LONGMANS' MODERN MATHEMATICAL SERIES

General Editors

P. ABBOTT, B.A., C. S. JACKSON, M.A.
F. S. MACAULAY, M.A., D.Sc.

RULER AND COMPASSES

Longmans' Modern Mathematical Series

RULER & COMPASSES

BY

HILDA P. HUDSON

M.A., Sc.D.

WITH DIAGRAMS

LONGMANS, GREEN AND CO.

39 PATERNOSTER ROW, LONDON

FOURTH AVENUE & 30TH STREET, NEW YORK

BOMBAY, CALCUTTA, AND MADRAS

1916

CONTENTS

REFERENCES.

Enriques, Fragen der Elementargeometrie. Deutsche Aufgabe von Fleischer. II Teil. Leipzig, 1907.

Vahlen, Konstruktionen und Approximationen. Leipzig, 1911.

Klein, Ausgewählte Fragen der Elementargeometrie. Göttingen, 1895.

Hilbert, Grundlagen der Geometrie. Leipzig, 2te Aufl., 1903.

Richmond, To construct a regular polygon of 17 sides. Math. Ann. LXVII. p. 459, 1909.

Gérard, Construction du polygone régulier de 17 côtés au moyen du seul compas. Math. Ann. XLVIII. p. 390, 1896.

Mathews, Projective geometry. London, 1914.

Zühlke, Ausführung elementargeometrischer Konstruktionen bei ungünstigen Lageverhältnissen. Leipzig, 1906.

Lemoine, Géométrographie ou art des constructions géométriques. Paris, 1902.

Poncelet, Traité des propriétés projectives des figures. Paris, 1822.

Steiner, Die geometrischen Konstructionen ausgeführt mittlest der geraden Linie und Eines festen Kreises. Berlin, 1833.

Mascheroni, Geometria del Compasso. Pavia, 1797.

Adler, Zur Theorie der Mascheronischen Constructionen. Wiener Sitzungsber., 1890.

Hobson, On geometrical constructions by means of the compass. Math. Gazette, March 1913.

These works will be cited by the name of the author only. Valuable help has also been given by Mr. C. S. Jackson and Dr. F. S. Macaulay, to whom the writer offers her thanks.

CHAPTER I.

INTRODUCTION.

AT the beginning of his *Elements*, Euclid places his three Postulates: "Let it be granted

(i) that a straight line may be drawn from any one point to any other point;

(ii) that a terminated straight line may be produced to any length in a straight line;

(iii) that a circle may be described from any centre, at any distance from that centre";

and all the constructions used in the first six books are built up from these three operations only. The first two tell us what Euclid could do with his ruler or straight edge. It can have had no graduations, for he does not use it to carry a distance from one position to another, but only to draw straight lines and produce them. The first postulate gives us that part of the straight line AB which lies between the given points A, B; and the second gives us the parts lying beyond A and beyond B; so that together they give the power to draw the whole of the straight line which is determined by the two given points, or rather as much of it as may be required for any problem in hand.

The last postulate tells us what Euclid could do with his compasses. Again, he does not use them to carry distance, except from one radius to another of the same circle; his instrument, whatever it was, must have collapsed in some way as soon as the centre was shifted, or either point left the plane (see p. 70). The three postulates then amount to granting the use of ruler and compasses, in order to draw a straight line through two given points, and to describe a circle with a given centre to pass through a given point ; and these two operations carry us through all the plane constructions of the *Elements*. The term *Euclidean construction* is used of any construction, whether contained in his works or not, which can be carried out with Euclid's two operations repeated any finite number of times.

In fact, Euclid gives only very few of the constructions which can be carried out with ruler and compasses, and probably every student of geometry has at some time or other constructed a figure which no one else had ever made before. But from very early times there were certain figures which everyone tried to make with ruler and compasses, and no one succeeded. The most famous of these baffling figures are the square equal in area to a given circle, and the angle equal to the third part of a given angle ; and it has at last been proved that neither of these can possibly be constructed by a finite series of Euclidean operations.

The set of figures which it is possible to construct with ruler and compasses is thus on the one hand infinite, and on the other hand limited. It is easy to see that it is infinite : even if we consider only the very simple type of figure consisting of a set of points at equal distances on a straight line, which can certainly be constructed with ruler and compasses, the figure may contain either three or four or a greater number of points without any upper limit, so that there are an infinite number of figures even of this one simple type ; much more is the whole set of possible figures infinite. And yet the set is limited, for many figures can be thought of which do not belong to it, and require apparatus other than ruler and compasses for their construction ; besides those mentioned above, there are for example the regular heptagon and nonagon ; or the ellipse, which can be drawn as a continuous curve with the help of two pins and a thread, but of which we can only obtain an unlimited number of separate points by Euclidean constructions.

The first question treated in this book is the one which naturally arises here : what constructions can be built upon Euclid's postulates, and what cannot ? or, in other words, what problems can be solved by ruler and compasses only ? For centuries, vain attempts were made to square the circle and trisect an angle by Euclidean constructions, and these attempts were often of use in other ways, though their immediate object failed; but it was only through the growth of analysis that it was proved once for all that they must fail. The ancient or classical geometry lends itself curiously little to any general treatment; and even modern geometry lacks a notation or calculus by which to examine its own powers and limitations. Occasionally we can be sure beforehand that a certain class of problem will yield to a certain type of method, but as a rule each problem has to be taken on its own merits and a separate method invented for it. There can be endless variety in the methods, and it might seem a hopeless task to sum them up, and impossible to say of any one problem that no ruler and compass construction for it ever will be found by some ingenious person yet unborn. Yet this is just what can be said in the case of the circle-squarers, and the final word came not from a geometer but from an analyst. The methods of coordinate geometry allow us to translate any geometrical statement into the language of algebra, and though this language is less elegant, it has a larger vocabulary; it can discuss problems in general as well as in particular, and it can give us the complete answers to the questions: what constructions are possible with ruler only, or with ruler and compasses?

In the next chapter we shall show how each step of a ruler and compass construction is equivalent to a certain analytical process; it is found that the power to use a ruler corresponds exactly to the power to solve linear equations, and the power to use compasses to the power to solve quadratics. For this reason, problems that can be solved with ruler only are called *linear problems*, and those that can be solved with ruler and compasses are called *quadratic problems*. Since each step of a ruler and compass construction is equivalent to the solution of an equation of the first or second degree, we consider what these algebraic processes can lead to, when combined in every possible way, and that enables us to answer the question before us and say (p. 19) that those problems and those problems alone can be solved by ruler only, which can be made to depend on a linear equation, whose root can be calculated by carrying out rational operations only; and (pp. 24, 30) that .those problems and those problems alone can be solved by ruler and compasses, which can be made to depend on an algebraic equation, whose degree must be a power of 2, and whose roots can be calculated by carrying out rational operations together with the extraction of square roots only.

This is a complete answer to our question, but it is stated entirely in algebraic language, and in the general form in which it stands it cannot be translated into the language of pure geometry, for the words are lacking. But they are not really needed; for if a definite problem is before us, stated geometrically, we can always apply the test to the algebraic equivalent of the problem, though the application in some particular cases may not be easy; and if the test is satisfied, we can from the analysis deduce a Euclidean construction for the geometrical problem. As an example, we shall consider, at the end of chapter II, what regular polygons are within our powers of construction. It is a curious and unexpected result that the regular polygon of 17 sides is included, and a construction for this is given on p. 34.

