INTRODUCTION TO GEOMETRY

By H. S. M. Coxeter

JOHN WILEY & SONS, Inc., Publishers

Introduction to Geometry

Introduction to Geometry

H.S.M. Coxeter

John Wiley and Sons, Inc. | New York • London • Sydney

ISBN: 978-0000000000
ISBN-10: 0000000000

Updated: 2024-12-19 12:29:36-06:002024-12-19

By the Same Author

PROJECTIVE GEOMETRY

Blaisdell, New York

THE REAL PROJECTIVE PLANE

Cambridge University Press

NON-EUCLIDEAN GEOMETRY

University of Toronto Press

REGULAR POLYTOPES

Methuen, London

H. S. M. COXETER, F.R.S. *Professor of Mathematics, University of Toronto*

John Wiley & Sons, Inc., New York • London • Sydney

ENCODED IN THE UNITED STATES OF AMERICA

Dedication

Preface

Contents

Plates

Epigraph

Part I.

I

1 Triangles

In this chapter, we review some of the well-known propositions of elementary geometry, stressing the role of symmetry. We refer to Euclid's propositions by his own numbers, which have been used throughout the world for more than two thousand years. Since the time of F. Commandino (1509–1575), who translated the works of Archimedes, Apollonius, and Pappus, many other theorems in the same spirit have been discovered. Such results were studied in great detail during the nineteenth century. As the present tendency is to abandon them in favor of other branches of mathematics, we shall be content to mention a few that seem particularly interesting.

1.1 Euclid

> Euclid's work will live long after all the text-books of the present day are superseded and forgotten. It is one of the noblest monuments of antiquity.
>
> Sir Thomas L. Heath (1861–1940)[1]

About 300 b.c., Euclid of Alexandria wrote a treatise in thirteen books called the *Elements*. Of the author (sometimes regrettably confused with the earlier philosopher, Euclid of Mégara) we know very little. Proclus (410–485 a.d.) said that he "put together the Elements, collecting many of Eudoxus's theorems, perfecting many of Theaetetus's, and also bringing to irrefragable demonstration the things which were only somewhat loosely proved by his predecessors." This man lived in the time of the first Ptolemy, [who] once asked him if there was in geometry any shorter way than that of the Elements, and he answered that "there was no royal road to geometry."

1 Heath **1**, p. vi. (Such references are collected at the end of the book, pp. 415–417.)

Heath quotes a story by Stobaeus, to the effect that someone who had begun to read geometry with Euclid asked him “What shall I get by learning these things?” Euclid called his slave and said “Give him a dime, since he must make gain out of what he learns.”

Of the thirteen books, the first six may be very briefly described as dealing respectively with triangles, rectangles, circles, polygons, proportion, and similarity. The next four, on the theory of numbers, include two notable achievements: IX.2 and X.9, where it is proved that there are infinitely many prime numbers, and that $\sqrt{2}$ is irrational [Hardy **2**, pp. 32–36]. Book XI is an introduction to solid geometry, XII deals with pyramids, cones, and cylinders, and XIII is on the five regular solids.

According to Proclus, Euclid “set before himself, as the end of the whole Elements, the construction of the so-called Platonic figures.” This notion of Euclid’s purpose is supported by the Platonic theory of a mystical correspondence between the four solids

$$\left\{\begin{array}{l}\text{cube,}\\ \text{tetrahedron,}\\ \text{octahedron,}\\ \text{icosahedron}\end{array}\right\} \text{and the four “elements”} \left\{\begin{array}{l}\text{earth,}\\ \text{fire,}\\ \text{air,}\\ \text{water}\end{array}\right\}$$

[cf. Coxeter **1**, p. 18]. Evidence to the contrary is supplied by the arithmetical books VII-X, which were obviously included for their intrinsic interest rather than for any application to solid geometry.

1.2 Primitive Concepts and Axioms

> ``When I use a word,'' Humpty-Dumpty said, ``it means just what I choose it to mean—neither more nor less.''
>
> Lewis Carroll (1832–1898)
>
> [Dodgson 2, Chap. 6]

In the logical development of any branch of mathematics, each definition of a concept or relation involves other concepts and relations. Therefore the only way to avoid a vicious circle is to allow certain *primitive* concepts and relations (usually as few as possible) to remain undefined [Synge 1, pp. 32–34]. Similarly, the proof of each proposition uses other propositions, and therefore certain primitive propositions, called *postulates* or *axioms*, must remain unproved. Euclid did not specify his primitive concepts and relations, but was content to give definitions in terms of ideas that would be familiar to everybody. His five Postulates are as follows:

1.2.1 1.21

A straight line may be drawn from any point to any other point.

1.2.2 1.22

A finite straight line may be extended continuously in a straight line.

1.2.3 1.23

A circle may be described with any center and any radius.

1.2.4 1.24

All right angles are equal to one another.

1.2.5 1.25

If a straight line meets two other straight lines so as to make the two interior angles on one side of it together less than two right angles, the other straight lines, if extended indefinitely, will meet on that side on which the angles are less than two right angles.[2]

It is quite natural that, after a lapse of about 2250 years, some details are now seen to be capable of improvement. (For instance, Euclid I.1 constructs an equilateral triangle by drawing two circles; but how do we know that these two circles will intersect?) The marvel is that so much of Euclid's work remains perfectly valid. In the modern treatment of his geometry [see, for instance, Coxeter 3, pp. 161–187], it is usual to recognize the primitive concept *point* and the two primitive relations of *intermediacy* (the idea that one point may be between two others) and *congruence* (the idea that the

2 In Chapter 15 we shall see how far we can go without using this unpleasantly complicated Fifth Postulate.

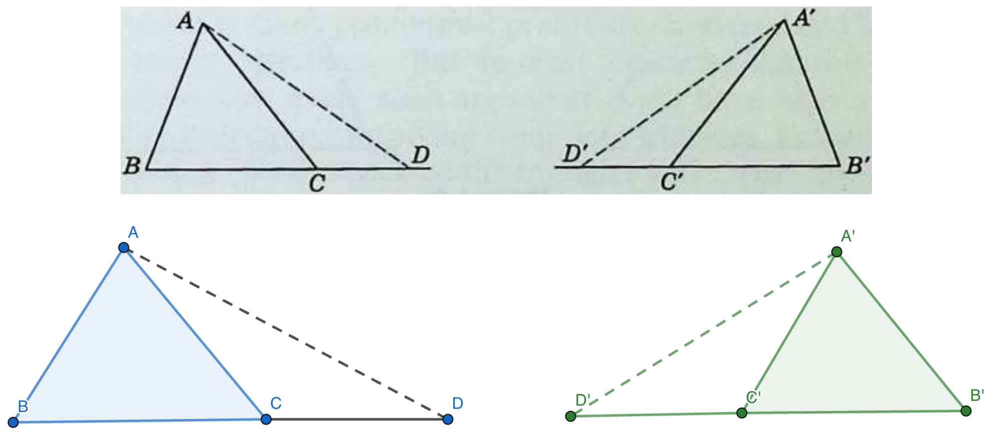


Figure 1.1.: 1.2a

distance between two points may be equal to the distance between two other points, or that two line segments may have the same length). There are also various versions of the axiom of *continuity*, one of which says that every convergent sequence of points has a limit.

Euclid's "principle of superposition, " used in proving I.4, raises the question whether a figure can be moved without changing its internal structure. This principle is nowadays replaced by a further explicit assumption such as the axiom of "the rigidity of a triangle with a tail" (Figure 1.2a 1.1):

1.2.6 1.26

If ABC is a triangle with D on the side BC extended, while D' is analogously related to another triangle $A'B'C'$, and if $BC = B'C'$, $CA = C'A'$, $AB = A'B'$, $BD = B'D'$, then $AD = A'D'$.

This axiom can be used to extend the notion of congruence from line segments to more complicated figures such as angles, so that we can say precisely what we mean by the relation

$$\angle ABC = \angle A'B'C'.$$

Then we no longer need the questionable principle of superposition in order to prove Euclid I.4:

If two triangles have two sides equal to two sides respectively, and have the angles contained by the equal sides equal, they will also have their third sides equal, and their remaining angles equal respectively; in fact, they will be congruent triangles.

1.3 Pons Asinorum

Minos: If is proposed to prove 1.5 by taking up the isosceles Triangle, turning it over, and then laying it down again upon itself.

Euclid: Surely that has too much of the Irish Bull about it, and reminds one a little too vividly of the man who walked down his own throat, to deserve a place in a strictly philosophical treatise?

Minos: I suppose its defenders would say that it is conceived to leave a trace of itself behind, and that the reversed Triangle is laid down upon the trace so left.

C. L. Dodgson (1832–1898)

[Dodgson 3, p. 48]

1.3.1 I.5.

The angles at the base of an isosceles triangle are equal.

The name *pons asinorum* for this famous theorem probably arose from the bridgelike appearance of Euclid's figure (with the construction lines required in his rather complicated proof) and from the notion that anyone unable to cross this bridge must be an ass. Fortunately, a far simpler proof was supplied by Pappus of Alexandria about 340a.d. (Figure 1.3a 1.2):

Let ABC be an isosceles triangle with AB equal to AC. Let us conceive this triangle as two triangles and argue in this way. Since $AB = AC and AC = AB$, the two sides AB, AC are equal to the two sides AC, AB. Also the angle BAC is equal to the angle CAB, for it is the same. Therefore all the corresponding parts (of the triangles ABC, ACB) are equal. In particular,

$$\angle ABC = \angle ACB.$$

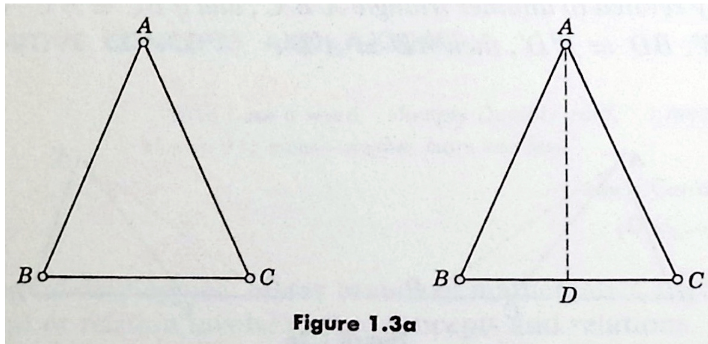


Figure 1.2.

The pedagogical difficulty of comparing the isosceles triangle *ABC* with itself is sometimes avoided by joining the apex A to D, the midpoint of the base BC. The median AD may be regarded as a *mirror* reflecting B into C. Accordingly, we say that an isosceles triangle is symmetrical by *reflection,* or that it has *bilateral symmetry.* (Of course, the idealized mirror used in geometry has no thickness and is silvered on both sides, so that it not only reflects B into C but also reflects C into B.)

Any figure, however irregular its shape may be, yields a symmetrical figure when we place it next to a mirror and waive the distinction between object and image. Such bilateral symmetry is characteristic of the external shape of most animals.

Given any point P on either side of a geometrical mirror, we can construct its reflected image P' by drawing the perpendicular from P to the mirror and extending this perpendicular line to an equal distance on the other side, so that the mirror perpendicularly bisects the line segment PP'. Working in the plane (Figure 1.3b 1.3) with a line AB for mirror, we draw two circles with centers A, B and radii AP, BP. The two points of intersection of these circles are P and its image P'.

We shall find that many geometrical proofs are shortened and made more vivid by the use of reflections. But we must remember that this procedure is merely a short cut: every such argument could have been avoided by means of a circumlocution involving congruent triangles. For instance, the above construction is valid because the triangles ABP, ABP' are congruent.

Pons asinorum has many useful consequences, such as the following five:

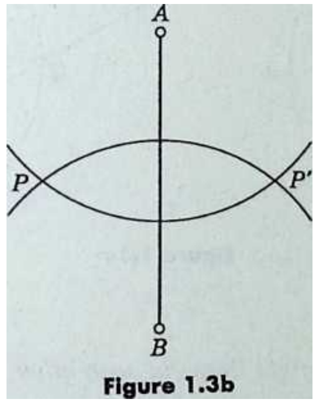


Figure 1.3.

1.3.2 III.3.

If a diameter of a circle bisects a chord which does not pass through the center, it is perpendicular to it; or, if perpendicular to it, it bisects it.

1.3.3 III.20.

In a circle the angle at the center is double the angle at the circumference, when the rays forming the angles meet the circumference in the same two points.

1.3.4 III.21.

In a circle, a chord subtends equal angles at any two points on the same one of the two arcs determined by the chord (e.g., in Figure 1.3c, $\angle PQQ' = \angle PP'Q'$*).*

1.3.5 III.22.

The opposite angles of any quadrangle inscribed in a circle are together equal to two right angles.

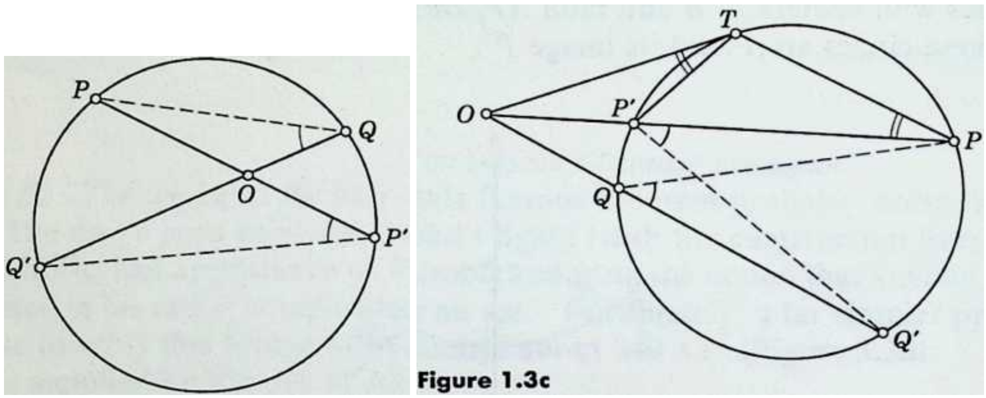


Figure 1.4.: 1.3c

1.3.6 III.32.

If a chord of a circle be drawn from the point of contact of a tangent, the angle made by the chord with the tangent is equal to the angle subtended by the chord at a point on that part of the circumference which lies on the far side of the chord (e.g., in Figure 1.3c, $\angle OTP' = \angle TPP'$).

We shall also have occasion to use two familiar theorems on similar triangles:

1.3.7 VI.2.

If a straight line be drawn parallel to one side of a triangle, it will cut the other sides proportionately; and, if two sides of the triangle be cut proportionately, the line joining the points of section will be parallel to the remaining side.

1.3.8 VI.4.

If corresponding angles of two triangles are equal, then corresponding sides are proportional.

