$

such that

$$\mathbf{r}_\alpha = a_\alpha \mathbf{e} + b_\alpha \mathbf{f} + c_\alpha \mathbf{g}, \quad \mathbf{e} = \Sigma A_\alpha \mathbf{r}_\alpha, \quad \mathbf{f} = \Sigma B_\alpha \mathbf{r}_\alpha, \quad \mathbf{g} = \Sigma C_\alpha \mathbf{r}_\alpha$$

and therefore

$$\mathbf{r}_\alpha = a_\alpha \Sigma A_\beta \mathbf{r}_\beta + b_\alpha \Sigma B_\beta \mathbf{r}_\beta + c_\alpha \Sigma C_\beta \mathbf{r}_\beta,$$

whence

$$a_\alpha A_\beta + b_\alpha B_\beta + c_\alpha C_\beta = \begin{cases} 1 & \text{if } \alpha = \beta, \\ 0 & \text{if } \alpha \neq \beta. \end{cases}$$

Since the product of two determinants is obtained by combining the rows of one with the columns of the other, we have

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \begin{vmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1.$$

Since the two determinants on the left are integers whose product is 1, each must be ± 1. Conversely, if a_α, b_α, c_α are given so that their determinant is ± 1, we can derive A_α, B_α, C_α by "inverting the matrix," and the given basis **e**, **f**, **g** yields the equally effective basis $\mathbf{r}_\alpha$. Hence

A necessary and sufficient condition for two triads of independent vectors

$$\mathbf{e}, \mathbf{f}, \mathbf{g} \quad \textit{and} \quad a_\alpha \mathbf{e} + b_\alpha \mathbf{f} + c_\alpha \mathbf{g} \qquad (\alpha = 1, 2, 3)$$

to be alternative bases for the same lattice is

13.92 $$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \pm 1$$

[cf. Hardy and Wright **1**, p. 28; Neville **1**, p. 5].

In other words, a lattice is derived from any one of its points by applying a *discrete group of translations:* one-, two-, or three-dimensional according as the translations are collinear, coplanar but not collinear, or not coplanar. In the one-dimensional case the generating translation is unique (except that it may be reversed), but in the other cases the two or three generators, that is, the basic vectors, may be chosen in infinitely many ways. When they have

been chosen, we can use them to set up a system of affine coordinates so that, in the three-dimensional case, the vector 13.91 goes from the origin (0, 0, 0) to the point (x, y, z), and the lattice consists of the points whose coordinates are integers. The eight points

$$(0,0,0), (1,0,0), (0,1,0), (0,0,1), (0,1,1), (1,0,1), (1,1,0), (1,1,1),$$

derived from the eight vectors

$$\mathbf{0}, \quad \mathbf{e}, \quad \mathbf{f}, \quad \mathbf{g}, \quad \mathbf{f}+\mathbf{g}, \quad \mathbf{g}+\mathbf{e}, \quad \mathbf{e}-\mathbf{f}, \quad \mathbf{e}+\mathbf{f}+\mathbf{g},$$

evidently form a parallelepiped, which is a *unit cell* of the lattice. By an argument analogous to that used for a two-dimensional lattice in § 4.1, *any two unit cells for the same lattice have the same volume.*

Any line joining two of the lattice points contains infinitely many of them, forming a one-dimensional sublattice of the three-dimensional lattice. In fact, the line joining (0, 0, 0) and (x, y, z) contains also (nx, ny, nz) for every integer n. If x, y, z have the greatest common divisor d, the lattice point

$$(x/d, y/d, z/d)$$

lies on this same line, and the corresponding translation generates the group of the one-dimensional lattice. The lattice point (x, y, z) is *visible* if and only if the three integers x, y, z have no common divisor greater than 1.

Any triangle of lattice points determines a plane containing a two-dimensional sublattice. For, if vectors

$$\mathbf{r}_1 = x_1\mathbf{e} + y_1\mathbf{f} + z_1\mathbf{g} \quad \text{and} \quad \mathbf{r}_2 = x_2\mathbf{e} + y_2\mathbf{f} + z_2\mathbf{g}$$

have integral components, so also does $t_1\mathbf{r}_1 + t_2\mathbf{r}_2$ for any integers t_1 and t_2. The parallel plane through any other lattice point will contain a congruent sublattice. Thus we may regard all the lattice points as being distributed among an infinite sequence of parallel planes, called *rational planes* [Buerger **1**, p. 7].