When we have agreed that the set of possible Euclidean constructions is both infinite and limited, and when we have found out in some measure what its limitations are, it is natural to seek a clearer view of the set, and to ask what are the best ways of classification. The first main subdivision has already been brought to our notice; it consists of linear problems, which can be solved with ruler only, and chapter III is devoted to these. We try to show how the data of a problem control its construction, and how the properties and relations of the data fall into distinct classes, each of which allows a particular set of constructions to be carried out.

Now though Euclid's compasses can to some extent carry distances, we are not making use of compasses in the section referred to; and Euclid's ruler cannot carry distances at all. So we find that in general, that is, if the data of the problem do not have some exceptional relations, we are not able to carry a distance from one part of the figure to another, nor to compare the lengths of two segments, even of the same straight line, unless one is a part of the other, and then all that we can say is that the whole is greater than the part. We can never say that two segments are equal in length unless they coincide; for the very idea of comparing the lengths of two segments which have different positions, involves the idea of making them lie alongside of one another in order that we may compare them, and therefore of carrying one at least into another position; and this we cannot in general do with ruler only. Now the property of a parallelogram, that its opposite sides are equal in length, shows that if we can draw parallels, we can carry a distance from a straight line to a parallel straight line, so that there is a close connection between carrying distances and drawing parallels; and we shall show that, to a certain extent, if we can do either we can do both. But even so, it is only the parallel sides of a parallelogram that are equal, and not the adjacent sides, and drawing parallels does not help us to transfer the length of a segment into any direction other than that of the original segment. It is only in the special case in which the data allow us to construct a rhombus, whose adjacent sides are equal as well as its parallel sides, that we can obtain equal lengths on any of two different sets of parallel straight lines. We could hardly expect, with ruler only, to be able to turn a given length into any direction, for this is one of the chief uses of a pair of compasses.

In this way we get the usual classification of linear constructions according to the projective and metrical properties of the data, properties of length and properties of angle. The idea of cross-ratio is fundamental to them all, and when we have to compare the cross-ratios of two different ranges, we are led to the theories of homography and involution; but in connection with these we come upon several problems that require the common or double points; and the construction of these points is equivalent to the solution of a quadratic equation, and therefore impossible with ruler only.

In chapter IV therefore, in which we admit the use of compasses, the comparison is worked out between describing a circle and solving a quadratic, and then we are able to carry the treatment of cross-ratio and involution to a more satisfactory stage. A digression is introduced at this point (p. 70) to show that the ordinary modern instruments, dividers, parallel ruler and set-square, that are commonly used along with ruler and compasses, amount to short cuts in Euclidean constructions without extending the range of soluble problems. We also consider which ways of using these other instruments can completely replace the use of compasses.

But instead of classifying the data of a construction, we may classify the methods, and ask afterwards in what sort of problem each method is likely to be useful. Two fundamental ideas are put forward in chapter V, one or other or both of which are prominent in very many constructions: these are separation of properties and transformation. They give rise to half a dozen fairly well-marked lines of attack which are illustrated in that chapter. There is the method of loci (p. 78), when some of the conditions to be fulfilled convince us that a required point must lie on a certain locus, which must be made up of straight lines and circles if the method is to succeed, and the other conditions convince us that the same point must also lie on another such locus, so that it can only be a point of intersection of the two loci. There is the method of trial and error (p. 80), when from a finite number of unsuccessful attempts at a construction we are able to discover the way to begin which is bound to lead to success, if a solution exists; the method of projection (p. 90), and several of its particular cases, in which the two principles of separation of properties and of transformation are both present; the method of inversion (p. 92), which is very appropriate to ruler and compass constructions because of the way in which it relates circles and straight lines; and the method of reciprocation (p. 98), which rests upon the principle of duality.

Besides these intrinsic ways of classification, which come from considering the constructions themselves, there are others that come from considering them from some quite external point of view, according to how far they avoid certain draw-backs which are practical rather than mathematical, and arise from faults in our actual apparatus. Yet there is some theoretical interest also in devising constructions that shall as far as possible get over the difficulties of a small sheet of paper and a blunt pencil (chapter VI). The idea of the last section of the same chapter is to make a numerical estimate of the length of a construction, by reckoning up all the different operations with ruler and compasses that it requires, so as to be able to say which is the shortest of different solutions of the same problem. This plan of "giving marks" is little more than a pastime, and the scale of marking is very arbitrary; but Lemoine's book on Geometrography deserves to be better known, and some account of the matter is given here in the hope of introducing more English readers to his original work.

The subjects of the last two chapters are also now mere curiosities, though one at least arose as a practical point in machine construction. Any Euclidean problem can be solved by drawing only one circle and the requisite number of straight lines, usually a large number; or else by drawing the requisite number of circles and no straight line at all. This is proved, and examples of the methods are given, in chapters VII and VIII.

Thus the connecting link throughout the book is the idea of the whole set of ruler and compass constructions, its extent, its limitations and its divisions. But the matter of the following pages consists largely of examples, which have been brought in wherever possible, and it is hoped that those readers who are not attracted by general or analytical discussions, may yet find something to interest them among these problems and their geometrical solutions.

CHAPTER II.

POSSIBLE CONSTRUCTIONS.

BEFORE beginning to treat of any particular ruler and compass constructions, we shall discuss in this chapter the whole set of such constructions; and for this we need the help of analysis. It is only by finding the equivalent analytical process for each geometrical step, that we are able to tell what constructions can or cannot be carried out by using certain definite instruments in certain definite ways.

Coordinates.

We must first see how to express the data of a problem in arithmetical form, by means of some system of co-ordinates, and then see how each operation with ruler or compasses combines the numerical data in a certain way. The laws of algebra then show exactly what results can be obtained with either instrument or both.

It will simplify matters to suppose that all the geometrical data are points. This is convenient, and it is perfectly general. By Euclid's first two postulates, any straight line which is given ready drawn could have been constructed if, instead, two points lying on it had been given; so if there is a straight line among the data we can discard it, provided we take instead, as data, two points lying on the straight line. These must be definite points, not just "any" points. None of Euclid's postulates enable us to take any point on a straight line (see p. 13); a point must be either given, or constructed as the intersection of two lines. So unless two definite points on the straight line are among the other data, before we discard the straight line we must first intersect it by two other given straight lines, or by the joins of two pairs of given points; then we can take these two points of intersection as data, instead of the given straight line. This assumes that there are at least two straight lines or three points not in a straight line among the other data, or directly obtainable from the data; there are some trivial cases in which so few elements are given, that only a finite number of points and lines can be constructed from them (see p. 19). For example, if all that is given is one pair of straight lines, we obtain their point of intersection and nothing more; it is the only point which is completely determined. All the other points on the given lines are partially determined, but not completely; they form a definite class of points, distinguished from all the other points of the plane but not from one another, and therefore not sufficiently determined to be ready for use in further constructions. Such cases are excluded from the present discussion. In the same way, a circle can be replaced either by the centre and one point on the circumference, or by the centre and two other points whose distance apart is equal to the radius, or by three points on the circumference; again we have to exclude a few trivial cases. Also, any of the problems we are considering has for its object the construction of points, straight lines and circles having certain required properties; and it can be considered as solved when we have obtained the required points and also points from which the required straight lines and circles can be immediately constructed; so that not only what is given but also what is required may be taken to be a set of points.