Combining this last result with III.21 and 32, we deduce two significant properties of secants of a circle (Figure 1.3c 1.4):

1.3.9 III.35

If in a circle two straight lines cut each other, the rectangle contained by the segments of the one is equal to the rectangle contained by the segments of the other (i.e., $OP \times OP' = OQ \times OQ'$.)

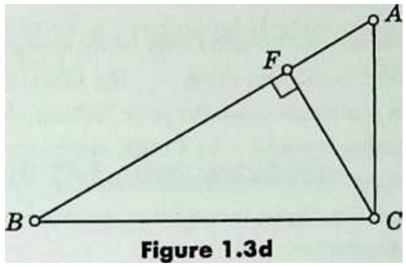


Figure 1.5.: 1.3d

1.3.10 III.6

If from a point outside a circle a secant and a tangent be drawn, the rectangle contained by the whole secant and the part outside the circle will be equal to the square on the tangent (i.e., $OP \times OP' = OT^2$.)

Book VI also contains an important property of area:

1.3.11 VI.19.

Similar triangles are to one another in the squared ratio of their corresponding sides (i.e., if ABC and $A'B'C'$ are similar triangles, their areas are in the ratio $AB^2 : A'B'^2$.)

This result yields the following easy proof for the theorem of Pythagoras [see Heath **1**, p. 353; **2**, p. 270]:

1.3.12 I.47.

In a right-angled triangle, the square on the hypotenuse is equal to the sum of the squares on the two catheti.

In the triangle ABC, right-angled at C, draw CF perpendicular to the hypotenuse AB, as in Figure 1.3d1.5. Then we have three similar right-angled triangles ABC, ACF, CBF, with hypotenuses AB, AC, CB. By VI.19, the areas satisfy

$$\frac{ABC}{AB^2} = \frac{ACF}{AC^2} = \frac{CBF}{CB^2}$$

Evidently, $ABC = ACF + CBF$. Therefore $AB^2 = AC^2 + CB^2$.

Exercises

1. Using rectangular Cartesian coordinates, show that the reflection in the $y-axis(x=0)$ reverses the sign of x. What happens when we reflect in the line $x=y$?

2. Deduce I.47 from III.36 (applied to the circle with center A and radius AC)

3. Inside a square $ABDE$, take a point C so that CDE is an isosceles triangle with angles 15° at D and E. What kind of triangle is ABC? (*Hint:* Inside the triangle BCD, take a point F so that FBD is congruent to CDE).

4. Prove the Erdös-Mordell theorem: If O is any point inside a triangle ABC and P, Q, R are the feet of the perpendiculars from O upon the respective sides BC, CA, AB, then

$$OA+OB+OC \geq 2(OP+OQ+OR).$$

Hint:[3] Let P_1 and P_2 be the feet of the perpendiculars from R and Q upon BC. Define analogous points Q_1 and Q_2 and R_1 and R_2 on the other sides. Using the similarity of the triangles PRP_1 and OBR, express P_1P in terms of RP, OR, and OB. After substituting such expressions into

$$OA+OB+OC \geq OA(P_1P+PP_2)/RQ+OB(Q_1Q+QQ_2)/PR+OC(R_1R+RR_2)/QP_3$$

collect the terms involving OP, OQ, OR, respectively.

5. Under what circumstances can the sign $\geq$ in Ex. 4 be replaced by $=$?

6. In the notation of Ex. 4,

$$OA \times OB \times OC \geq (OQ+OR)(OR+OP)(OP+OQ).$$

(A. Oppenheim, *American Mathematical Monthly*, **68** (1961), p. 230. See also L. J. Mordell, *Mathematical Gazette*, **46** (1962), pp. 213–215.)

3 Leon Bankoff, *American Mathematical Monthly*, **65** (1958), p. 521. For other proofs see G. R. Veldkamp and H. Brabant, *Nieuw Tijdschrift voor Wiskunde*, **45** (1958), pp. 193–196; **46** (1959), p. 87.

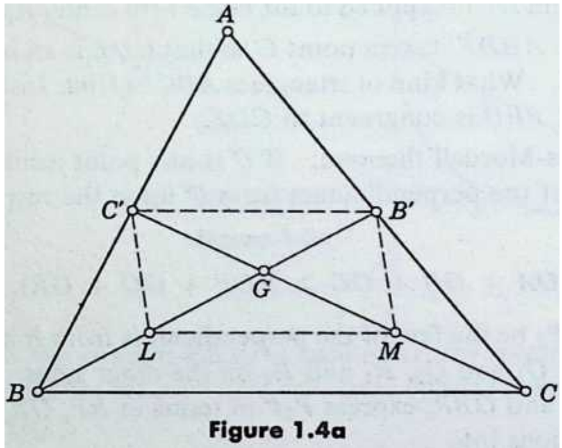


Figure 1.6.

1.4 The Medians and the Centroid

> Oriental mathematics may be an interesting curiosity, but Greek mathematics is the real thing. ...The Greeks, as Littlewood said to me once, are not clever schoolboys or ``Scholarship candidates,'' but ``Fellows of another college.'' So Greek mathematics is ``permanent,'' more permanent even than Greek literature. Archimedes will be remembered when Aeschylus is forgotten, because languages die and mathematical ideas do not.
>
> G. H. Hardy (1877–1947) [Hardy 2, p. 21]

The line joining a vertex of a triangle to the midpoint of the opposite side is called a *median.*

Let two of the three medians, say BB' and CC', meet in G (Figure 1.4a 1.6). Let L and M be the midpoints of GB and GC. By Euclid VI.2 and 4 (which were quoted on page 8), both $C'B'$ and LM are parallel to BC and half as long. Therefore $B'C'LM$ is a parallelogram. Since the diagonals of a parallelogram bisect each other, we have

$$B'G = GL = LB, C'G = GM = MC.$$

Thus the two medians BB', CC' trisect each other at G. In other words, this point G, which could have been defined as a point of trisection of one median, is also a point of trisection of another, and similarly of the third. We have thus proved [by the method of Court 1, p. 58] the following theorem:

1.4.1 1.41

The three medians of any triangle all pass through one point.

This common point G of the three medians is called the *centroid* of the triangle. Archimedes (c. 287–212 B.C.) obtained it as the center of gravity of a triangular plate of uniform density.

Exercises

1. Any triangle having two equal medians is isosceles.[4]

2. The sum of the medians of a triangle lies between $\frac{3}{4}p$ and p, where p is the sum of the sides. [Court 1, pp. 60–61.]

1.5 The Incircle and the Circumcircle

Alone at nights,

I read my Bible more and Euclid less.

Robert Buchanan (1841–1901)

(An Old Dominie's Story)

Euclid III.3 tells us that a circle is symmetrical by reflection in any diameter (whereas an ellipse is merely symmetrical about two special diameters: the major and minor axes). It follows that the angle between two intersecting tangents is bisected by the diameter through their common point.

4 It is to be understood that any exercise appearing in the form of a theorem is intended to be *proved.* It saves space to omit the words "Prove that" or "Show that."

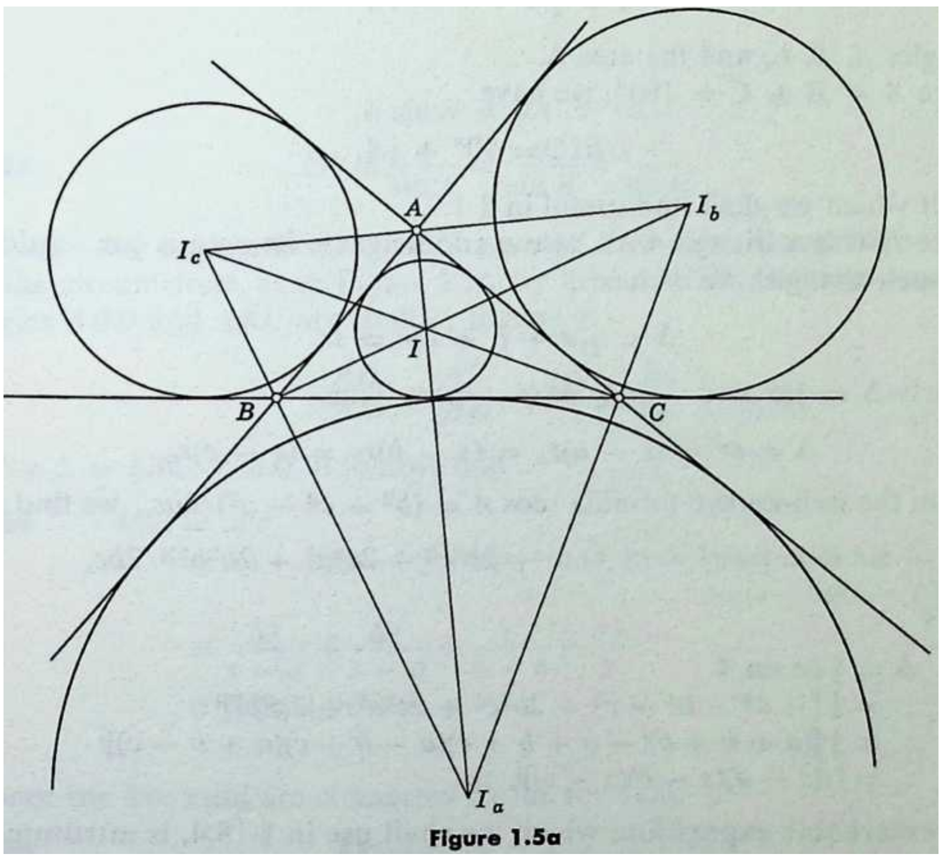


Figure 1.7.

By considering the loci of points equidistant from pairs of sides of a triangle ABC, we see that the internal and external bisectors of the three angles of the triangle meet by threes in four points I, I_a, I_b, I_c, as in Figure 1.5a 1.7. These points are the centers of the four circles that can be drawn to touch the three lines BC, CA, AB. One of them, the *incenter* I, being inside the triangle, is the center of the inscribed circle or *incircle* (Euclid IV.4). The other three are the *excenters* I_a, I_b, I_c: the centers of the three escribed circles or *excircles* [Court **2**, pp. 72–88]. The radii of the incircle and excircles are the *inradius* r and the *exradii* r_a, r_b, r_c.

In describing a triangle ABC, it is customary to call the sides

$$a = BC, b = CA, c = AB,$$

the semiperimeter

$$s = \frac{1}{2}(a + b + c),$$

the angles A, B, C, and the area Δ.

Since $A + B + C = 180^{\circ}$, we have

1.5.1 1.51

$$\angle BIC = 90^{\circ} + \frac{1}{2}A$$

,

a result which we shall find useful in §1.9.

Since IBC is a triangle with base a and height r, its area is $\frac{1}{2}ar$. Adding three such triangles we deduce

$$\Delta = \frac{1}{2}(a + b + c)r = sr.$$

Similarly

$$\Delta = \frac{1}{2}(b + c + -a)r_a = (s - a)r_a$$

Thus

1.5.2 1.52

$$\Delta = sr = (s-a)r_a = (s-b)r_b = (s-c)r_c.$$

From the well-known formula

$$\cos A = \frac{(b^2+c^2-a^2)}{2bc},$$

we find also

$$\sin A = \frac{[-a^4-b^4-c^4+2b^2c^2+2c^2a^2+2a^2b^2]^{\frac{1}{2}}}{2bc},$$

whence

$$\Delta = \frac{1}{2}bc\sin A$$

1.5.3 1.53

$$= \frac{1}{4}[-a^4-b^4-c^4+2b^2c^2+2c^2a2+2a^2b^2]^{\frac{1}{2}}$$
$$= \frac{1}{4}[(a+b+c)(-a+b+c)(a-b+c)(a+b-c)]^{\frac{1}{2}}$$
$$= [s(s-a)(s-b)(s-c)]^{\frac{1}{2}}.$$

This remarkable expression, which we shall use in §18.4, is attributed to Heron of Alexandria (about 60 a.d.), but it was really discovered by Archimedes. (See B. L. van der Waerden,*Science Awakening*, Oxford University Press, New York, 1961, pp. 228, 277.)

Another consequence of the symmetry of a circle is that the perpendicular bisectors of the three sides of a triangle all pass through the *circumcenter* O, which is the center of the circumscribed circle or *circumcircle* (Euclid IV.5). This is the only circle that can be drawn through the three vertices A, B, C. Its radius R is called the *circumradius* of the triangle. Since the "angle at the center," $\angle BOC$ (Figure 1.5b 1.8), is double the angle A, the congruent right-angled triangles OBA', OCA' each have an angle A at O, whence

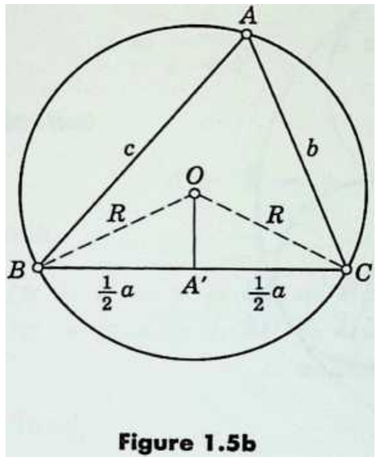


Figure 1.8.

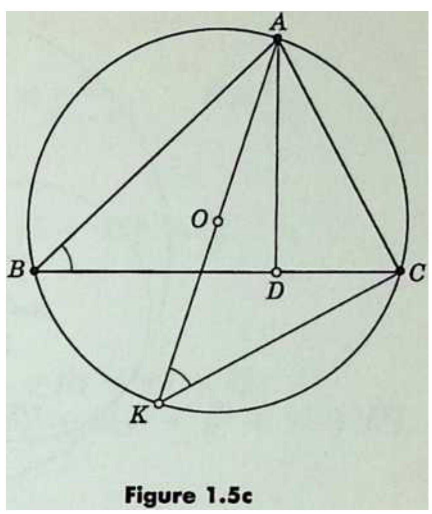


Figure 1.9.: 1.5c

$$R \sin A = BA' = \frac{1}{2}a,$$

1.5.4 1.54

$2R = \frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C}$

Draw AD perpendicular to BC, and let AK be the diameter through A of the circumcircle, as in Figure 1.5c 1.9. By Euclid III.21, the right-angled triangles ABD and AKC are similar; therefore

$$\frac{AD}{AB} = \frac{AC}{AK}, AD = \frac{bc}{2R}.$$

Since

$$\Delta = \frac{1}{2}BC \times AD$$

, it follows that

1.5.5 1.55

$$\begin{aligned} 4\Delta R = abc &= s(s-b)(s-c) + s(s-c)(s-a) + s(s-a)(s-b) - (s-a)(s-b)(s-c) \\ &= \frac{\Delta^2}{(s-a)} + \frac{\Delta^2}{(s-b)} + \frac{\Delta^2}{(s-c)} - \frac{\Delta^2}{s} \\ &\Delta(r_a + r_b + r_c - r) \end{aligned}$$

.