Any such plane, being the join of three points whose coordinates are integers, has an equation of the form

13.93 $$Xx + Yy + Zz = N,$$

where the coefficients X, Y, Z, N are integers, so that the intercepts on the coordinate axes have the rational values N/X, N/Y, N/Z. (This is the reason for the name "rational" planes.) We may assume that the greatest common divisor of X, Y, Z is 1; for, any common factor of X, Y, Z would be a factor of N too, and then we could divide both sides of the equation by this number, obtaining a simpler and equally effective equation for the same plane.

Conversely, any such equation (in which the greatest common divisor of X, Y, Z is 1) represents a plane containing a two-dimensional sublattice. This is obvious when $X = 1$, since then we can assign arbitrar integral values to y, z, and solve 13.93 for x. When X, Y, Z are all greater than 1, we consider the set of numbers

$$xX + yY + zZ,$$

where x, y, z are variable integers while X, Y, Z remain constant. This set (like the set of lattice vectors) is an *ideal:* it contains the difference of any two of its members and (therefore) all the multiples of any one. Let d denote its smallest positive member, and N any other member. Then N is a multiple of d: for otherwise we could divide N by d and obtain a remainder $N - qd$, which would be a member smaller than d. Thus every member of the set is a multiple of d. But X, Y, Z are members. Therefore d, being a common divisor, must be equal to 1, and the set simply consists of all the integers. In other words, the equation 13.93 has one integral solution (and therefore infinitely many) [cf. Uspensky and Heaslet **1**, p. 54].

For each triad of integers X, Y, Z, coprime in the above sense (but not necessarily coprime in pairs), we have a sequence of parallel planes 13.93, evenly spaced, one plane for each integer N. Since every lattice point lies in one of the planes, the infinite region between any two consecutive planes is completely empty. One of the planes, namely that for which $N = 0$, passes through the origin. The nearest others, given by $N = \pm 1$, are appropriately called *first* rational planes [Buerger **1**, p. 9]. We shall have occasion to consider them again in § 18.3.

EXERCISES

1. How can a parallelepiped be dissected into six tetrahedra all having the same volume?

2. Identify the transformation $(x, y, z) \to (x, y, -z)$ with the affine reflection that leaves invariant the plane $z = 0$ while interchanging the points $(0, 0, \pm 1)$.

3. A lattice is transformed into itself by the central inversion that interchanges two of its points.

4. Every lattice point in a first rational plane is visible.

5. Is every rational plane through a visible point a first rational plane?

6. Find a triangle of lattice points in the first rational plane

$$6x + 10y + 15z = 1.$$

7. Obtain a formula for all the lattice points in this plane.

8. The origin is the only lattice point in the plane

$$x + \sqrt{2}y + \sqrt{3}z = 0.$$

14 Projective geometry

In affine geometry, as we have seen, parallelism plays a leading role. In projective geometry, on the other hand, there is no parallelism: every pair of coplanar lines is a pair of intersecting lines. The conflict with 12.61 is explained by the fact that the projective plane is not an "ordered" plane. The set of points on a line, like the set of lines through a point, is closed: given three, we cannot pick out one as lying "between" the other two. At first sight we might expect a geometry having no circles, no distances, no angles, no intermediacy, and no parallelism, to be somewhat meagre. But, in fact, a very beautiful and intricate collection of propositions emerges: propositions of which Euclid never dreamed, because his interest in measurement led him in a different direction. A few of these nonmetrical propositions were discovered by Pappus of Alexandria in the fourth century A.D. Others are associated with the names of two Frenchmen: the architect Girard Desargues (1591–1661) and the philosopher Blaise Pascal (1623–1662). Meanwhile, the related subject of perspective [Yaglom **2**, p. 31] had been studied by artists such as Leonardo da Vinci (1452–1519) and Albrecht Dürer (1471–1528).

Kepler's invention of points at infinity made it possible to regard the projective plane as the affine plane plus the line at infinity. A converse relationship was suggested by Poncelet's *Traité des propriétés projectives des figures* (1822) and von Staudt's *Geometrie der Lage* (1847), in which projective geometry appeared as an independent science, making it possible to regard the affine plane as the projective plane minus an arbitrary line o, and then to regard the Euclidean plane as the affine plane with a special rule for associating pairs of points on o (in "perpendicular directions") [Coxeter **2**, pp. 115, 138]. This standpoint became still clearer in 1899, when Mario Pieri placed the subject on an axiomatic foundation. Other systems of axioms, slightly different from Pieri's, have been proposed by subsequent authors. The particular system that we shall give in § 14.1 was suggested by Bachmann [**1**, pp. 76–77]. To test the consistency of a system of axioms, we apply it to a "model," in which the primitive concepts are represented by familiar concepts whose properties we are prepared to accept [Coxeter **2**, pp.