We begin by referring all the points, given or required, to some axes of coordinates or other frame of reference; then every point has a set of coordinates, and every set of coordinates determines a point. These coordinates might be lengths, or areas, or angles, according to the system chosen; but we shall assume that they are pure numbers, the ratios of the geometrical quantities which determine the point to the corresponding units of the same kind. In a Cartesian system then, the coordinates are ratios of lengths, namely, of the abscissa to the unit of length on the axis of x, and of the ordinate to the unit of length on the axis of y; the two units may be the same or different. We shall use wherever possible an oblique Cartesian system, and we may take its frame of reference to consist of two axes together with a point on each at the corresponding unit distance from the origin. Later in the chapter we shall replace this by a more general system of projective coordinates (p. 19), for which the frame of reference consists of four points. In both systems, a point has two unique coordinates, and a straight line is represented by a linear equation between them.

Thus when a set of points is given, their coordinates are also given ; the whole set of data in any problem can be replaced by a set of numbers, and in the same way what is required can also be taken to be a set of numbers, the coordinates of the points which determine the points, straight lines and circles which are to be constructed. The question before us takes the form : what is the relation of this second set of numbers to the first, if the construction is one which can be carried out with ruler, or compasses, or both ?

I. Ruler Alone.

If the ruler alone is used, the only lines which can be drawn are the straight lines joining pairs of points already given or obtained ; and the only way of obtaining a new point is as the intersection of two such joins.

Let x_1, y_1 be the Cartesian coordinates of P_1, the point of intersection of the straight lines joining $A_1(a_1, b_1)$, $A_2(a_2, b_2)$ and A_3, A_4 respectively ; then x_1, y_1 satisfy both the equations of these two straight lines. The first of these equations is

$$\begin{vmatrix} x, & y, & 1 \\ a_1, & b_1, & 1 \\ a_2, & b_2, & 1 \end{vmatrix} = 0 ;$$

for it is formed from the general equation of a straight line,

$$lx + my + 1 = 0,$$

by putting for the undetermined coefficients l, m their values found from the two conditions that this line passes through the points A_1 and A_2,

$$la_1 + mb_1 + 1 = 0$$

and $$la_2 + mb_2 + 1 = 0,$$

and the result of eliminating l, m between the last three equations is the determinantal equation first written.

So x_1, y_1 are given by the simultaneous equations

$$\begin{vmatrix} x, & y, & 1 \\ a_1, & b_1, & 1 \\ a_2, & b_2, & 1 \end{vmatrix} = 0, \qquad \begin{vmatrix} x, & y, & 1 \\ a_3, & b_3, & 1 \\ a_4, & b_4, & 1 \end{vmatrix} = 0,$$

that is,

$$x(b_1 - b_2) + y(a_2 - a_1) + a_1b_2 - a_2b_1 = 0,$$
$$x(b_3 - b_4) + y(a_4 - a_3) + a_3b_4 - a_4b_3 = 0,$$

whence

$$\frac{x_1}{(a_2 - a_1)(a_3b_4 - a_4b_3) - (a_4 - a_3)(a_1b_2 - a_2b_1)} = \frac{y_1}{(b_2 - b_1)(a_3b_4 - a_4b_3) - (b_4 - b_3)(a_1b_2 - a_2b_1)}$$
$$= \frac{1}{(a_1 - a_2)(b_3 - b_4) - (a_3 - a_4)(b_1 - b_2)}.$$

What is important for our purpose is that x_1, y_1 are rational functions of each of the eight coordinates $a_1 \dots b_4$ of the four points A.

If then we have a set of given points (A), taking two pairs of them in every possible way, we can construct a new set of points (P) whose coordinates (x, y) are formed from the given coordinates (a, b) by equations of the above types. These points P can then be used in further constructions, and the numbers (x, y) can be added to the numerical data (a, b), and with these we can go on to construct fresh points and find fresh coordinates, at each step adding the fresh point to the data, and adding the fresh pair of coordinates to the set of numbers that may take the place of $a_1 \dots b_4$ in the above equations.

This process will terminate if we ever arrive at a figure in which every point obtained is already joined to every other by a straight line, as for example, a triangle, so that no new point or straight line can be obtained. We shall see later (p. 19) that this only happens in a few special cases; in general the process can be carried on indefinitely, and the number of points and straight lines which can be obtained is infinite.

At each step, the coordinates of the point just constructed are rational functions of those of the four points used in its construction. Hence, by successive substitution, the coordinates x and y of any point P, which can be constructed with ruler only from the given points A(a, b), are rational functions of the set of numbers a, b. The steps of the substitution correspond exactly to the steps of the ruler construction taken in reverse order; to the geometrical process of drawing the straight lines A_1A_2, A_3A_4 to meet in P_1 there corresponds the process of substituting for x_1, y_1 their values in terms of $a_1 \ldots b_4$ given above.

For example, let there be four given points, A(0, 0), B(0, 1), C(1, 0), D(a, b).

(i) Join AB, CD to meet in P(p, q), and join AC, BD to meet in Q(r, s).

(ii) Join AD, PQ to meet in X(x, y).

Here is a very simple construction in two stages: (i) from the four given points A, B, C, D we obtain two others, P, Q; (ii) from these points, together with two of the given set, we obtain the final point X. Corresponding to the three pairs of straight lines drawn, there are three pairs of equations:

$$p = 0, \quad q = \frac{b}{1-a}; \quad r = \frac{a}{1-b}, \quad s = 0; \quad \ldots\ldots\ldots\ldots\text{(i)}$$

$$x = \frac{a(ps-qr)}{b(p-r)-a(q-s)}, \quad y = \frac{b(ps-qr)}{b(p-r)-a(q-s)}. \quad \ldots\ldots\text{(ii)}$$

In order to express the coordinates x, y of X in terms of those of A, B, C, D, we start with equations (ii), corresponding to the last step of the construction, which express x, y in terms of the coordinates of P, Q as well as of A, D. In these expressions we substitute for p, q, r, s from equations (i), which correspond to the preceding stage of the construction, and so obtain the final expressions of x, y in terms of the coordinates of A, B, C, D only, which reduce to

$$x = \frac{a}{2-a-b}, \quad y = \frac{b}{2-a-b}.$$

This first general result may be briefly stated thus:

The points which can be constructed with ruler only from a given set of points have coordinates which are rational functions of the coordinates of the given points.

Indeterminate Constructions.