Hence the five radii are connected by the formula

1.5.6 1.56

$$4R = r_a + r_b + r_c - r$$

[Court **1**, p. 73].

The lengths $s - a, s-, s - c$, which appear in 1.52 as well as in Heron's formula for Δ, are easily recognized as the radii of three mutually tangent circles with centers A, B, C. Frederick Soddy (1877–1956, who is famous for his pioneering work on isotopes and for his original approach to economics) initiated a fascinating discussion of the two circles that can be drawn to touch all these three, as in Figure 1.5d 1.10, namely a small circle surrounded by the three, and a large one which usually encloses the three (though it fails to do so if the triangle is "very obtuse"). Let these two circles have centers S, S' and radii σ, σ', so that

$$SA = \sigma + s - a, SB = \sigma + s - b, SC = \sigma + s - c.$$

Also let S_a, S_b, S_c denote the angles at S in the three triangles SBC, SCA, SAB. Applying to these triangles the familiar formulas

$$\cos^2 \frac{1}{2}A = \frac{s(s-a)}{bc}, \sin^2 \frac{1}{2}A = \frac{(s-b)(s-c)}{bc}$$

for the angle A of any triangle ABC, we obtain

$$\cos^2 \frac{1}{2}S_a = \frac{(\sigma + a)\sigma}{(\sigma + s - b)(\sigma + s - c)}, \sin^2 \frac{1}{2}S_a = \frac{(s-b)(s-c)}{(\sigma + s - b)(\sigma + s - c)}$$

,

and so on.

By 1.54, we can write $\sin A, \sin B, \sin C$ in place of a, b, c in

$$a^2 - b^2 - c^2 + 2bc \cos A = 0.$$

Then we can replace A, B, C by any three angles whose sum is 180°, such as $\frac{1}{2}S_a, \frac{1}{2}S_b, \frac{1}{2}S_c$. Thus

$$\frac{(s-b)(s-c)}{(\sigma+s-b)(\sigma+s-c)} - \frac{(s-c)(s-a)}{(\sigma+s-c)(\sigma+s-a)} - \frac{(s-a)(s-b)}{(\sigma+s-a)(\sigma+s-b)} + 2\left[\frac{(s-c)(s-a)}{(\sigma+s-c)(\sigma+s-a)} \cdot \frac{(s-a)(s-b)}{(\sigma+s-a)(\sigma+s-b)}\right.$$

whence

$$\frac{\sigma+s-a}{s-a} - \frac{\sigma+s-b}{s-b} - \frac{\sigma+s-b}{s-b} - \frac{\sigma+s-c}{s-c} + 2\left[\frac{\sigma(\sigma+s-b+s-c)}{(s-b)(s-c)}\right]^{\frac{1}{2}} = 0$$

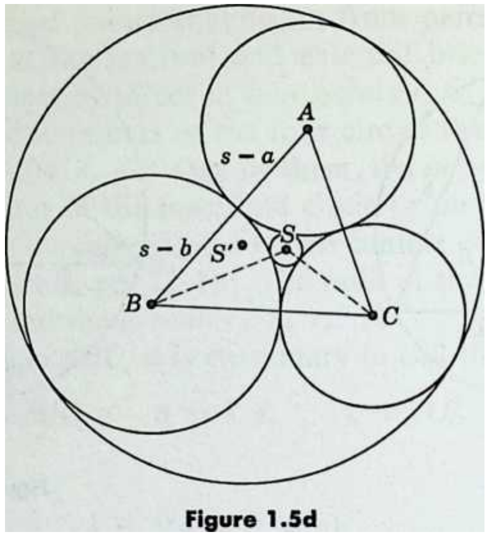


Figure 1.10.

Dividing by σ and using the abbreviations

$$\alpha = \frac{1}{s-a}, \beta \frac{1}{s-b}, \gamma \frac{1}{s-c}, \delta 1 \alpha,$$

we deduce

$$\alpha - \beta - \gamma - \delta + 2[\beta\gamma + \gamma\delta + \delta\beta]^{\frac{1}{2}} = 0,$$

whence

$$(\alpha - \beta - \gamma - \delta)^2 = 4(\beta\gamma + \gamma\delta + \delta\beta)$$
$$(\alpha - \beta - \gamma - \delta)^2 = 4(\alpha\beta + \alpha\gamma + \alpha\delta + \beta\gamma + \gamma\delta + \delta\beta)$$
$$= 2(\alpha + \beta + \gamma + \delta)^2 - 2(\alpha^2 + \beta^2 + \gamma^2 + \delta^2)$$

,

and finally

1.5.7 1.57

$$2(\alpha^2 + \beta^2 + \gamma^2 + \delta^2) = (\alpha + \beta + \gamma + \delta)^2$$

.

We have now found a perfectly symmetrical formula connecting the four quantities $\alpha, \beta, \gamma, \delta$, which are the reciprocals of the radii of four mutually tangent circles. The reciprocal of the radius of a circle is often called its *curvature.* Soddy preferred the simpler term *bend*, as in his poem *The Kiss Precise*[5] of which the middle verse runs as follows:

Four circles to the kissing come,

The smaller are the benter.

The bend is just the inverse of

The distance from the centre.

Though their intrigue left Euclid dumb

There's now no need for rule of thumb.

Since zero bend's a dead straight line

And concave bends have minus sign,

The sum of the squares of all four bends

Is half the square of their sum.

Solving 1.57 as a quadratic equation for δ, we obtain the two roots

$$\alpha + \beta + \gamma \pm 2(\beta\gamma + \gamma\alpha + \alpha\beta)^{\frac{1}{2}}.$$

The upper sign yields the larger bend, that is, the smaller circle. Thus the radii are[6]

1.5.8 1.58

$$\sigma = \left(\alpha + \beta + \gamma + 2\sqrt{\beta\gamma + \gamma\alpha + \alpha\beta}\right)^{-1} \tag{1.1}$$

5 *Nature*, 137(1936), p. 1021.

6 Steiner [1, pp. 60–63, 524]. See also Hobson [1, p. 216, Ex. 29] and J. Satterly, *Mathematics Teacher*, **53** (1960), pp. 90–95.

and

$$\sigma' = \left[\alpha + \beta + \gamma - 2\,(\beta\gamma + \gamma\alpha + \alpha\beta)^{\frac{1}{2}}\right]^{-1}$$

This last expression is usually negative, indicating a "concave bend": the circle with center S' *encloses* the circles with centers A, B, C.

Writing $(s-a)^{-1}, (s-b)^{-1}, (s-c)^{-1}$ for α, β, γ, we find

$$\sigma = \Delta/\{\frac{\Delta}{s-a} + \frac{\Delta}{s-b} + \frac{\Delta}{s-c} + 2[s(s-a+s-b+s-c)]^{\frac{1}{2}}\}$$

1.5.9 1.59

$$\Delta/(r_a + r_b + r_c + 2s)$$
$$= \Delta/(4R + r + 2s).$$

Similarly, $\sigma' = \Delta/(4R + r - 2s)$.

Exercises

1. Find the locus of the image of a fixed point P by reflection in a variable line through another fixed point O.

2. For the general triangle ABC, establish the identities

$$\frac{1}{r_a} + \frac{1}{r_b} + \frac{1}{r_c} = \frac{1}{r}, \frac{1}{\sigma} = \frac{1}{s-a} + \frac{1}{s-b} + \frac{1}{s-c} + \frac{2}{r}.$$

3. The lengths of the tangents from the vertex A to the incircle and to the three excircles are respectively

$$s-a, s, s-c, s-b.$$

4. Any triangle having two equal internal angle bisectors (each measured from a vertex to the opposite side) is isosceles. (*Hint:* If the angles B and C are not equal, one must be less, say $B < C$. Then, if the equal angle bisectors are BM and CN, there is a point P on AN such that $\angle PCN = \frac{1}{2}B$, and a point Q on PN such that $BQ = CP$. Compare the angles at P and Q in the congruent triangles BMQ and CNP.)

5. The circumcenter of an obtuse-angled triangle lies outside the triangle.

6. Where is the circumcenter of a right-angled triangle?

7. Let U, V, W be three points on the respective sides BC, CA, AB of a triangle *ABC*. The perpendiculars to the sides at these points are concurrent if and only if

$$AW^2 + BU^2 + CV^2 = WB^2 + UC^2 + VA^2.$$

8. Given a triangle ABC, for what value of x is there a point whose distances from A, B, C are equal to $x - a, x - b, x - c$? (J.A.H. Hunter.)

9. In Figure 1.54, what happens to S' if

$$2(\alpha^2 + \beta^2 + \gamma^2) = (\alpha + \beta + \gamma)^2?$$

   Sketch the case in which $a = 8$, $b = c = 5$, so that $\alpha = 1$ and $\beta = \gamma = \frac{1}{4}$.

10. A triangle is right-angled if and only if $2R + r = s$.

11. Given a point P on the circumcircle of a triangle, the feet of the perpendiculars from P to the three sides all lie on a straight line. (This line is commonly called the *Simson line* of P with respect to the triangle, although it was first mentioned by W. Wallace, thirty years after Simson's death [Johnson **1**, p. 138].)

12. Given a triangle ABC and a point P in its plane (but not on a side nor on the circumcircle), let $A_1B_1C_1$ be the derived triangle formed by the feet of the perpendiculars from P to the sides BC, CA, AB. Let $A_2B_2C_2$ be derived analogously from $A_1B_1C_1$ (using the same P), $A_3B_3C_3$ from $A_2B_2C_2$. Then $A_3B_3C_3$ is directly similar to ABC. [Casey **1**, p. 253.] (*Hint:* $\angle PBA = \angle PA_1C_1 = \angle PC_2B_2 = \angle PB_3A_3$.) This result has been extended by A. Oppenheim from the third derived triangle of a triangle to the nth derived n-gon of an n-gon.

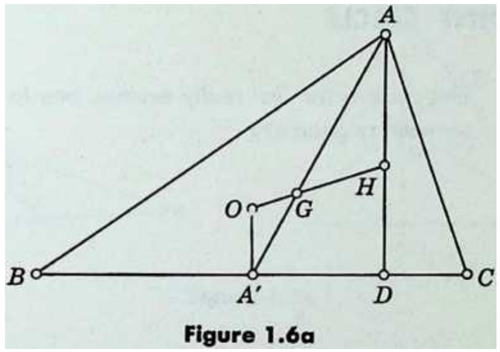


Figure 1.11.

1.6 The Euler Line and the Orthocenter

> Although the Greeks worked fruitfully, not only in geometry but also in the most varied fields of mathematics, nevertheless we today have gone beyond them everywhere and certainly also in geometry.
>
> F. Klein (1849–1925)
>
> [Klein 2, p. 189]

From now on, we shall have various occasions to mention the name of L. Euler (1707–1783), a Swiss who spent most of his life in Russia, making important contributions to all branches of mathematics. Some of his simplest discoveries are of such a nature that one can well imagine the ghost of Euclid saying, "Why on earth didn't I think of that?"

If the circumcenter O and centroid G of a triangle coincide, each median is perpendicular to the side that it bisects, and the triangle is "isosceles three ways," that is, equilateral. Hence, if a triangle ABC is not equilateral, its circumcenter and centroid lie on a unique line OG. On this so-called *Euler line*, consider a point H such that $OH = 3OG$, that is, $GH = 2OG$ (Figure 1.6a 1.11). Since also $GA = 2A'G$, the latter half of Euclid VI.2 tells us that AH is parallel to $A'O$, which is the perpendicular bisector of BC. Thus AH is perpendicular to BC. Similarly BH is perpendicular to CA, and CH to AB.

The line through a vertex perpendicular to the opposite side is called an *altitude.* The above remarks [cf. Court 2, p. 101] show that

> The three altitudes of any triangle all pass through one point on the Euler line.

This common point H of the three altitudes is called the *orthocenter* of the triangle.

Exercises

1. Through each vertex of a given triangle ABC draw a line parallel to the opposite side. The perpendicular bisectors of the sides of the triangle so formed suggest an alternative proof that the three altitudes of ABC are concurrent. [Gauss 1, vol. 4, p. 396.]

2. The orthocenter of an obtuse-angled triangle lies outside the triangle.

3. Where is the orthocenter of a right-angled triangle?

4. Any triangle having two equal altitudes is isosceles.

5. Construct an isosceles triangle ABC (with base *BC)*, given the median BB' and the altitude BE. (*Hint:* The centroid is two-thirds of the way from B to B'.) (H. Freudenthal.)

6. The altitude AD of any triangle ABC is of length

$$2R \sin B \sin C$$

.

7. Find the perpendicular distance from the centroid G to the side BC.

8. If the Euler line passes through a vertex, the triangle is either right-angled or isosceles (or both).

9. If the Euler line is parallel to the side BC, the angles B and C satisfy

$$\tan B \tan C = 3$$

.

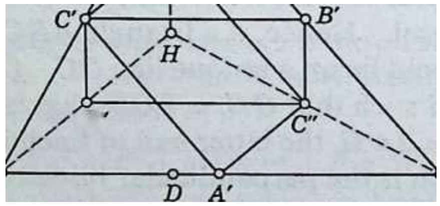


Figure 1.12.

1.7 The Nine-point Circle

This circle is the first really exciting one to appear in any course on elementary geometry.

Daniel Pedoe (1910–)

[Pedoe 1, p. 1]

The feet of the altitudes (that is, three points like D in Figure 1.6a 1.11) form the *orthic triangle* (or "pedal triangle") of ABC. The circumcircle of the orthic triangle is called the *nine-point circle* (or "Feuerbach circle") of the original triangle, because it contains not only the feet of the three altitudes but also six other significant points. In fact,

1.7.1 1.71

The midpoints of the three sides, the midpoints of the lines joining the orthocenter to the three vertices, and the feet of the three altitudes, all lie on a circle.

Proof [Coxeter **2**, 9.29]. Let $A', B', C', A'', B'', C''$ be the midpoints of BC, CA, AB, HA, HB, HC, and let D, E, F be the feet of the altitudes, as in Figure 1.7a 1.12. By Euclid VI.2 and 4 again, both CB' and $B''C''$ are parallel to BC while both $B'C''$ and $C'B''$ are parallel to AH. Since AH is perpendicular to BC, it follows that $B'C'B''C''$ is a rectangle. Similarly $C'A'C'A''$ is a rectangle. Hence $A'A'', B'B'', C'C''$ are three diameters of a circle. Since these diameters subtend right angles at D, E, F, respectively, the same circle passes through these points too.

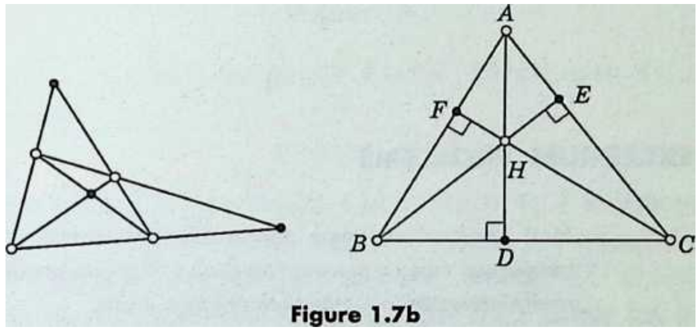


Figure 1.13.

If four points in a plane are joined in pairs by six distinct lines, they are called the *vertices* of a *complete quadrangle*, and the lines are its six *sides.* Two sides are said to be *opposite* if they have no common vertex. Any point of intersection of two opposite sides is called a *diagonal point.* There may be as many as three such points (see Figure 1.7b 1.13).

If a triangle ABC is not right-angled, its vertices and orthocenter form a special kind of quadrangle whose opposite sides are perpendicular. In this terminology, the concurrence of the three altitudes can be expressed as follows:

1.7.2 1.72

If two pairs of opposite sides of a complete quadrangle are pairs of perpendicular lines, the remaining sides are likewise perpendicular.

Such a quadrangle $ABCH$ is called an *orthocentric quadrangle.* Its six sides

$$BC, CA, AB, HA, HB, HC$$

are the sides and altitudes of the triangle ABC, and its diagonal points D, E, F are the feet of the altitudes. Among the four vertices of the quadrangle, our notation seems to give a special role to the vertex H. Clearly, however,

1.7.3 1.73

Each vertex of an orthocentric quadrangle is the orthocenter of the triangle formed by the remaining three vertices.

The four triangles (just one of which is acute-angled) all have the same orthic triangle and consequently the same nine-point circle.

It is proved in books on affine geometry [such as Coxeter 2, 8.71] that the midpoints of the six sides of any complete quadrangle and the three diagonal points all lie on a conic. The above remarks show that, when the quadrangle is orthocentric, this "nine-point conic" reduces to a circle.

Exercises

1. Of the nine points described in 1.71, how many coincide when the triangle is (a) isosceles, (b) equilateral?

2. The feet of the altitudes decompose the nine-point circle into three arcs. If the triangle is scalene, the remaining six of the nine points are distributed among the three arcs as follows: One arc contains just one of the six points, another contains two, and the third contains three.

3. On the arc $A'D$ of the nine-point circle, take the point X one-third of the way from A' to D. Take points Y, Z similarly, on the arcs $B'E, C'F$. Then XYZ is an equilateral triangle.

4. The incenter and the excenters of any triangle form an orthocentric quadrangle. [Casey 1, p. 274.]

5. In the notation of §1.5, the Euler line of $I_a lblc$ is IO.

6. The four triangles that occur in an orthocentric quadrangle have equal circumradii.

1.8 Two Extremum Problems

> Most people have some appreciation of mathematics, just as most people can enjoy a pleasant tune; and there are probably more people really interested in mathematics than in music.

G. H. Hardy **[2**, p. 26]

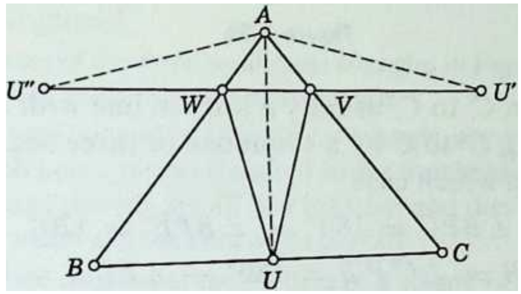


Figure 1.14.: Figure 1.8a

> Their interest will be stimulated if only we can eliminate the aversion toward mathematics that so many have acquired from childhood experiences.

Hans Rademacher (1892–)

[Rademacher and Toeplitz 1, p. 5]

We shall describe the problems of Fagnano and Fermat in considerable detail because of the interesting methods used in solving them. The first was proposed in 1775 by J. F. Toschi di Fagnano, who solved it by means of differential calculus. The method given here was discovered by L. Fejer while he was a student [Rademacher and Toeplitz 1, pp. 30–32].

FAGNANO′S PROBLEM. *In a given acute-angled triangle ABC, inscribe a triangle UVW whose perimeter is as small as possible.*

Consider first an arbitrary triangle UVW with U on BC, V on CA, W on AB. LetU', U'' be the images of U by reflection in CA, AB, respectively. Then

$$UV + VW + WU = U'V + VW + WU'',$$

which is a path from U' to U'', usually a broken fine with angles at V and W. Such a path from U' to U'' is minimal when it is straight, as in Figure 1.8a.

Hence, among all inscribed triangles with a given vertex U on BC, the one with smallest perimeter occurs when V and W lie on the straight line $U'U''$. In this way we obtain a definite triangle UVW for each choice of U on *BC.* The problem will be solved when we have chosen Uso as to minimize $U'U''$, which is equal to the perimeter.

Since AU' and AU'' are images of AU by reflection in AC and AB, they are congruent and

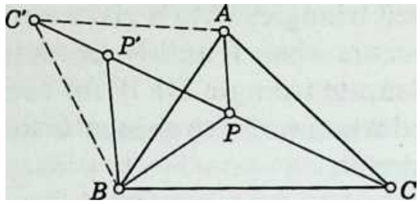


Figure 1.15.

$$\angle U'AU'' = 2A.$$

Thus $AU'U''$ is an isosceles triangle whose angle at A is independent of the choice of U. The base $U'U''$ is minimal when the equal sides are minimal, that is, when AU is minimal. In other words, AU is the shortest distance from the given point A to the given line BC. Since the hypotenuse of a right-angled triangle is longer than either cathetus, the desired location of U is such that AU is perpendicular to BC. Thus AU is the altitude from A. This choice of U yields a unique triangle UVW whose perimeter is smaller than that of any other inscribed triangle. Since we could equally well have begun with B or C instead of A, we see that BV and CW are the altitudes from B and C. Hence *The triangle of minimal perimeter inscribed in an acute-angled triangle* ABC *is the orthic triangle of* ABC*.*

The same method can be used to prove the analogous result for spherical triangles[Steiner 2, p. 45, No. 7].

The other problem, proposed by Pierre Fermat (1601–1665), likewise seeks to minimize the sum of three distances. The solution given here is due to J. E. Hofmann.[7]

FERMAT'S PROBLEM. *In a given acute-angled triangle ABC, locate a point* P *whose distances from* A, B, C *have the smallest possible sum.*

Consider first an arbitrary point P inside the triangle. Join it to A, B, C and rotate the inner triangle APB through 60° about B to obtain $C'P'B$, so that ABC' and PBP' are equilateral triangles, as in Figure 1.8b. Then

$$AP + BP + CP = C'P' + P'P + PC,$$

7 Elementare Losung einer Minimumsaufgabe , *Zeitschrift fUr mathematischen und naturwissenschaftlichen Unterricht*, **60** (1929), pp. 22–23.

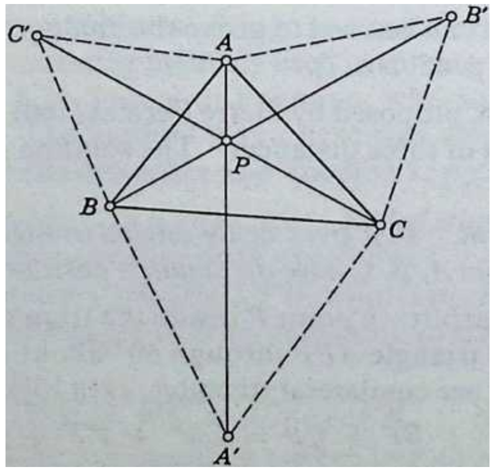


Figure 1.16.

which is a path from C to C, usually a broken line with angles at P' and P. Such a path (joining C' to C by a sequence of three segments) is minimal when it is straight, in which case

$$\angle BPC = 180^\circ - \angle BPP' = 120^\circ$$

and $\angle APB = \angle CP'B = 180^\circ - \angle PP'B = 120^\circ$.

Thus the desired point P, for which $AP + BP + CP$ is minimal, is the point from which each of the sides BC, CA, AB subtends an angle of 120°. This "Fermat point" is most simply constructed as the second intersection of the line CC' and the circle ABC' (that is, the circumcircle of the equilateral triangle ABC').

It has been pointed out [for example by Pedoe **1**, pp. 11–12] that the triangle ABC need not be assumed to be acute-angled. The above solution is valid whenever there is no angle greater than 120°.

Instead of the equilateral triangle ABC' on AB, we could just as well have drawn an equilateral triangle BCA' on BC, or CAB' on CA, as in Figure 1.8c. Thus the three lines AA', BB', CC' all pass through the Fermat point P, and any two of them provide an alternative construction for it. Moreover, the line segments AA', BB', CC' are all equal to $AP + BP + CP$. Hence

If equilateral triangles BCA', CAB', ABC are drawn outwards on the sides of any triangle ABC, the line segments AA', BB', CC' are equal, concurrent, and inclined at 60° to one another.

Exercises

1. In Figure 1.8a, UV and VW make equal angles with CA. Deduce that the orthocenter of any triangle is the incenter of its orthic triangle. (In other words, if ABC is a triangular billiard table, a ball at U, hit in the direction UV, will go round the triangle UVW indefinitely, that is, until it is stopped by friction.)

2. How does Fagnano's problem collapse when we try to apply it to a triangle ABC in which the angle A is obtuse?

3. The circumcircles of the three equilateral triangles in Figure 1.8c all pass through P, and their centers form a fourth equilateral triangle.[8]

4. Three holes, at the vertices of an arbitrary triangle, are drilled through the top of a table. Through each hole a thread is passed with a weight hanging from it below the table. Above, the three threads are all tied together and then released. If the three weights are all equal, where will the knot come to rest?

5. Four villages are situated at the vertices of a square of side one mile. The inhabitants wish to connect the villages with a system of roads, but they have only enough material to make $\sqrt{3}+1$ miles of road. How do they proceed? [Courant and Robbins **1**, p. 392.]

6. Solve Fermat's problem for a triangle ABC with $A > 120^\circ$, and for a convex quadrangle $ABCD$.

7. If two points P, P', inside a triangle ABC, are so situated that $\angle CBPZPBP' = ZP'BA$, $\angle ACP' = \angle P'CP = \angle PCB$, then $\angle BP'P = \angle PP'C$.

8. If four squares are placed externally (or internally) on the four sides of any parallelogram, their centers are the vertices of another square. [Yaglom 1, pp. 96–97.]

9. Let X, Y, Z be the centers of squares placed externally on the sides BC, CA, AB of a triangle ABC.Then the segment AX is congruent and perpendicular to YZ (also BY to ZX and CZ to XY).(W. A. J. Luxemburg.)

8 Court [1, pp. 105–107]. See also *Mathesis* 1938, p. 293 (footnote, where this theorem is attributed to Napoleon); and Forder **[2**, p. 40] for some interesting generalizations.

10. Let Z, X, U, V be the centers of squares placed externally on the sides AB, BC, CD, DA of any simple quadrangle (or "quadrilateral") $ABCD$. Then the segment ZU (joining the centers of two "opposite" squares) is congruent and perpendicular to XV. [Forder **2**, p. 40.]

1.9 Morley's Theorem

> Many of the proofs in mathematics are very long and intricate. Others, though not long, are very ingeniously constructed.
>
> E. C. Titchmarsh (1899–1963)
> [Titchmarsh 1, p. 23]

One of the most surprising theorems in elementary geometry was discovered about 1899 by F. Morley (whose son Christopher wrote novels such as *Thunder on the Left*.) He mentioned it to his friends, who spread it over the world in the form of mathematical gossip. At last, after ten years, a trigonometrical proof by M. Satyanarayana and an elementary proof by M. T. Naraniengar were published.[9]

MORLEY′S THEOREM. *The three points of intersection of the adjacent trisectors of the angles of any triangle form an equilateral triangle.*

In other words, any triangle ABC yields an equilateral triangle PQR if the angles A, B, C are trisected by AQ and AR, BR and BP, CP and CQ, as in Figure 1.9a. (Much trouble is experienced if we try a direct approach, but the difficulties disappear if we work backwards, beginning with an equilateral triangle and building up a general triangle which is afterwards identified with the given triangle ABC.)

On the respective sides QR, RP, PQ of a given equilateral triangle PQR, erect isosceles triangles $P'QR, Q'RP, R'PQ$ whose base angles α, β, γ satisfy the equation and inequalities

$$\alpha + \beta + \gamma = 120^\circ, \alpha < 60^\circ, \beta < 60^\circ, \gamma < 60^\circ.$$

9 *Mathematical Questions and their Solutions from the Educational Times* (New Series), 1**5** (1909), pp. 23–24,47. See also C. H. Chepmell and R. F. Davis, *Mathematical Gazette*, **11** (1923), pp. 85–86; F. Morley, *American Journal of Mathematics*, **51** (1929), pp. 465–472, H. D. Grossman, *American Mathematical Monthly*, **50** (1943), p. 552, and L. Bankoff, *Mathematics Magazine*, **35** (1962), pp. 223–224. The treatment given here is due to Raoul Bricard, *Nouvelles Annates de Mathmatiques* (5), **1** (1922), pp. 254—258. A similar proof was devised independently by Bottema [1, p. 34].

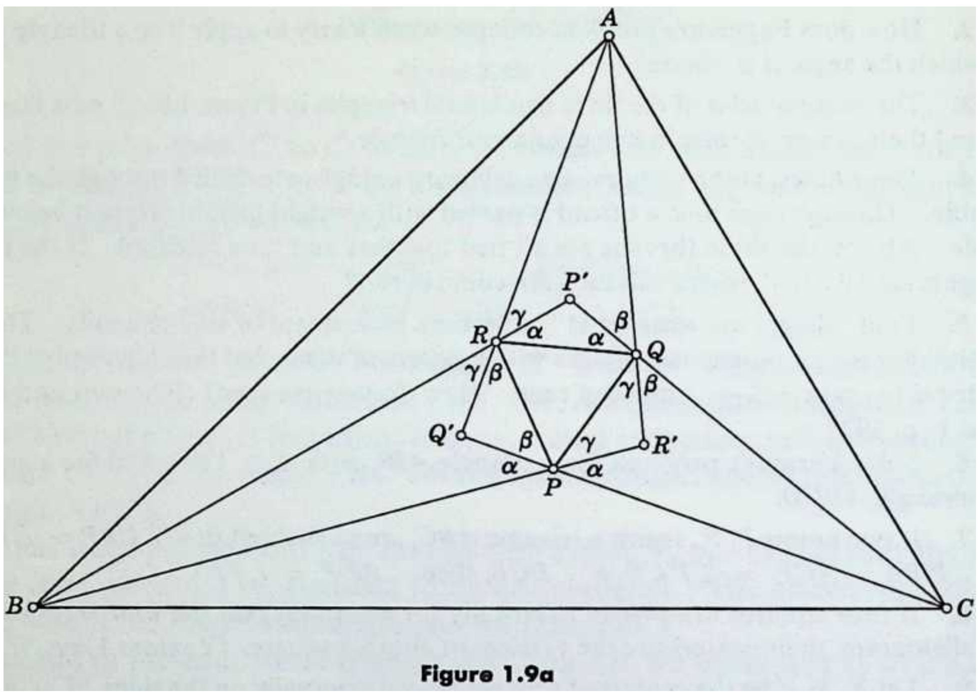


Figure 1.17.