Often in the course of the solution of a problem we find instructions which are indeterminate, such as: "take any point" or "any straight line," "take any point upon a given straight line," "draw any straight line through a given point." These indeterminate operations are quite different from those we have been considering, in which each new point is completely determined by two straight lines already drawn, and each new straight line by two points. These vague operations are all equivalent to taking some arbitrary points $X(\xi, \eta)$ and carrying out definite operations upon them. Now these ill-defined points are of two classes: either the position of X does, or it does not, affect the position of the points to be finally constructed. If it does, the problem as it stands is indeterminate, and X must be regarded as belonging to the data, for the construction cannot be carried out until X has been chosen; and ξ, η must be included in the set of given coordinates $\mathbf{a}$, $\mathbf{b}$. If, on the other hand, X is a true auxiliary point, so that its position has no effect upon the final result, then the coordinates of the points required in the problem are the same wherever X may be chosen, and cannot involve ξ, η; in particular, they are the same as when X is chosen so that ξ, η are rational functions of $\mathbf{a}$, $\mathbf{b}$; as for example when X coincides with one of the given points A, or with one of the points P that can be immediately and definitely obtained from the points A. In this case, the final coordinates are certainly rational functions of the set of numbers $\mathbf{a}$, $\mathbf{b}$ only; therefore they are so, whatever the position of the auxiliary point X. It may quite well happen, when the given points are specially situated, that none of the points A or P will serve for X; as for example when all the given points lie on one straight line, and X has to be any point in the plane not lying on that line. If this line is taken to be the axis of $\mathbf{x}$, we have every $\mathbf{b}=\mathbf{0}$; but we can still take ξ, η to be rational functions of the $\mathbf{a}$, with the condition that η does not vanish. Strictly speaking, no ruler construction at all can be carried out on a set of given points all lying on one straight line; but by means of auxiliary points not lying on the line, we can construct other points, on the line, whose positions depend only on the given points, and which may therefore be said, in a slightly extended sense, to be constructed from the given points alone.

Converse Theory.

The next thing to do is to examine the converse theory, and determine whether all, or only some, of the points whose coordinates are rational functions of $\mathfrak{a}$, $\mathfrak{b}$ can be constructed with ruler only.

Now if our coordinates are Cartesian, the answer is that they cannot all be constructed. When a Cartesian frame of reference is given, the coordinates of a given point are two definite numbers; but though they are uniquely determined, they are defined by a hypothetical construction, namely by drawing a parallel to one of the axes; and if the data are a perfectly general set of points, we cannot draw parallels with ruler only. Though the coordinates are determinate, they cannot be constructed, and we cannot carry out an assigned rational operation on them. We must therefore eventually discard the Cartesian coordinates.

But we shall first show that, provided we can draw parallels, we can carry out any rational operation on the Cartesian coordinates. Since this is not so in general, we are now confining our attention to the particular case in which the data of the problem are not a perfectly general set of points, but are so specialized as to allow us to draw parallels. The simplest special property to assume is that among the data, or directly obtainable from them, are four points at the corners of a parallelogram; because when a parallelogram is given, we can with ruler only draw through any point a parallel to any straight line. This fundamental theorem is fully discussed in the next chapter, p. 51; for the present we take it for granted.

It is now convenient to choose a Cartesian frame as follows: let the origin O coincide with a given point A_1; let the axes pass through two other given points A_2, A_3, and let A_1A_2, A_1A_3 be taken as units of measurement along the respective axes, so that the points A, B at unit distances from O along the axes coincide with A_2, A_3. Then we must put

$$\mathfrak{a}_1 = \mathfrak{b}_1 = 0 ; \quad \mathfrak{a}_2 = 1, \mathfrak{b}_2 = 0 ; \quad \mathfrak{a}_3 = 0, \mathfrak{b}_3 = 1.$$

In the particular case which we are now considering, we can complete the parallelogram OAEB, where E is the unit point, whose coordinates are (**1**, **1**) ; and further, by drawing parallels through each of the given points to the sides and diagonals of this unit parallelogram, we can construct the points each of whose coordinates is one of the numbers **0**, ±**1**, ±**a**, ±**b**.

First Four Rules.

Next, if P, Q are the given points (**a**, **0**), (**b**, **0**), we can construct the point R, whose coordinates are (**a** + **b**, **0**), and hence, by repeating the process, all the points whose coordinates are linear functions of **a**, **b**, ... with positive or negative integral coefficients.

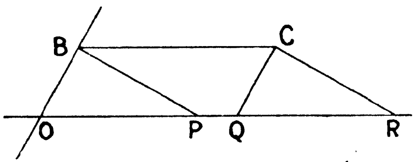

FIG. 1.

The construction of R is as follows :
Complete the parallelogram OQCB, and draw CR parallel to BP to meet OPQ in R, which is the required point.

For the triangles RCQ, PBO are congruent, and QR = OP ;

$$\therefore \quad OR = OQ + QR = OQ + OP = \mathbf{a} + \mathbf{b}.$$

If the three points O(**0**, **0**), A(**1**, **0**) and X(**x**, **0**) are given, and we wish to construct the point whose coordinate is **7x** + **3**, we can lay off along the axis, beginning at X, six adjacent lengths XX_1, X_1X_2, ... X_5X_6, each equal to OX, and three more X_6Y_1, Y_1Y_2, Y_2Y_3, each equal to OA ; then Y_3 is the point required.

We can also construct the points S(**ab**, **0**) and T(**a**/**b**, **0**). These are such that OS is a fourth proportional to OA, OP, OQ, while OT is a fourth proportional to OQ, OP, OA. Then to construct the point whose coordinate is **3**/**x**, we need to find in succession points A_2, A_3, Y along the axis, where $AA_2 = A_2A_3 = OA$, and OY is a fourth proportional to OX, OA, OA_3.

The construction of S is as follows :

Take any point C on OB, for example, the point $(0, a)$; join CA, CP to meet BE (or any other parallel to OPQ) in

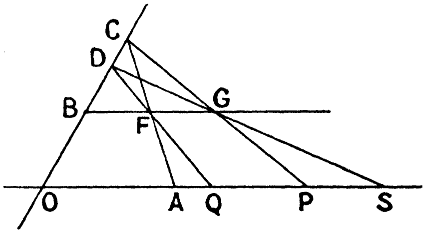

FIG. 2.

F, G; join QF to meet OC in D, and join DG to meet OPQ in S, which is the required point.

For, from similar figures, $\frac{OQ}{OS}=\frac{BF}{BG}=\frac{OA}{OP}$, so that OS is the required fourth proportional to OA, OP, OQ.

The construction of T is similar.

If in fig. 2 we allow P and Q to coincide, we obtain the point S, whose coordinate is a^2. To construct the point Y whose coordinate is x^2+3x+2, we can first find the point $S(x^2, 0)$, and then lay off from S along the axis three lengths equal to OX and two equal to OA; the end of the last segment is the point Y. Or it is shorter to construct the points H, K, whose coordinates are $x+1$, $x+2$, and then OY is a fourth proportional to OA, OH, OK.

Thus when once a parallelogram is drawn, we can find a ruler construction for any point whose coordinates are formed from those of points given or already obtained, by the processes of addition, subtraction, multiplication or division. But it is the definition of a rational function that it is formed from its arguments by a finite number of these four elementary operations. So that what we have just said is that we can find a ruler construction for any point whose coordinates are rational functions of the given coordinates. This set of functions includes all rational numbers, which are rational functions of the coordinates $(0, 1)$ of the single point A.

Combining this second general result with the first, we can now state that when a parallelogram is given it is possible to construct, with ruler only, all those and only those points whose Cartesian coördinates are rational functions of the coordinates of the given points.