Extend the sides of the isosceles triangles below their bases until they meet again in points A, B, C. Since $\alpha + \beta + \gamma + 60^\circ = 180^\circ$, we can immediately infer the measurement of some other angles, as marked in Figure 1.9a. For instance, the triangle AQR must have an angle $60^\circ—\alpha$ at its vertex A, since its angles at Q and R are $\alpha + \beta$ and $\gamma + \alpha$.

Referring to 1.51, we see that one way to characterize the incenter I of a triangle ABC is to describe it as lying on the bisector of the angle A at such a distance that

$$\angle BIC = 90^\circ + \frac{1}{2}A.$$

Applying this principle to the point P in the triangle $P'BC$, we observe that the line PP' (which is a median of both the equilateral triangle PQR and the isosceles triangle $P'Q$) bisects the angle at P'. Also the half angle at $P' 90^\circ—\alpha$, and

$$\angle BPC = 180^\circ - \alpha = 90^\circ + (90^\circ - \alpha).$$

Hence P is the incenter of the triangle$P'BC$. Likewise Q is the incenter of $Q'CA$, and R of$R'AB$. Therefore all the three small angles at C are equal; likewise at A and at B. In other words, the angles of the triangleABC are trisected.

The three small angles at A are each $\frac{1}{3}A = 60^\circ - \alpha$; similarly at$B$and C. Thus

$$\alpha - 60^\circ - \frac{1}{3}A, \beta = 60^\circ - \frac{1}{3}B, \gamma = 60^\circ - \frac{1}{3}C.$$

By choosing these values for the base angles of our isosceles triangles, we can ensure that the above procedure yields a triangle ABC that is similar to any given triangle.

This completes the proof.

Exercises

1. The three lines PP', QQ', RR' (Figure 1.9a) are concurrent. In other words, the trisectors of A, B, C meet again to form another triangle $P'Q'R'$ which is perspective with the equilateral triangle PQR. (In general $P'Q'R'$ is *not* equilateral.)

2. What values of α, β, γ will make the triangle ABC (i) equilateral, (ii) right-angled isosceles? Sketch the figure in each case.

3. Let P_1 and P_2 (on CA and AB) be the images of P by reflection in CP' and BP'. Then the four points P_1, Q, R, P_2 are evenly spaced along a circle through A. In the special case when the triangle ABC is equilateral, these four points occur among the vertices of a regular enneagon (9-gon) in which A is the vertex opposite to the side QR.

2 Regular Polygons

We begin this chapter by discussing (without proofs) the possibility of constructing certain regular polygons with the instruments allowed by Euclid. We then consider all these polygons, regardless of the question of constructibility, from the standpoint of symmetry. Finally, we extend the concept of a regular polygon so as to include star polygons.

2.1 Cyclotomy

> One, two! One, two! And through and through The vorpal blade went snicker-snack!
>
> Lewis Carroll
>
> [Dodgson 2, Chap. 1]

Euclid's postulates imply a restriction on the instruments that he allowed for making constructions, namely the restriction to ruler (or straightedge) and compasses. He constructed an equilateral triangle (I.1), a square (IV.6), a regular pentagon (IV.11), a regular hexagon (IV.15), and a regular 15-gon (IV. 16). The number of sides may be doubled again and again by repeated angle bisections. It is natural to ask which other regular polygons can be constructed with Euclid's instruments. This question was completely answered by Gauss (1777–1855) at the age of nineteen [see Smith 2, pp. 301–302]. Gauss found that a regular n-gon, say $\{n\}$, can be so constructed if and only if the odd prime factors of n are distinct "Fermat primes"

$$F_k = 2^{2^k} + 1.$$

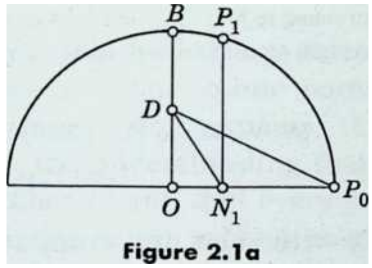


Figure 2.1.

The only known primes of this kind are

$$F_0 = 2^1 + 1 = 3, F_1 = 2^2 + 1 = 5, F_2 = 2^4 + 1 = 17, F_3 = 2^8 + 1 = 257, F_4 = 2^{16} + 1 = 65537.$$

Since 7 is not a Fermat prime, Euclid's instruments will not suffice for the construction of a regular heptagon {7}. Since the factors of 9 are not distinct, the same is true for a regular enneagon {9}.

To inscribe a regular pentagon in a given circle, simpler constructions than Euclid's were given by Ptolemy and Richmond.[1] The former has been repeated in many textbooks. The latter is as follows (Figure 2.1a 2.1).

To inscribe a regular pentagon $P_0P_1P_2P_3P_4$ in a circle with center O: draw the radius OB perpendicular to OP_Q; join P_0 to D, the midpoint of OB; bisect the angle ODP_0 to obtain N_1 on OP_0; and draw N_1P_1 perpendicular to OP_0 to obtain P_1 on the circle. Then P_0P_1 is a side of the desired pentagon.

Richmond also gave a simple construction for the {17} $P_0P_1 \ldots P_1 6$ (Figure 2.1b 2.2). Join P_0 to J, one quarter of the way from O to B. On the diameter through P_0 take E, F, so that $\angle OJE$ is one quarter of OJP_0 and $\angle FJE$ is 45°. Let the circle on FP_0 as diameter cut OB in K, and let the circle with center E and radius EK cut OP_0 in N_3 (between O and P_0) and N_5. Draw perpendiculars to OP_0 at these two points, to cut the original circle in P_3 and P_5. Then the arc P_3P_5 (and likewise P_1P_3) is $\frac{2}{17}$ of the circumference. (The proof involves repeated application of the principle that the roots of the equation $x^2 + 2x \cot 2C - 1 = 0$ are $\tan C$ and $-\cot C$.)

1 H. W. R_1chmond, *Quarterly Journal of Mathematics,* **26** (1893), pp. 296–297; see also H. E. Dudeney, *Amusements in Mathematics* (London 1917), p. 38.

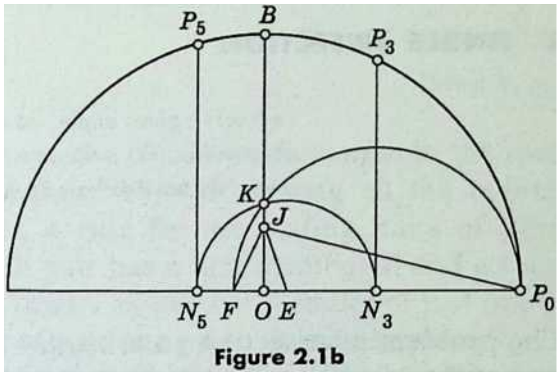


Figure 2.2.

Richelot and Schwendenwein constructed the regular 257-gon about 1898. J. Hermes spent ten years on the regular 65537-gon and deposited the manuscript in a large box in the University of Göttingen, where it may still be found.

The next number of the form $F_k = 2^{2^k} + 1$ is $F_5 = 4294967297$. Fermat incorrectly assumed it to be prime. G. T. Bennett gave the following neat proof[2] that it is composite [Hardy and Wright **1,** p. 14]: the number

$$641 = 5^4 + 2^4 = 5 \cdot 2^7 + 1$$

dividing both

$$5^4 \cdot 2^{28} + 2^{32}$$

and

$$5^4 \cdot 2^{28} - 1$$

divides their difference, which is F_5.

The question naturally arises whether F_k may be prime for some greater value of k. It is now known that this can happen only if F_k divides $3^{(F_k-1)/2} + 1$. Using this criterion, electronic computing machines have shown that F_k is composite for $5 \le k \le 16$. Therefore Hermes's construction is the last of its kind that will ever be undertaken!

2 Rediscovered by P. Kanagasabapathy, *Mathematical Gazette,* **42** (1958), p. 310.

EXERCISES

1. Verify the correctness of Richmond's construction for {5} (Figure 2.1a).

2. Assuming Richmond's construction for {17}, how would you inscribe {51} in the same circle?

2.2 Angle Trisection

> To trisect a given angle, we may proceed to find the sine of the angle—say a—then, if x is the sine of an angle equal to one-third of the given angle, we have $4x^3 = 3x - a$.
>
> W. W. Rouse Ball (1850--1925)
>
> [Ball 1 p. 327]

The problem of trisecting an arbitrary angle with ruler and compasses exercised the ingenuity of professional and amateur mathematicians for two thousand years [Ball 1, pp. 333–335]. It is, of course, easy to trisect certain particular angles, such as a right angle. But any construction for trisecting an arbitrary angle could be applied to an angle of 60°, and then we could draw a regular enneagon. Since the number 9 has 3 as a repeated factor, this polygon cannot be drawn with ruler and compasses. In view of Gauss's discovery, we may say that it has been known since 1796 that the classical trisection problem can never be solved.

This is probably the reason why Morley's Theorem (§1.9) was not discovered till the twentieth century: people felt uneasy about mentioning the trisectors of an angle. However, although the trisectors cannot be constructed by means of the ruler and compasses, they can be found in other ways [Cundy and Rollett 1, pp. 208–211]. Even if these more versatile instruments had never been discovered, the theorem would still be meaningful. Most mathematicians are willing to accept the existence of things that they have not been able to construct. For instance, it was proved in 1909 that the Fermat numbers F_7 and F_8 are composite, but their smallest prime factors still remain to be computed.

EXERCISE

The number $2^n + 1$ is composite whenever n is not a power of 2.

2.3 Isometry

> One way of describing the structure of space, preferred by both Newton and Helmholtz, is through the notion of congruence. Congruent parts of space V, V' are such as can be occupied by the same rigid body in two of its positions. If you move the body from one into the other position the particle of the body covering a point P of V will afterwards cover a certain point P of V, and thus the result of the motion is a mapping $P \to P'$ of V upon V'. We can extend the rigid body either actually or in imagination so as to cover an arbitrarily given point P of space, and hence the congruent mapping $P \to P'$ can be extended to the entire space.
>
> Hermann Weyl (1885--1955)
>
> [Weyl 1, p, 43]

We shall find it convenient to use the word transformation in the special sense of a one-to-one correspondence $P \to P'$ among all the points in the plane (or in space), that is, a rule for associating pairs of points, with the understanding that each pair has a first member P and a second member P' and that every point occurs as the first member of just one pair and also as the second member of just one pair. It may happen that the members of a pair coincide, that is, that P' coincides with P; in this case P is called an *invariant* point (or "double point") of the transformation.

In particular, an *isometry* (or "congruent transformation," or "congruence") is a transformation which preserves length, so that, if (P, P') and (Q, Q') are two pairs of corresponding points, we have $PQ = P'Q'$: PQ and $P'Q'$ are congruent segments. For instance, a *rotation* of the plane about P (or about a line through P perpendicular to the plane) is an isometry having P as an invariant point, but a *translation* (or "parallel displacement") has no invariant point: every point is moved.

A *reflection* is the special kind of isometry in which the invariant points consist of all the points on a line (or plane) called the *mirror*.

A still simpler kind of transformation (so simple that it may at first seem too trivial to be worth mentioning) is the *identity,* which leaves every point unchanged. The result of applying several transformations successively is called their *product.* If the product of two transformations is the identity, each is called the *inverse* of the other, and their product in the reverse order is again the identity.

2.3.1 2.31

If an isometry has more than one invariant point, it must be either the identity or a reflection.

To prove this, let A and B be two invariant points, and P any point not on the line AB (Figure 1.36). The corresponding point P', satisfying

$$AP' = AP, BP' = BP,$$

must He on the circle with center A and radius AP, and on the circle with center B and radius BP. Since P is not on AB, these circles do not touch each other but intersect in two points, one of which is P. Hence P' is either P itself or the image of P by reflection in AB.

2.4 Symmetry

> Tyger! Tyger! burning bright
> In the forests of the night,
> What immortal hand or eye
> Dare frame thy fearful symmetry?
>
> William Blake (1757--1827)

When we say that a figure is "symmetrical" we mean that we can apply certain isometries, called *symmetry operations,* which leave the whole figure unchanged while permuting its parts. For example, the capital letters E and A (Figure 2.4a) have bilateral symmetry, the mirror being horizontal for the former, vertical for the latter. The letter N (Figure 2.46) is symmetrical by a *half-turn,* or rotation through 180° (or "reflection in a point," or "central inversion"), which may be regarded as the result of reflecting horizontally and then vertically, or vice versa. The swastika (Figure 2.4c) is symmetrical by rotation through any number of right angles.

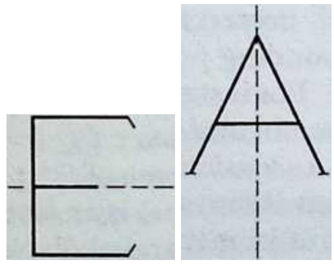


Figure 2.3.

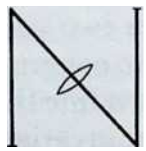


Figure 2.4.

In counting the symmetry operations of a figure, it is usual to include the identity; any figure has this trivial symmetry. Thus the swastika admits four distinct symmetry operations: rotations through $1, 2, 3, or 4$ right angles. The last is the identity. The first and third are inverses of each other, since their product is the identity.

This use of the word "product" suggests an algebraic symbolism in which the transformations are denoted by capital letters while 1 denotes the identity. (Instead of 1, some authors write E.) Thus if S is the counterclockwise quarter-turn, the four symmetry operations of the swastika are

$$S, S^2, S^3 = S^{-1} and S^4 = 1.$$

Since the smallest power of S that is equal to the identity is the fourth power, we say that S is of *period* 4. Similarly S^2, being a half-turn, is of period 2 [see Coxeter **1,** p. 39]. The only transformation of period 1 is the identity. A translation is aperiodic (that is, it has no period), but it is conveniently said to be of infinite period.

Some figures admit both reflections and rotations as symmetry operations. The letter H (Figure 2.4c?) has a horizontal mirror (like E) and a vertical mirror (like A), as well as a center of rotational symmetry (like N) where the two mirrors intersect. Thus it has four symmetry operations: the identity 1, the horizontal reflection R_1, the vertical reflection R_2, and the half-turn $R_1R_2 = R_2R_1$.

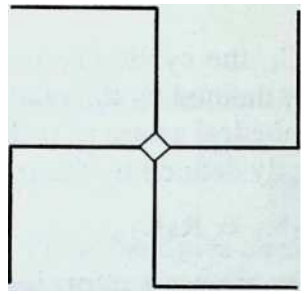


Figure 2.5.

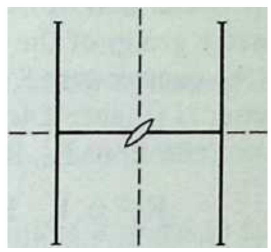


Figure 2.6.

EXERCISES

1. Every isometry of period 2 is either a reflection or a half-turn [Bachmann **1,** pp. 2–3],

2. Express (a) a half-turn, (b) a quarter-turn, as transformations of (i) Cartesian coordinates, (ii) polar coordinates. (Take the origin to be the center of rotation.)

2.5 Groups

Symmetry, as wide or as narrow as you may define its meaning, is one idea by which man through the ages has tried to comprehend and create order, beauty, and perfection.

Hermann Weyl [1, p. 5]

A set of transformations [Birkhoff and MacLane **1,** pp. 119–122] is said to form a *group* if it contains the inverse of each and the product of any two (including the product of one with itself or with its inverse). The number of distinct transformations is called the *order of* the group. (This may be either finite or infinite.) Clearly the symmetry operations of any figure form a group. This is called the *symmetry group* of the figure. In the extreme case when the figure is completely irregular (like the numeral 6) its symmetry group is of order one, consisting of the identity alone.

The symmetry group of the letter E or A (Figure 2.4a) is the so-called *dihedral* group of order 2, generated by a single reflection and denoted by D_1. (The name is easily remembered, as the Greek origin of the word "dihedral" is almost equivalent to the Latin origin of "bilateral.") The symmetry group of the letter N (Figure 2.4b) is likewise of order 2, but in this case the generator is a half-turn and we speak of the *cyclic* group, C_2. The two groups D_1 and C_2 are abstractly identical or *isomorphic;* they are different geometrical representations of the single abstract group of order 2, defined by the relation

2.5.1 2.51

$$R^2 = 1$$

or $R = R^{-1}$ [Coxeter and Moser 1, p. 1].

The symmetry group of the swastika is C_4, the cyclic group of order 4, generated by the quarter-turn S and abstractly defined by the relation $S^4 = 1$. That of the letter H (Figure 2.4d) is D_2, the dihedral group of order 4, generated by the two reflections R_1, R_2 and abstractly defined by the relations

2.5.2 2.52

$$R_1^2 = 1, R_2^2 = 1, R_1R_2 = R_2R_1.$$

Although C_4 and D_2 both have order 4, they are *not* isomorphic: they have a different structure, different "multiplication tables." To see this, it suffices to observe that C_4 contains two operations of period 4, whereas all the operations in D_2 (except the identity) are of period 2: the generators obviously, and their product also, since

$$(R_1R_2)^2 = R_1R_2R_1R_2 = R_1R_2R_2R_1 = R_1R_2^2R_1 = R_1R_1 = R_1^2 = 1.$$

This last remark illustrates what we mean by saying that 2.52 is an *abstract definition* for D_2, namely that every true relation concerning the generators R_1, R_2 is an algebraic consequence of these simple relations. An alternative abstract definition for the same group is

2.5.3 2.53

$$R_1^2 = 1, R_22 = 1, (R_1R_2)2 = 1,$$

from which we can easily deduce $R_1R_2 = R_2R_1$.

The general cyclic group C_n, of order n, has the abstract definition

2.5.4 2.54

$$S^n = 1.$$

Its single generator S, of period n, is conveniently represented by a rotation through $360^\circ/n$. Then S^k is a rotation through k times this angle, and the n operations in C_n are given by the values of k from 1 to n, or from 0 to $n-1$. In particular, C_5 occurs in nature as the symmetry group of the periwinkle flower.

EXERCISE

Express a rotation through angle a about the origin as a transformation of (i) polar coordinates, (ii) Cartesian coordinates. If $f(r, 0) = 0$ is the equation for a curve in polar coordinates, what is the equation for the transformed curve?

2.6 The Product Of Two Reflections

> Thou in thy lake dost see
> Thyself.
>
> J.M. Legare (1823--1859)
>
> (To a Lily)

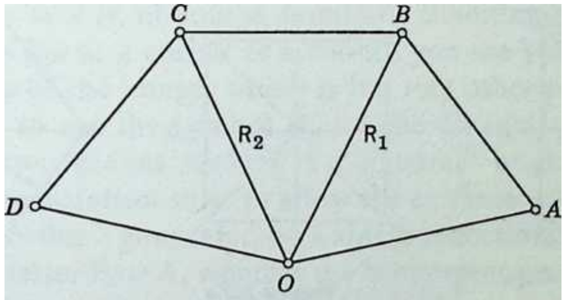


Figure 2.6a

Figure 2.7.

In any group of transformations, the associative law

$$(RS)T = R(ST)$$

is automatically satisfied, but the commutative law

$$RS = SR$$

does not necessarily hold, and care must be taken in inverting a product, for example,

$$(RS)^{-1} = S^{-1}R^{-1},$$

not $R^{-1}S^{1}$. (This becomes clear when we think of R and S as the operations of putting on our socks and shoes, respectively.)

The product of reflections in two intersecting fines (or planes) is a rotation through twice the angle between them. In fact, if $A, B, C, D, \ldots$ are evenly spaced on a circle with center O, let R_1 and R_2 be the reflections in OB and OC (Figure 2.6a). Then R_1 reflects the triangle OAB into OAB, which is reflected by R_2 to OCD; thus R_1R_2 is the rotation through $\angle AOC$ or $\angle BOD$, which is twice $\angle BOC$. Since a rotation is completely determined by its center and its angle, R_1R_2 is equal to the product of reflections in any two lines through O making the same angle as OB and OC. (The reflections in OA and OB are actually $R_1R_2R_1$ and R_1 whose product is $R_1R_2R_1^2 = R_1R_2$.) In particular, the half-turn about $O$is the product of reflections in any two perpendicular lines through O.

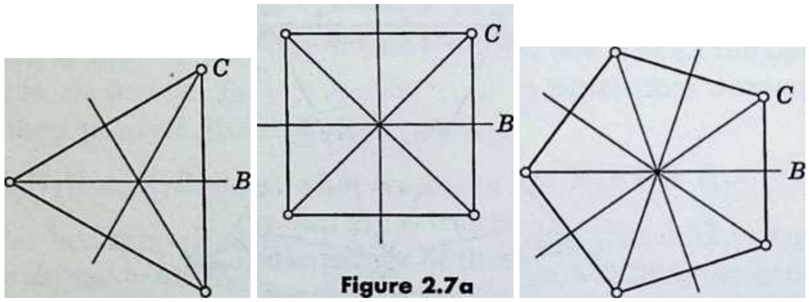


Figure 2.8.

Since R_1R_2 is a counterclockwise rotation, R_2R_1 is the corresponding clockwise rotation; in fact,

$$R_2R_1 = R_2^{-1}R_1^{-1} = (R_1R_2)^{-1}.$$

This is the same as R_1R_2 if the two mirrors are at right angles, i n which case R_1R_2 is a half-turn and $(R_1R_2)^2 = 1$.

EXERCISES

1. The product of quarter-turns (in the same sense) about C and B is the half-turn about the center of a square having BC for a side.

2. Let $ACPQ$ and $BARS$ be squares on the sides AC and BA of a triangle ABC. If B and C remain fixed while A varies freely, PS passes through a fixed point.

2.7 The Kaleidoscope

$D_2$is a special case of the general dihedral group D_n, which is , for$n > 2$, the symmetry group of the regular n-gon, $\{n\}$. (See Figure 2.7a for the cases $n = 3, 4, 5$.) This is evidently a group of order $2n$, consisting of n rotations (through the n effectively distinct multiples of $360^\circ/n$) and n reflections. When n is odd, each of the n mirrors joins a vertex to the midpoint of the opposite side; when n is even, $\frac{1}{2}n$ mirrors join pairs of opposite vertices and bisect pairs of opposite sides[see Birkhoff and MacLane 1, pp. 117–118, 135].

The n rotations are just the operations of the cyclic group C_n.Thus the operations of D_n include all the operations of C_n:in technical language, C_n is a *subgroup* of D_n. The rotation through $360^\circ/n$, which generates the subgroup, may be described as the product $S = R_1R_2$ of reflections in two adjacent mirrors (such as OB and OC in Figure 2.7a) which are inclined at $180^\circ/n$.

Let $R_1, R_2, \ldots, R_n$ denote the n reflections in their natural order of arrangement. Then R_1R_{fc+1}, being the product of reflections in two mirrors inclined at k times 180°/m, is a rotation through k times $360^\circ/n$:

$$R_1R_{k+1} = S^k.$$

Thus $R_{k+1} = R_1S^k$, and the n reflections may be expressed as

$$R_1, R_1S, R_1S^2, \ldots, R_1S^{n-1}.$$

In other words, D_n is generated by R_1 and S. By substituting R_1R_2 for S, we see that the same group is equally well generated by R_1 and R_2, which satisfy the relations

2.7.1 2.71

$$R_1^2 = 1, R_2^2 = 1, (R_1R_2)^n = 1.$$

(The first two relations come from 2.51 and the third from 2.54.) These relations can be shown to suffice for an abstract definition[see Coxeter and Moser 1, pp. 6, 36].

A practical way to make a model of D_n is to join two ordinary mirrors by a hinge and stand them on the lines OB, OC of Figure 2.7a so that they are inclined at $180^\circ/n$. Any object placed between the mirrors yields $2n$ visible images (including the object itself). If the object is your right hand, n of the images will look like a left hand, illustrating the principle that, since a reflection reverses sense, the product of any even number of reflections preserves sense, and the product of any odd number of reflections reverses sense.

The first published account of this instrument seems to have been by Athanasius Kircher in 1646. The name *kaleidoscope* (from /caAos, beautiful; etS o s, a form; and *aKOTreiv,* to see) was coined by Sir David Brewster, who wrote a treatise on its theory and history. He complained[Brewster 1, *p.*147] that Kircher allowed the angle between the two mirrors to be any submultipl e of 360° instead of restricting it to submultiples of 180°.

The case when $n = 2$ is , of course, familiar. Standing between two perpendicular mirrors (as at a corner of a room), you see your image in each and also the image of the image, which is the way other people see you.

Having decided to use the symbol D_n for the dihedral group generated by reflections in two planes making a "dihedral " angle of $180^\circ/n$, we naturally stretch the notation so as to allow the extreme val ue $1n = 1$. Thus D_1 is the group of order 2 generated by a single reflection, that is, the symmetry group of the letter E or A, whereas the isomorphic group C_2, generated by a half-turn, is the symmetry group of the letter N.

According to Weyl[1, pp. 66,99], it was Leonardo da Vinci who discovered that the only finite groups of isometries in the plane are

$$C_1, C_2, C_3, \dots,$$
$$D_1, D_2, D_3, \dots,$$

His interest in them was from the standpoint of architectural plans. Of course, the prevalent groups in architecture have always been D_1 and D_2. But the pyramids of Egypt exhibit the group D_4, and Leonardo's suggestion has been followed to some extent in modem times: the Pentagon Building in Washington has the symmetry group D_5,and the Bahai Temple near Chicago has D_9. In nature, many flowers have dihedral symmetry groups such as D_6.The symmetry group of a snowflake is usually D_6 but occasionally only D_3.[Kepler 1, pp. 259–280.]

If you cut an apple the way most people cut an orange, the core is seen to have the symmetry group D_5. Extending the five-pointed star by straight cuts in each ha lf, you divide the whole apple into ten pieces from each of which the core can be removed in the form of two thin flakes.

EXERCISES

1. Describe the symmetry groups of (a) a scalene triangle, (b) an isosceles triangle, (c) a parabola, (d) a parallelogram, (e) a rhombus, (f) a rectangle, (g) an ellipse.
2. Use inverses and the associative law to prove algebraically the "cancellation rule" which says that the relation
$$RT = ST$$
implies R = S.
3. Show how the usual defining relations for $D3$, namely 2.71 with $n = 3$, may be deduced by algebraic manipulation from the simpler relations
$$R1^2 = 1, R_1 R_2 R1 = R2 R_1 R_2$$
4. The cyclic group C_m is a subgroup of C_n if and only if the number m is a divisor of n. In particular, if n is prime, the only subgroups of C_n are C_n itself and C_1.

2.8 Star Polygons

Instead of deriving the dihedral group D_n from the regular polygon $\{n\}$, we could have derived the polygon from the group: the vertices of the polygon are just the n images of a point P_0 (the C of Figure 2.7a) on one of the two mirrors of the kaleidoscope. In fact, there is no need to use the whole group D_n: its subgroup C_n will suffice. The vertex P_k of the polygon $P_0 P_1 \dots P_{n-1}$ can be derived from the initial vertexP_0 by a rotation through k times $360^\circ / n$.

More generally, rotations about a fixed point O through angles $\Theta, 2\Theta, 3\Theta \dots$ transform any point P_0(distinct from O) into other points $P_1, P_2, P_3, \dots$ on the circle with center O and radius OP_0. In general, these points become increasingly dense on the circle; but if the angle Θ is commensurable with a right angle, only a finite number of them will be distinct. In particular, if $\Theta = 360^\circ / n$, where n is a positive integer greater than 2, then there will be n points P_k whose successive joins

$$P_0 P_1, P_1 P_2, \dots, P_{n-1} P_0$$

are the sides of an ordinary regular n-gon.

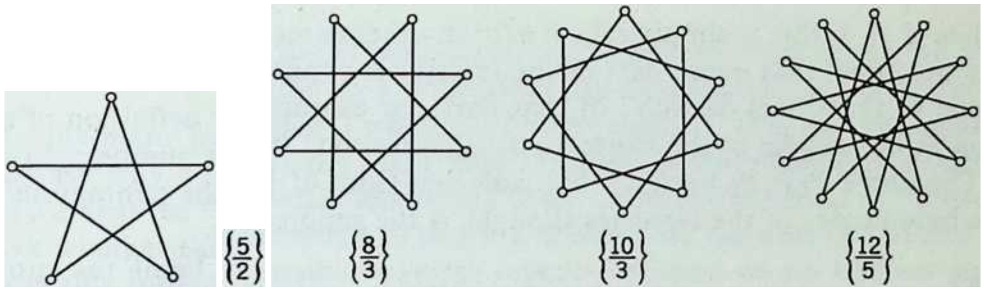


Figure 2.9.