Generalization by Projection.

So far we have only considered the case in which a parallelogram is given, and this allowed us to use Cartesian coordinates; we have now to remove the restriction. This can be done by the method of projection. If the whole figure, with which we have been dealing hitherto, is the projection of one in another plane, the four corners of the parallelogram correspond to a quadrangle, or set of four points, which need have no special properties, and may be any four points in that plane, provided that no three of them lie on one straight line. For if we project a plane α on to a plane β from a vertex of projection V, then the straight line fg, in which α is cut by the plane through V parallel to β, projects into the straight line at infinity in β. Thus by properly choosing β, we can project any straight line of α to infinity, and then any two straight lines in α meeting on fg project into two straight lines in β meeting at infinity, that is, into two parallels. Now we choose fg to be the straight line joining the points of intersection f, g of two pairs of opposite sides of the quadrangle in α; then the projection of the quadrangle has two pairs of parallel opposite sides, and is a parallelogram.

Then the join of any two points, or the intersection of any two straight lines, corresponds to the join or the intersection of their projections, so that ruler constructions in α correspond step by step to ruler constructions in β. If then, in any plane α in which points are given, four of these are chosen as a quadrangle of reference, we can project upon another plane β, choosing the projection so that the line at infinity in β corresponds to a diagonal of the quadrangle, and the quadrangle itself projects into a parallelogram. Then the points of α, which can be constructed with a ruler from the given points in α, project into the points of β which can be constructed with a ruler from the projections of the given points; and since these given projections include the four corners of a parallelogram, we can choose as above a Cartesian frame of reference in β, such that the corners of the parallelogram have the coordinates $(\mathbf{0}, \mathbf{0})$, $(\mathbf{1}, \mathbf{0})$, $(\mathbf{0}, \mathbf{1})$, $(\mathbf{1}, \mathbf{1})$ respectively;

and then we know that we can construct in β exactly those points whose coordinates are rational functions of the coordinates of the given projections; and hence we can construct in α exactly those points of which this set is the projection. We have therefore to find a meaning for these Cartesian coordinates in β, with reference only to the points of α and the quadrangle chosen from among them. This is done by expressing the coordinates as cross-ratios, which are not altered by projection (p. 38).

In the first plane α, let **oaeb** be the chosen quadrangle, and **fg** the one of its three diagonals which is projected into the line at infinity FG in β. Let capital letters denote the projections of the same small letters; then OAEB is

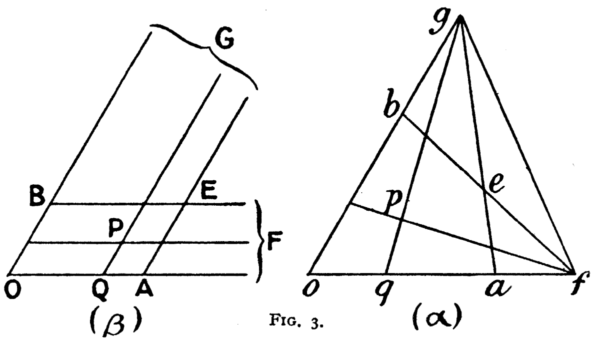

FIG. 3.

the parallelogram. F is the point at infinity in which the parallels OA, BE meet, and G is the point at infinity in which the parallels OB, AE meet. Let P(**x**, **y**) be any point in β, and PQ its ordinate, drawn parallel to OB to meet OA in Q; then QP produced also passes through G at infinity, and it is the projection of the straight line **qp**, joining **p** to the diagonal point **g** of the quadrangle, produced to meet **oa** in **q**.

Now the coordinates of P are **x**, **y**, the numerical measures of OQ and QP; and since OA is the unit of abscissae, the numerical measure of OQ is the ratio $\frac{OQ}{OA}$. Since F is at infinity, the ratio $\frac{QF}{AF}=1$, and we may divide the coordinate by this without altering its value.

$$\therefore \quad x = \frac{OQ}{OA} = \frac{OQ}{OA} \cdot \frac{AF}{QF}$$

= the cross-ratio of the two pairs of points OF, QA, which is written {OF, QA},

= {**of**, **qa**}, since a cross-ratio is unaltered by projection,

= **g**{**of**, **qa**}, that is, the cross-ratio of the pencil subtended at **g** by **of**, **qa**, so that the Cartesian coordinate of the projection of **p** is equal to the cross-ratio of this pencil, one of whose rays is **gp**, and the others are the sides and diagonal of the quadrangle of reference which pass through **g**.

Similarly, **y** = **f**{**og**, **pb**}.

We can therefore take the cross-ratios of these two pencils as the *projective coordinates* of the point **p**. They are unaltered by projection, and become the same as the Cartesian coordinates for the particular projection in which **fg** passes to infinity.

Then the points	**o** ;	**a** ;	**b** ;	**e** ;	**f** ;	**g** ;
have the coordinates	**0, 0** ;	**1, 0** ;	**0, 1** ;	**1, 1** ;	**∞, 0** ;	**0, ∞**.

We suppose that any four of the given points are taken to be **o**, **a**, **b**, **e** ; then **f** is determined as the intersection of **oa**, **be**, and **g** as the intersection of **ob**, **ae**. The system of projective coordinates is determined, and to them can be applied the results obtained before for Cartesian coordinates. The final conclusion is therefore :

If a set of points is given in a plane, then by ruler constructions those points and those points only can be obtained whose projective coordinates are rational functions of those of the given points, where any four of the given points, no three of which are in a straight line, are taken to determine the frame of reference, that is, to have the coordinates (**0, 0**) ; (**1, 0**) ; (**0, 1**) ; (**1, 1**).

Trivial Cases.

We have to assume that the data furnish four points, no three of which are in a straight line ; if there are not so many, the constructions fail. When all that is given is a single point, we can do nothing ; when two points are given, we can join them ; when we have three points, we can draw the sides (which may all coincide) of the triangle of which they are the vertices. More generally, when all the given points except one lie on the same straight line, we can draw this line and also the rays of the pencil subtended at the last point by the others. In all these cases we are not able to construct any new points.

Domains.

If the system consists of exactly four points, no three of which are on a straight line, they furnish the quadrangle of reference; the set of given coordinates is ($\mathbf{0}$, $\mathbf{1}$, $\boldsymbol{\infty}$), and the coordinates which can be obtained are the set of all rational numbers, positive, negative, or zero. All the numbers of this set are rational functions of $\mathbf{1}$ alone, for $\mathbf{0}$ is the difference $\mathbf{1}-\mathbf{1}$, and $\boldsymbol{\infty}$ is the reciprocal of $\mathbf{0}$. We say that the *domain* of the coordinates in this case is the set of all rational numbers, which we denote by $[\mathbf{1}]$.