Let us now extend this notion by allowing n to be any rational number greater than 2, say the fraction p/d (where p and d are coprime). Accordingly, we define a (generalized) *regular polygon* $\{n\}$, where $n = p/d$. Its p vertices are derived from P_0 by repeated rotations through $360^\circ/n$,and its p sides (enclosing the center d times) are

$$P_0P_1, P_1P_2, \dots, P_{p-1}P_0$$

Since a ray coming out from the center without passing through a vertex will cross d of the p sides, th is denominator d is called the *density* of the polygon [Coxeter 1, pp. 93–94]. When $d = 1$, so that $n = p$, we have the ordinary regular p-gon, $\{n\}$. When $d > 1$, the sides cross one another, but the crossing points are not counted as vertices. Since d may be any positive integer relatively prime to p and less than $\frac{1}{2}p$, there is a regular polygon $\{n\}$ for each rational number $n > 2$. In fact, it is occasionally desirable to include also the *digon*$\{2\}$, although its two sides coincide.

When $p = 5$, we have the pentagon $\{5\}$ of density 1 and the *pentagram*$\{\frac{5}{2}\}$ of density 2, which was used as a special symbol by the Babylonians and by the Py thagoreans. Similarly, the *octagram*$\{\frac{8}{3}\}$ and the *decagram* have density 3, while the *dodecagram*$\{\frac{10}{3}\}$ has density 5 (Fi gure 2.8a). These particular polygons have names as well as symbols because they occur as faces of interesting polyhedra and tessellations.[3]

Polygons for which$d > 1$are known as *star*polygons. They are frequently used in decoration. The earliest mathematical discussion of them was by Thomas Bradwardine (1290–1349), who became archbishop of Canterbury for the last month of his life. They were also studied by the great German scientist Kepler (1571–1630)[see

3 H. S. M. Coxeter , M. S. Longuet-Higgins, and J. C. P. Miller, Uniform polyhedra, *Philosophical Transactions of the Royal Society*, **A, 246** (1954), pp. 401–450.

Coxeter **1,**p. 114]. It was the Swiss mathematician L. Schlafli (1814–1895) who first used a numerical symbol such as $\{p/d\}$. This notation is justified by the occurrence of formulas that hold for $\{n\}$ equally well whether n be an integer or a fraction. For instance, any side of $\{n\}$ forms with the center O an isosceles triangle OP_0P_1 (Figure 2.86) who se angle at O is $2\pi/n$. (As we are introducing trigonometrical ideas, it is natural to use radian measure and write 2π instead of 360°.) The base of this isosceles triangle, being a side of the polygon, is conveniently denoted by $2l$. The other sides of the triangle are equal to the circumradius R of the polygon. The altitude or median from O is the in radius r of the polygon. Hence

2.8.1 2.81

$$R = l \csc \frac{\pi}{n}, r = l \cot \frac{\pi}{n}.$$

If $n = p/d$, the area of the polygon is naturally defined to be the sum of t he areas of the$p$isosc eles triangles, namely

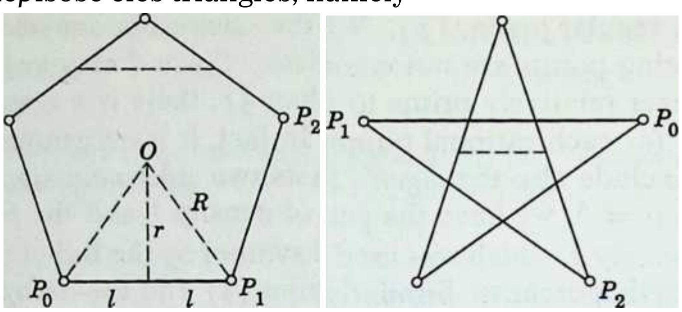


Figure 2.8b

2.8.2 2.82

$$plr = pl^2 \cot \frac{\pi}{n}.$$

When d =1, this is simply $pl^2 \cot \pi/p$; in other cases our definition of area has the effect that every part of the interior is counted a number of times equal to the "local density" of that part; for example, the pentagonal region in the middle of the pentagram $\{\frac{5}{2}\}$ is counted twice.

The angle $P_0P_1P_2$ between two adjacent sides of $\{n\}$, being the sum of the base angles of the isosceles triangle, is the supplement of $2\pi/n$, namely

2.8.3 2.83

$$\left(1 - \frac{2}{n}\right)\pi.$$

The line segment joining the midpoints of two adjacent sides is called the *vertex figure* of *{n}*. Its length i s clearly

2.8.4 2.84

$$2l\cos\frac{\pi}{n}$$

[Coxeter **1,** pp. 16, 94],

EXERCISES

1. If the sides of a polygon inscribed in a circle are all equal, the polygon is regular.
2. If a polygon inscribed in a circle has an odd number of vertices, and all its angles are equal, the polygon is regular. (Marcel Riesz.)
3. Find the angles of the polygons

$$\{5\}, \{\tfrac{5}{2}\}, \{9\}, \{\tfrac{9}{2}\}, \{\tfrac{9}{4}\}.$$

4. Find the radii and vertex figures of the polygons

$$\{8\}, \{\tfrac{8}{3}\}, \{12\}, \{\tfrac{12}{3}\}$$

5. Give polar coordinates for the kth vertex of a polygon $\{n\}$ of circumradius 1 with its center at the pole, taking P_0 to be $(1, 0)$.
6. Can a square cake be cut into nine slices so that everyone gets the same amount of cake and the same amount of icing?

3 Isometry in the Euclidean Plane

Having made some use of reflections, rotations, and translations, we naturally ask why a rotation or a translation can be achieved as a continuous displacement (or “motion”) while a reflection cannot. It is also reasonable to ask whether there is any other kind of isometry that resembles a reflection in this respect. After answering these questions in terms of “sense,” we shall use the information to prove a remarkable theorem (§3.6) and to describe the seven possible ways to repeat a pattern on an endless strip (§3.7).

3.1 Direct and Opposite Isometries

> ` `Take care of the sense, and the sounds will take care of themselves.''
>
> Lewis Carroll
>
> [Dodgson 1,Chap. 9]

By several applications of Axiom 1.26, it can be proved that any point P in the plane of two congruent triangles $ABC, A'B'C$ determines a corresponding point P' such that $AP = A'P', BP = B'P', CP = CP'$. Likewise, another point Q yields Q', and $PQ = P'Q'$. Hence

3.2 3.11

Any two congruent triangles are related by a unique isometry.

In § 1.3, we saw that Pappus’s proof of *Pons asinorum* involved the comparison of two coincident triangles ABC, ACB. We see intuitively that this is a distinction of *sense:* if one is counter-clockwise the other is clockwise. It is a “topological” property of the Euclidean plane that this distinction can be extended from coincident triangles to distinct triangles: any two “directed” triangles, ABC and $A'B'C'$, either agree or disagree in sense. (For a deeper investigation of this intuitive idea, see Veblen and Young **[2,** pp. 61–62] or Denk and Hofmann **[1,** p. 56].)

If ABC and $A'B'C'$ are congruent, the isometry that relates them is said to be *direct or opposite* according as it preserves or reverses sense, that is, according as ABC and $A'B'C'$ agree or disagree. It is easily seen that this property of the isometry is independent of the chosen triangle *ABC:* if the same isometry relates DEF to $D'E'F'$, where DEF agrees with ABC, then also $D'E'F'$ agrees with $A'B'C'$. Clearly, direct and opposite isometries combine like positive and negative numbers (e.g., the product of two opposite isometries is direct). Since a reflection is opposite, a rotation (which is the product of two reflections) is direct. In particular, the identity is direct. Some authors call direct and opposite isometries "displacements and reversals" or "proper and improper congruences."

Theorem 2.31 can be extended as follows:

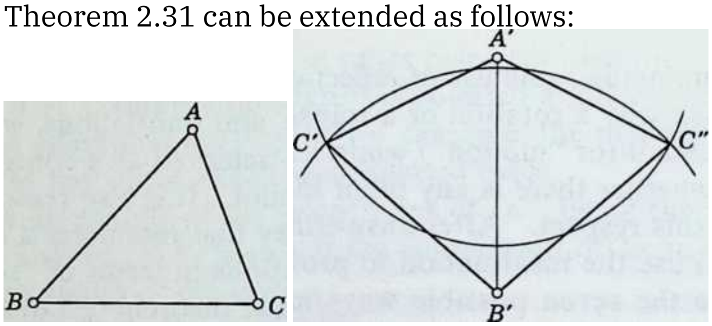


Figure 3.1a

12. *Two given congruent line segments (or point pairs) AB, A'B' are related by just two isometries: one direct and one opposite.*

To prove this, take any point C outside the line AB, and construct C'' so that the triangle $A'B'C''$ is congruent to ABC. The two possible positions of C (marked C', C'' in Figure 3. la) provide the two isometries. Since either can be derived from the other by reflecting in $A'B'$, one of the isometries is direct and the other opposite.

For a complete discussion we need the following theorem [Bachmann **1,** P-3]:

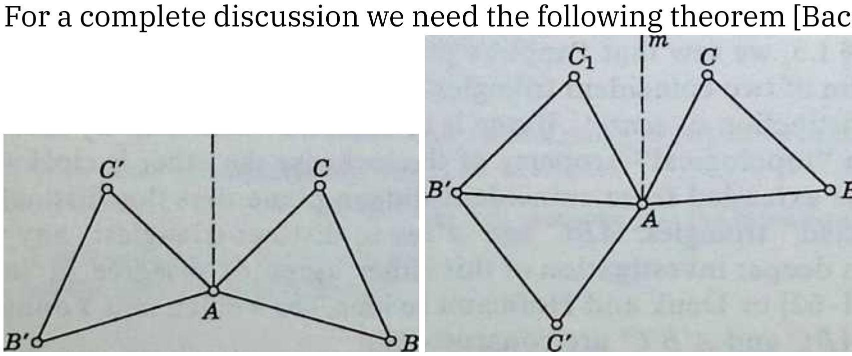


Figure 3.1b

13. *Every isometry of the plane is the product of at most three reflections. If there is an invariant point, "three" can be replaced by "two."*

We prove this in four stages, using 3.11. Trivially, if the triangles *ABC, A′B′C″* coincide, the isometry is the identity (which is the product of a reflection with itself). If *A* coincides with *A′,* and *B* with *B′,* while *C* and *C″* are distinct, the triangles are related by the reflection in *AB*. The case when only *A* coincides with *A′* can be reduced to one of the previous cases by reflecting *ABC* in *m,* the perpendicular bisector of *BB′* (see Figure 3.16). Finally, the general case can be reduced to one of the first three cases by reflecting *ABC* in the perpendicular bisector of *AA′* [Coxeter **1,** p. 35].

Since a reflection reverses sense, an isometry is direct or opposite according as it is the product of an even or odd number of reflections.

Since the identity is the product of two reflections (namely of any reflection with itself), we may say simply that any isometry is the product of *two* or *three* reflections, according as it is direct or opposite. In particular,

14. *Any isometry with an invariant point is a rotation or a reflection according as it is direct or opposite.*

Exercises

1. Name two direct isometries.
2. Name one opposite isometry. Is there any other kind?
3. If *AB* and *A'B'* are related by a rotation, how can the center of rotation be constructed? *{Hint:* The perpendicular bisectors of *AA'* and *PB′* are not necessarily distinct.)
4. The product of reflections in three lines through a point is the reflection in another line through the same point [Bachmann 1, p. 5].

3.3 TRANSLATION

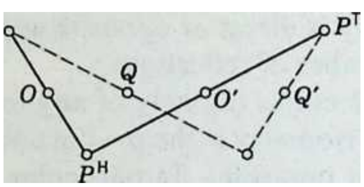
Enoch walked with God; and he was not, for God took him.

Genesis V, 24

The particular isometries so far considered, namely reflections (which are opposite) and rotations (which are direct), have each at least one invariant point. A familiar isometry that leaves no point invariant is a *translation* [Bachmann 1, p. 7], which may be described as the product of half-turns about two distinct points *O, O′* (Figure 3.2a). The first half-turn transforms an arbitrary point P into P^H, and the second transforms this into *Pf* with the final result that PP^T *is parallel to OO′ and twice as long.* Thus the length and direction of PP^T are constant: independent of the position of P. Since a translation is completely determined by its length and direction, the product of half-turns about O and O' is the same as the product of half-turns about Q and Q', provided QQ' is equal and parallel to *OO′.* (This means that $OO'Q'Q$ is a parallelogram, possibly collapsing to form four collinear points, as in Figure 3.2a.) Thus, for a given translation, the center of one of the two half-turns may be arbitrarily assigned.

21. *The product of two translations is a translation.*

For, we may arrange the centers so that the first translation is the product of half-turns about *Oi* and *O2,* while the second is the product of half-turns about O^2 and O^3. When they are combined, the two half-turns about *O2* cancel, and we are left with the product of half-turns about (9i and O^3.



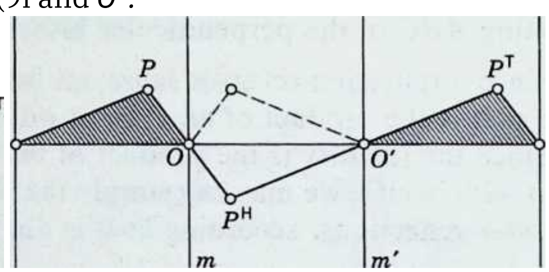


Figure 3.2a Figure 3.2b

Similarly, if *m* and *m′* (Figure 3.26) are the lines through O and O' perpendicular to *OO′,* the half-turns about O and O' are the products of reflections in *m* and *OO′, OO′* and *m′.* When they are combined, the two reflections in OO' cancel, and we are left with the product of reflections in *m* and *m′.* Hence

22. *The product of reflections in two parallel mirrors is a translation through twice the distance between the mirrors.*

If a translation T takes P to P^T and Q to *Qf,* the segment QQ^T is equal and parallel to PP^T; therefore $PQQfP^T$ is a parallelogram. Similarly, if another translation U takes P to *Q,* it also takes P^T to 2^T; therefore

$$TU = UT.$$

(In detail, if Q is P^u, *Qf* is P^{UT}. But U takes P^T to *P*™. Therefore P^{TU} and P^{UT} coincide, for all positions of P.) In other words,

23. *Translations are commutative.*

The product of a half-turn H and a translation T is another half-turn; for we can express the translation as the product of two half-turns, one of which is H, say T = HH′, and then we have

$$HT = H^2H' = H' :$$

24. *The product of a half-turn and a translation is a half-turn.*

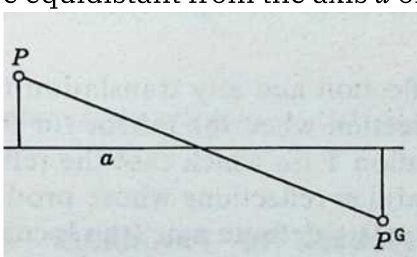
Exercises

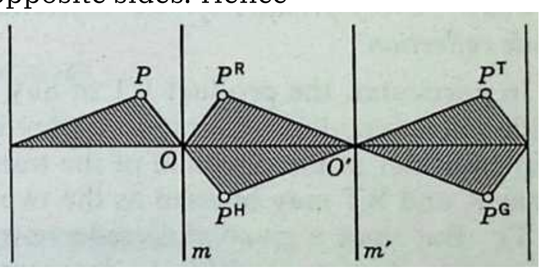
1. If T is the product of half-turns about O and *O′,* what is the product of half-turns about *O′* and *07*
2. When a translation is expressed as the product of two reflections, to what extent can one of the two mirrors be arbitrarily assigned?
3. What is the product of rotations through opposite angles (a and—a) about two distinct points?
4. The product of reflections in three parallel lines is the reflection in another line belonging to the same pencil of parallels.
5. Every product of three half-turns is a half-turn [Bachmann 1, p. 7],
6. If Hi, H^2, H^3 are half-turns, $HiH^2H^3 = H^3H^2Hi$.
7. Express the translation through distance *a* along the x-axis as a transformation of Cartesian coordinates. If/(x, y) = 0 is the equation for a curve, what is the equation for the transformed curve? Consider, for instance, the circle $x^2 + 7^2 - 1 = 0$.