Now let a fifth point $(\mathbf{a}, \mathbf{b})$ be added to the data; then if $\mathbf{a}$ and $\mathbf{b}$ are both rational, nothing is added to the domain of the coordinates; for if the point had not been given, it could have been constructed. But if $\mathbf{a}$, say, is an irrational number, we are now able to construct a much larger set of points; the domain is extended from all rational numbers to all rational functions of $\mathbf{a}$. This domain is denoted by $[\mathbf{1}, \mathbf{a}]$; it includes the former domain $[\mathbf{1}]$. If $\mathbf{b}$ is an independent irrational number, and not a rational function of $\mathbf{a}$, it also extends the domain, which becomes $[\mathbf{1}, \mathbf{a}, \mathbf{b}]$; but if either $\mathbf{a}$ or $\mathbf{b}$ is a rational function of the other, it can be omitted, for its inclusion adds nothing. In the same way, if there is any greater number of given points, whose coordinates are $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, ..., the corresponding domain is $[\mathbf{1}, \mathbf{a}, \mathbf{b}, \mathbf{c}, ...]$, in which any rational coordinates can be omitted, and also any which can be expressed as rational functions of the remaining irrational coordinates; and $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, ... may be replaced by another set of numbers $\mathbf{a}'$, $\mathbf{b}'$, $\mathbf{c}'$, ..., provided that each of these two sets can be expressed as a set of rational functions of the other set of numbers.

Suppose, for example, that we are given a square with an equilateral triangle described on one of its sides. Referred to this side and an adjacent side of the square as axes, the five given points may be taken to be

$$\mathrm{O}(\mathbf{0}, \mathbf{0}), \quad \mathrm{A}(\mathbf{1}, \mathbf{0}), \quad \mathrm{B}(\mathbf{0}, \mathbf{1}), \quad \mathrm{E}(\mathbf{1}, \mathbf{1}), \quad \mathrm{C}(\tfrac{1}{2}, \tfrac{1}{2}\sqrt{\mathbf{3}}),$$

and the domain is most simply expressed as [**1**, $\sqrt{3}$]. We can now construct equilateral triangles on all the other sides of the square and on its diagonals, for the coordinates of all the other points required involve no surd other than $\sqrt{3}$. For example, if ABF is the equilateral triangle on AB on the side remote from O, the coordinates of F are each $\frac{1}{2}(1+\sqrt{3})$, and F can be constructed by drawing CD

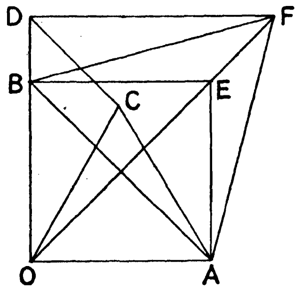

FIG. 4.

parallel to AB to meet OB in D, and DF parallel to OA to meet OE in F. But we cannot construct a regular pentagon on any straight line in the figure, for the coordinates of one or more of its corners would involve $\sqrt{5}$, which is not a rational function of $\sqrt{3}$, and therefore does not belong to the domain of the coordinates of points which can be constructed from the five given points with ruler only.

II. Ruler and Compasses.

When we also use compasses, a new point may be determined not only as the intersection of a pair of straight lines, but also as one of the intersections of a circle and a straight line, or of two circles. The case of two circles is much the same as that of a circle and a straight line, for the second circle may be replaced by the common chord or radical axis as below, p. 22.

When a circle is drawn, any two diameters cut the circumference in the corners of a rectangle, so that we can use rectangular Cartesian coordinates; and since we can cut off equal lengths from the two axes, we can use the same scale for abscissae and ordinates Now the equation of the straight line joining two points is linear, with coefficients rational in the coordinates 1, a, b, ... of the points already constructed, say

$$lx + my + 1 = 0.$$

The circle determined by its centre $A_1(a_1, b_1)$ and a point $A_2(a_2, b_2)$ on its circumference, which we call the circle $A_1(A_2)$, has for its equation

$$(x - a_1)^2 + (y - b_1)^2 = (a_2 - a_1)^2 + (b_2 - b_1)^2\,;$$

more generally, the circle determined by its centre A_1 and radius equal to A_2A_3, which we call the circle $A_1(A_2A_3)$, has for its equation

$$(x - a_1)^2 + (y - b_1)^2 = (a_2 - a_3)^2 + (b_2 - b_3)^2,$$

which is a quadratic equation in x, y, whose coefficients also belong to the domain $[1, a, b, ...]$. To find the abscissae of the points of intersection of this straight line and this circle, eliminate y between the two equations and solve for x. Substitute for y in the second equation the value $-\frac{lx+1}{m}$ obtained from the first. We find

$$(x - a_1)^2 + \left\{-\frac{lx+1}{m} - b_1\right\}^2 = (a_2 - a_3)^2 + (b_2 - b_3)^2,$$

or say $$px^2 + 2qx + r = 0,$$

which is a quadratic equation for x, whose coefficients p, q, r belong to the same domain as before. The expression which we obtain for x is $\frac{1}{p}(-q \pm \sqrt{q^2 - pr})$, which does not in general belong to $[1, a, b, ...]$, but involves the square root of the rational expression $q^2 - pr$; and y involves the same square root.

In the case of two circles we can take the equations to be

$$x^2 + y^2 - 2a_1x - 2b_1y + c_1 = 0,$$
$$x^2 + y^2 - 2a_2x - 2b_2y + c_2 = 0,$$

with coefficients belonging to $[1, a, b, ...]$. The coordinates of the points of intersection satisfy both these equations, and satisfy the equation found by subtracting,

$$2(a_2 - a_1)x + 2(b_2 - b_1)y + c_1 - c_2 = 0,$$

which is linear, and therefore represents the straight line joining the two real or imaginary points of intersection, that is, the common chord of the two circles. The intersections are therefore the same as those of either circle with this straight line, and since the coefficients of the linear equation also belong to [**1**, **a**, **b**, ...], so that it is of the type $lx + my + 1 = 0$ considered above, this case is included in the last.

Thus the geometrical operation of describing a circle with given centre and radius to cut a given straight line in two points, is equivalent to the analytical operation of taking a square root, together with some rational operations. Conversely, we can find a geometrical operation equivalent to taking the square root of any number **p**,

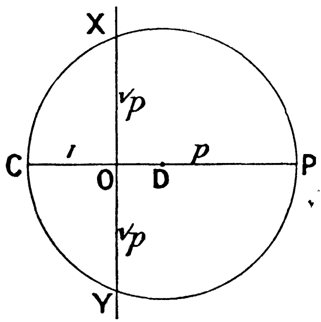

FIG. 5.

provided that **p** is a coordinate of a point which can be constructed.

Let P be the point (**p**, **0**) and let C be (−**1**, **0**). Bisect CP in D; then the circle D(P) meets the axis of **y** in two points X, Y, such that

$$YO \cdot OX = CO \cdot OP = 1 \cdot p;$$

$$\therefore\ OX = \sqrt{p}, \qquad OY = -\sqrt{p}.$$

Quadratic Surds.

Thus the analytical steps equivalent to the various steps of a ruler and compass construction are rational operations together with the new operation of taking a square root; and these may be combined and repeated in any way a finite number of times and carried out upon the coordinates

($\mathbf{1}$, $\mathbf{a}$, $\mathbf{b}$, ...) of the given points. The result of such an analytical process, if it does not happen to be rational, is called a *quadratic surd*; the simplest type is the square root of a rational expression, but a general quadratic surd may consist of several terms, and the quantity under each radical need not be rational, but may be another quadratic surd. But the whole expression must involve no roots but square roots, and others, such as fourth roots, which can be expressed by means of square roots only; and it must not involve an infinite series of operations.