3. GLIDE REFLECTION

We are now familiar with three kinds of isometry: reflection, rotation, and translation. We have not yet considered the product of the reflections in the sides of a triangle. We shall find that this is a *glide reflection:* the product of the reflection in a fine a and a translation along the same line. Clearly, a glide reflection is determined by its *axis a* and the extent of the component translation. Since a reflection is opposite whereas a translation is direct, their product is opposite. Thus a glide reflection is an opposite isometry having no invariant point [Coxeter 1, p. 36],

If a glide reflection G transforms an arbitrary point P into P^G (Figure 3.3a), P and P^G are equidistant from the axis a on opposite sides. Hence





The midpoint of the line segment PP^G lies on the axis for all positions of P.

Figure 3.3a

Figure 3.3b

Let Ri and T denote the component reflection and translation. They evidently commute, so that

$$G = R^X T = T R^X.$$

We have seen (Figure 3.26) that the translation T may be expressed as the product of two half-turns or of two parallel reflections. Identifying the line a in Figure 3.3a with the line *OO′* in Figure 3.3d, let R, R′denote the reflections in *m, m′.* Then the product of the two half-turns

$H = RRj = RiR, H' = R'Ri = R^X R'$

is $T = HH' = RRxRiR' = RR'$,

and the glide reflection is

$G = RiT = RiRR' = HR'$

$= TRi = RR'Ri = RH'$.

Thus a glide reflection may be expressed as the product of three reflections (two perpendicular to the third), or of a half-turn and a reflection, or of a reflection and a half-turn. Conversely, the product of any half-turn and any reflection (or vice versa) is a glide reflection, provided the center of the half-turn does not lie on the mirror. [Bachmann 1, p. 6.]

We saw in 3.13 that any direct isometry in the plane is the product of two reflections, that is, a translation or a rotation according as the two mirrors are parallel or intersecting; also that any opposite isometry with an invariant point is a reflection. To complete the catalog of isometries, the only remaining possibility is an opposite isometry with no invariant point. If such an isometry S transforms an arbitrary point A into Aconsider the half-turn H that interchanges these two points. The product HS, being an opposite isometry which leaves the point A' invariant, can only be a reflection R. Hence the given opposite isometry is the glide reflection

S = H-iR = HR:

Every opposite isometry with no invariant point is a glide reflection.

In other words,

3.31 *Every product of three reflections is either a single reflection or a glide reflection.*

In particular, the product RT of any reflection and any translation is a glide reflection, degenerating to a pure reflection when the mirror for R is perpendicular to the direction of the translation T (in which case the reflections R and RT may be used as the two parallel reflections whose product is T). But since a given glide reflection G has a definite axis (the locus of midpoints of segments PP^a), its decomposition into a reflection and a translation *along the mirror* is unique (unlike its decomposition into a reflection and a half-turn, where we may either take the mirror to be any line perpendicular to the axis or equivalently take the center of the half-turn to be any point on the axis).

Exercises

1. If B is the midpoint of $A C$, what kinds of isometry will transform

(i) AB into CB, (ii) AB into BC?

2. Every direct isometry is the product of two reflections. Every opposite isometry is the product of a reflection and a half-turn.

3. Describe the product of the reflection in OO' and the half-turn about O.

4. Describe the product of two glide reflections whose axes are perpendicular.

5. Every product of three glide reflections is a glide reflection.

6. The product of three reflections is a reflection if and only if the three mirrors are either concurrent or parallel.

7. If R_1, R_2, R3 are three reflections, $(R_1R_2R_3)^2$ is a translation [Rademacher and Toeplitz 1, p. 29].

8. Describe the transformation

(x,y) >(x + *a,—y).*
Justify the statement that this transforms the curve*f(x,y)* = 0 into *f(x—a,—y)* = 0.

4. REFLECTIONS AND HALF-TURNS

Thomsen* has developed a very beautiful theory in which geometrical properties of points $O, Oi, O^2, \ldots$ and lines $m, mi, m^2, \ldots$ (understood to be all distinct) are expressed as relations among the corresponding half-turns $H, Hi, H^2, \ldots$ and reflections $R, Ri, R_2, \ldots$. The reader can soon convince himself that the following pairs of statements are logically equivalent:
<—> *m* and *mi* are perpendicular.
$RRiHR = R^XR = RH$
O lies on *m.*
$RiR_2R^3 = R^3R_2R^X$ < - - - > mi, m^2, m^3
are either concurrent or parallel.

O is the midpoint of $Oi < 9^2$.

$HiH = HH^2$
$HiR = RH^2$
m is the perpendicular bisector of OiO^2.

EXERCISE

Interpret the relations (a) $H1H2H3H4 = 1; (b) R^X R = RR_2$.

5. SUMMARY OF RESULTS ON ISOMETRIES

And thick and fast they came at last. And more, and more, and more.

Lewis Carroll

[Dodgson **2,** Chap. 4]

Some readers may have become confused with the abundance of technical terms, many of which are familiar words to which unusually precise meanings have been attached. Accordingly, let us repeat some of the definitions, stressing both their analogies and their differences.

* G. Thomsen, The treatment of elementary geometry by a group-calculus, *Mathematical Gazette,* **17** (1933), p. 232. Bachmann [1] devotes a whole book to the development of this idea.

In all the contexts that concern us here, a *transformation* is a one-to-one correspondence of the whole plane (or space) with itself. An *isometry* is a special kind of transformation, namely, the kind that preserves length. A *symmetry operation* belongs to a given figure rather than to the whole plane: it is an isometry that transforms the figure into itself.

In the plane, a *direct* (sense-preserving) isometry, being the product of two reflections, is a rotation or a translation according as it does or does not have an invariant point, that is, according as the two mirrors are intersecting or parallel. In the latter case the length of the translation is twice the distance between the mirrors; in the former, the angle of the rotation is twice the angle between the mirrors. In particular, the product of reflections in two perpendicular mirrors is a half-turn, that is, a rotation through two right angles. Moreover, the product of two half-turns is a translation.

An *opposite* (sense-reversing) isometry, being the product of three reflections, is, in general, a *glide reflection-,* the product of a reflection and a translation. In the special case when the translation is the identity (i.e., a translation through zero distance), the glide reflection reduces to a single reflection, which has a whole line of invariant points, namely, all the points on the mirror.

To sum up:

3.51 *Any direct isometry is either a translation or a rotation. Any opposite isometry is either a reflection or a glide reflection.*

Exercises

1. If S is an opposite isometry, S^2 is a translation.

2. If Ri, R_2, R3 are three reflections, $(R2R3R_1R_2R3)^2$ is a translation along the first mirror. *(Hint:* Since $R^X R_2 R3$ and $R2R3R1$ are glide reflections, their squares are commutative, by 3.23; thus

$$(R^1 R_2 R^3)^2 (R_2 R3R1)^2 = (R_2 R3R1)^2 (R_1 R_2 R3)^2,$$

that is, Ri and $(R2R3R_1R_2R3)^2$ are commutative [cf. Bachmann 1, p. 13].)

6. HJELMSLEV'S THEOREM

> ...a very high degree of unexpectedness, combined with inevitability and economy.

G. H. Hardy **[2,** p. 53]

We saw, in 3.12, that two congruent line segments *AB, A'B',* are related by just two isometries: one direct and one opposite. Both isometries have the same effect on every point collinear with *A* and *B,* that is, every point on the infinite straight line *AB* (for instance, the midpoint of *AB* is transformed into the midpoint of *A'B').* The opposite isometry is a reflection or glide reflection whose mirror or axis contains all the midpoints of segments joining pairs of corresponding points. If two of these midpoints coincide, the

direct isometry is a half-turn, and they all coincide [Coxeter 3, p. 267]. Hence

HJELMSLEVS THEOREM. *When all the points P on one line are related by an isometry to all the points P' on another, the midpoints of the segments PP' are distinct and collinear or else they all coincide.*

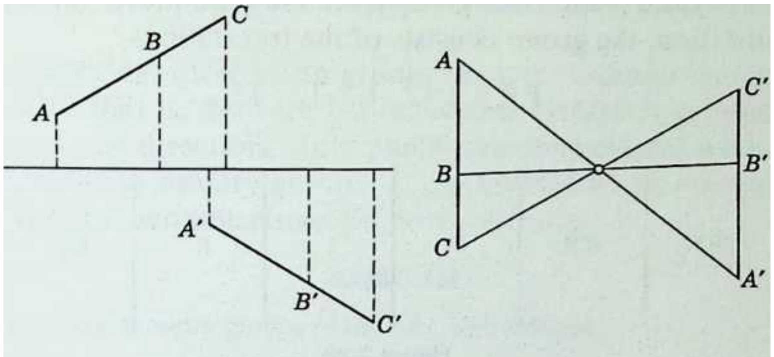


Figure 3.6a

In particular, if *A, B, C* are on one line and *A′, B′, C″* on another, with

3.61 *AB = A′B′, BC = B′C* (Figure 3.6a), then the midpoints of *AA′, BB′, CC″* are either collinear or coincident (J. T. Hjelmslev, 1873–1950).

7. PATTERNS ON A STRIP

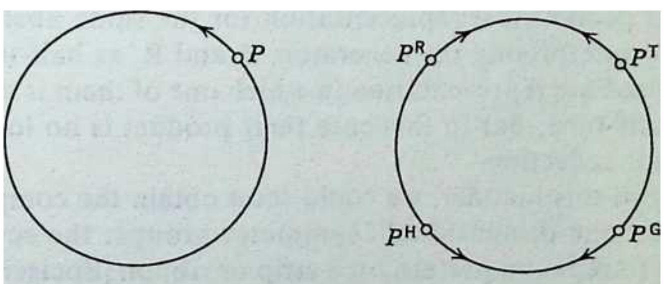


Figure 3.7a

Any kind of isometry may be used to relate two equal circles. For instance, the point P on the first circle of Figure 3.7*a* is transformed into P^T on the second circle by a translation, into P^R by a reflection, into P^H by a half-turn, and into P^G by a glide reflection. (Arrows have been inserted to indicate what happens to the positive sense of rotation round the first circle.) These four isometries have one important property in common: they leave invariant (as a whole) one infinite straight line, namely, the line joining the centers of the two circles. (In the fourth case this is the *only* invariant line.)

We have seen (Figure 3.26) that the product of reflections in two parallel mirrors *m, m′* is a translation. This may be regarded as the limiting case of a rotation whose center is very far away; for the two parallel mirrors are the limiting case of two mirrors intersecting at a very small angle. Accordingly, the infinite group generated by a single translation is denoted by C^x, and the infinite group generated by two parallel reflections is denoted by D^x. Abstractly, C^x is the "free group with one generator." If T is the generating translation, the group consists of the translations

$\dots, T-2, T-i, 1, T, T2, \dots.$

m m′

$RR'RR'RR1R'RR'$

Figure 3.7b

Similarly, D^x, generated by the reflections R, R' in parallel mirrors *m, m′* (Figure 3.76), consists of the reflections and translations

$\dots, RR'R, R'R, R, 1, R', RR', R'RR', \dots$

[Coxeter 1, p. 76]; its abstract definition is simply

$R2 = R'2^{=}i.$

This group can be observed when we sit in a barber's chair between two parallel mirrors (cf. the *New Yorker,* Feb. 23, 1957, p. 39, where somehow the reflection RR′RR′R yields a demon).

A different geometrical representation for the same abstract group $D^x$is obtained by interpreting the generators R and R′as half-turns. There is also an intermediate representation in which one of them is a reflection and the other a half-turn; but in this case their product is no longer a translation but a glide reflection.

Continuing in this manner, we could soon obtain the complete list of the seven infinite "one-dimensional" symmetry groups: the seven essentially distinct ways to repeat a pattern on a strip or ribbon [Speiser **1,** pp. 81–82]:

Typical pattern	Generators	*Abstract Group*
(i)...L L L L...	1 translation	
(ii)...L T L r...	1 glide reflection	r '-'oo
(iii)...V V V V...	2 reflections	
(iv)...N N N N...	2 half-turns	
(v)...V A V A...	1 reflection and 1 half-turn	>

(vi).. .D D D D...	1 translation and 1 reflection	c^{x} x Dr
(vii).. .H H H H...	3 reflections	D X Dr

In (iii), the two mirrors are both vertical, one in the middle of a V, reflecting it into itself, while the other reflects this V into one of its neighbors; thus one half of the V, placed between the two mirrors, yields the whole pattern. In (vi) and (vii) there is a horizontal mirror, and the symbols in the last column indicate "direct products" [Coxeter 1, p. 42]. For all these groups, except (i) and (ii), there is some freedom in choosing the generators; for example, in (iii) or (iv) one of the two generators could be replaced by a translation.

Strictly speaking, these seven groups are not "1-dimensional" but "Ildimensional;" that is, they are 2-dimensional symmetry groups involving translation in one direction. In a purely one-dimensional world there are only two infinite symmetry groups: C^{x}, generated by one translation, and D^{x}, generated by two reflections (in point mirrors).

Exercises

1. Identify the symmetry groups of the following patterns:

   ...b b b b...,

   ...b p b p...,

   ...b d b d...,

   ...b q b q...,

...bdpqbdpq....

2. Which are the symmetry groups of *(a)* a cycloid, *(b)* a sine curve?

Index

Todo list