Now if we want to know whether some definite problem can be solved by means of ruler and compasses, we can express the conditions laid upon the elements to be constructed, which, as before (see p. 8), we take to be points, as a set of simultaneous equations connecting their unknown coordinates with the known coordinates of the given points. If the problem is determinate, there are just as many equations as unknowns. Then the problem can be solved by means of ruler and compasses, if and only if the solution of the equations can be completely expressed by means of rational functions and quadratic surds. It is a necessary condition that the determination of the coordinates can be made to depend on a set of *algebraic* equations. Each step of a Euclidean construction determines a point as the intersection of straight lines or circles, either uniquely or as one of a pair of possible positions. Since a construction has only a finite number of steps, when all the alternatives are considered the final coordinates have one of a finite number of sets of possible values, and their determination can be made to depend on a set of equations having only a finite number of solutions, that is, of finite degree and therefore algebraic, of the form

$$\mathsf{F}(\mathbf{x}, \mathbf{y}, \ldots, \mathbf{a}, \mathbf{b}, \ldots) = \mathbf{0},$$

where $\mathbf{x}$, $\mathbf{y}$, ... are the unknown, and $\mathbf{a}$, $\mathbf{b}$, ... the known coordinates, and F is an algebraic expression.

For the sake of simplicity, let us now take the case in which all the given coordinates are rational, so that the domain is [$\mathbf{1}$]. In the general case, the whole argument follows the same course, and leads to corresponding results, if for the set of rational numbers we substitute the domain [$\mathbf{1}$, $\mathbf{a}$, $\mathbf{b}$, ...]. By the known processes of the theory of equations, $F=0$ can first be rationalized, so that we can take F to be a polynomial with rational coefficients; next, we can eliminate all but one of the unknown coordinates, and suppose each unknown given by a separate equation. Thirdly, $F(x)$ may be reducible, that is, it may fall into two or more factors with rational coefficients, of any degrees in x from 1 up to $n-1$ if n is the degree of F itself; in this case processes are known for finding the factors and reducing the equation. We therefore need only consider an algebraic, irreducible equation in a single unknown, with rational coefficients.

Now consider a quadratic surd x; suppose that its denominator has been rationalized (p. 27), and that it is written before us. It consists of a collection of rational numbers, of symbols denoting rational operations, and of radicals which are symbols denoting the extraction of square roots. In reading this collection, the first time we come to a radical let us put the expression underneath it equal to y_1^2, so that we can write y_1 instead of the whole square root, and pass on till we come to another square root, for which we write y_2, and so on. Since each square root may be multiplied by a rational factor, x is now expressed in the form

$$x = c_1 y_1 + c_2 y_2 + \dots + c_k y_k + c', \qquad \text{(I)}$$

where each c is rational and each y^2 is either rational or a quadratic surd simpler than x; for even in the extreme case when $k=1$, there is one radical fewer in y_1^2 than in x.

Now treat y_1^2, y_2^2, ... y_k^2 in the same way as we have treated x; then there are k equations of the type

$$y_k^2 = c_{k,1} y_{k,1} + c_{k,2} y_{k,2} + \dots + c_{k,l} y_{k,l} + c_k', \qquad \text{(II)}$$

where each c is rational and each $y_{k,l}^2$ is a quadratic surd simpler than y_k.

Continue this process with $y_{1,1}^2$, ... $y_{k,l}^2$, until there are no radicals left. We thus introduce as many new unknowns y as there are different square roots in x, and the last set of equations corresponding to (II) are of the type

$$y^2_{k,l,m,\dots} = c'_{k,l,m,\dots} \qquad \text{(III)}$$

For example, let

$$x = \sqrt{5a + \sqrt{6b}} + \sqrt{8b + \sqrt{7a}}.$$

We put $\quad x = y_1 + y_2, \quad \ldots\ldots\ldots\ldots \text{(I)}$

$$y_1^2 = 5a + y_{11}, \quad y_2^2 = 8b + y_{21}, \quad \ldots\ldots\ldots \text{(II)}$$

$$y_{11}^2 = 6b, \quad y_{21}^2 = 7a. \quad \ldots\ldots\ldots \text{(III)}$$

Or again, let $\quad x = \sqrt{1 + \sqrt{3}} + \sqrt{3}.$

Then $\quad x = y_1 + y_2,$

$$y_1^2 = 1 + y_2, \quad y_2^2 = 3.$$

Now take the equations in reverse order. Each of the last set of y's is connected with rational terms by one of the equations (III); that is, it is determined by a quadratic equation with rational coefficients; we can solve these equations and find this set of y's. Each of the set of y's which came just before is determined by an equation of type (II), whose coefficients are either rational, or rational functions of the y's which have just been found. At every step a set of y's are each found from a quadratic equation whose coefficients are rational in the y's already found, till finally we find $y_1, \ldots y_k$, and then x itself is given by a linear equation connecting it with these. Thus we can find x by solving a series of equations of degrees 1 and 2 only; hence *If an equation can be solved by quadratic surds, its solution can be made to depend on that of a series of equations of first and second degrees.* In the first place, all linear and quadratic equations can be so solved, but no irreducible cubic; for we shall now show, more generally, that the degree of an irreducible equation which can be solved by quadratic surds must be a power of 2; this is a necessary condition, but it is not sufficient.

Conjugate Surds.

If the signs of any particular set of these surds y are changed wherever they occur in x, the result is another quadratic surd x_1 which is said to be *conjugate* to x. In this way we obtain a set of 2^n conjugate surds, of which x is one. For example, if

$$x = \sqrt{2} + \sqrt{3} + \sqrt{2} \cdot \sqrt{3},$$

the other three conjugate surds are

$$-\sqrt{2} + \sqrt{3} - \sqrt{2} \cdot \sqrt{3}, \qquad \sqrt{2} - \sqrt{3} - \sqrt{2} \cdot \sqrt{3},$$

$$-\sqrt{2} - \sqrt{3} + \sqrt{2} \cdot \sqrt{3}.$$

But if we write $\mathbf{x} = \sqrt{2} + \sqrt{3} + \sqrt{6}$, we must not consider all three square roots as independent $\mathbf{y}$'s; for to change the sign of $\sqrt{2}$, without at the same time changing the sign of $\sqrt{6}$, does not give a conjugate surd. In this case there exists a relation $\sqrt{2} \cdot \sqrt{3} = \sqrt{6}$, which does not hold when the sign of $\sqrt{2}$ only is changed. We avoid the case in which the $\mathbf{y}$'s are dependent in this way; for a fuller discussion of independent surds, see Enriques, p. 140 ff. The equations (II), (III) may be used to eliminate all squares of $\mathbf{y}$'s from any relations that exist; if they do not then become identities, they are reduced to a form linear in each $\mathbf{y}$, and can be used first of all to eliminate some of the $\mathbf{y}$'s altogether. We suppose this carried out; the remaining $\mathbf{y}$'s are then linearly independent, that is, there exists no rational relation between them which is linear in each separately. It is a consequence of this, that any relation that exists remains true if the sign of any one $\mathbf{y}$ is changed throughout.

We shall now prove that if a given rational equation $f(X) = 0$ is satisfied by $X = \mathbf{x}$, then it is also satisfied by $X = \mathbf{x}_1$, where $\mathbf{x}_1$ is any of the surds conjugate to $\mathbf{x}$.

Rationalizing Factor.

Any symmetric function of a complete set of conjugate surds is rational. For if it were irrational, it would depend really, and not only apparently, on some one of the radicals, and would be altered in value by changing the sign of that radical. But to change the sign of one radical changes each of the surds into a conjugate surd, and permutes the set of conjugates among themselves, without changing the set as a whole, and therefore does not alter a symmetric function of the set; these symmetric functions therefore do not really depend on the sign of any radical, but are rational. In particular, the continued product of the whole set is rational, and any surd has as a *rationalizing factor* the product of all the other surds conjugate to it. We can rationalize the denominator of the fraction considered above (p. 25), which is the quotient of two rational integral functions of the $\mathbf{y}$'s, by multiplying numerator and denominator by the product of all the surds conjugate to the denominator. This rationalizing factor is by construction an integral function of the different surds y, of which the denominator is a function, and of their conjugates, which are other surds of the same nature.

Consider the result of substituting x for X in the function $f(X)$. In the first place, we substitute for X the expression (I) linear in each y, and carry out on this the operations indicated by f. Now, since $X = x$ is a root of $f(X) = 0$, this last expression vanishes identically, as a direct consequence of the defining equations (I), (II), which, combined in a certain way, lead to $f(x) = 0$.

Now instead of x let us substitute x_1 in f, where x_1 only differs from x in the sign of a single surd y_1. If y_1 occurs anywhere under a radical this has the effect of changing some of the other surds y into their conjugates y' wherever they occur, but these new surds y' are defined by equations of exactly the same forms (II), (III) as before, so that $f(x_1)$ vanishes as a direct consequence of these, combined in the same way as before.

For example, consider the surd

$$x = \sqrt{3} + \sqrt{1 + \sqrt{3}}.$$

In order to find the rational equation satisfied by this, we first transpose the term $\sqrt{3}$ and square,

$$(x - \sqrt{3})^2 = 1 + \sqrt{3},$$

then rearrange in the form

$$x^2 + 2 = \sqrt{3}(2x + 1)$$

and square again. Therefore x is a root of the equation

$$f(X) \equiv X^4 - 8X^2 - 12X + 1 = 0.$$

Substitute and expand:

$$\begin{aligned} f(x) &= 3^2 + 4 . 3\sqrt{3}\sqrt{1 + \sqrt{3}} + 6 . 3(1 + \sqrt{3}) \\ &\quad + 4\sqrt{3}(1 + \sqrt{3})\sqrt{1 + \sqrt{3}} + (1 + \sqrt{3})^2 \\ &\quad - 8\{3 + 2\sqrt{3}\sqrt{1 + \sqrt{3}} + 1 + \sqrt{3}\} - 12\{\sqrt{3} + \sqrt{1 + \sqrt{3}}\} + 1 \\ &= (9 + 18 + 1 + 3 - 24 - 8 + 1) + \sqrt{3}(18 + 2 - 8 - 12) \\ &\quad + \sqrt{1 + \sqrt{3}}(12 - 12) + \sqrt{3}\sqrt{1 + \sqrt{3}}(12 + 4 - 16), \end{aligned}$$

each term of which vanishes separately. If in this we change the sign of $\sqrt{3}$, we must replace the second independent surd $\sqrt{1+\sqrt{3}}$ by the conjugate $\sqrt{1-\sqrt{3}}$; we have four terms

$$f(x_1) = (0) - \sqrt{3}(0) + \sqrt{1-\sqrt{3}}(0) - \sqrt{3}\sqrt{1-\sqrt{3}}(0),$$

each of which vanishes as before. But the quartic equation is not satisfied by $-\sqrt{3}+\sqrt{1+\sqrt{3}}$, which is not one of the surds conjugate to x.

In the general case, it follows that $f(X)=0$ is also satisfied by $X=x_2$, where x_2 is a surd conjugate to x_1 formed by changing the sign of any other independent surd y_2 in x as well as the sign of y_1; and so on. If the signs of any number of the surds are changed wherever they occur in x, the result is a conjugate surd x_p, which also satisfies $f(X)=0$.

Thus, *if a rational equation has a quadratic surd for one root, all the conjugate surds are roots.* But these 2^n conjugate surds may not all have different values, so that the number of different roots of $f(X)=0$ given in this way is N say, where $N \leqq 2^n$. For example, if

$$x = \sqrt{5+\sqrt{3}} + \sqrt{5-\sqrt{3}},$$

its value is not altered by changing the sign of $\sqrt{3}$ in both terms.

Degree a Power of 2.

We next prove that N is a factor of 2^n, and is therefore itself a power of 2.

Let $x_q'(q=1, 2, \dots N)$ be the set of all the N different values of all the 2^n conjugate surds $x_p(p=1, 2, \dots 2^n)$.

Consider the two products

$$F(X) = \Pi(X - x_p), \quad F'(X) = \Pi(X - x_q')$$

of degrees 2^n, N respectively. A change of sign of any particular square root wherever it occurs in the x, interchanges the x in pairs, and therefore merely alters the order of the factors in F, and does not affect the value of the coefficient of any power of X in F; these coefficients therefore, being independent of the sign of any of the square roots that occur, are rational: F is a rational function of X with rational coefficients. Again, the change of sign of a square root changes the x' into another set of N of the surds x, all different, and therefore the same as the x' in another order; the change merely alters the order of the factors in F' and does not affect its coefficients, so that F' also is a rational function of X with rational coefficients. Moreover F' is irreducible; for if it fell into two rational factors F_1' and F_2', each would be the product of some but not all of the differences $X - x'$, and $F_1' = 0$ would have some but not all of the x as roots; but we have seen that if any rational equation has one of the x as roots, it has them all, so that this is a contradiction, and it follows that F' is irreducible. But F' consists of N from among the factors of F, so that F has F' as a factor, and the remaining factor F'' must also be rational. F'' may be a mere constant; in this case F and F' are of the same degree, $N = 2^n$, and the x are all different. But if not, F'' is of degree $2^n - N$; it is the product of some of the factors of F, and $F'' = 0$ has at least one of the x as a root; hence it has all the N different surds x' as roots, and, by the same argument as before, F'' has F' as a factor. Since both F'' and F' are rational, the remaining factor F''' of F'' is rational, and its degree is $2^n - 2N$. The same argument applies to F'''; it is either a constant or it contains F' as a factor, and then we can treat its other factor in the same way. We go on, at each step removing the factor F', and the process cannot stop as long as the remaining factor is of positive degree. Since the degree of F is finite, there can only be a finite number of steps, and we come at last to a factor of degree 0, that is, a constant. Hence F is the product of a power of F' and this constant, and its degree 2^n is a multiple of the degree N of F. Therefore N is a factor of 2^n, and is itself a power of 2, say $N = 2^k$.

We have now proved that any equation $f(X) = 0$ with rational coefficients which has a quadratic surd x as a root has as roots all the 2^k different values x' of the set of conjugate surds of which x is one; and f has all the differences $X - x'$ as factors, and therefore their product F', so that f has the rational factor F'. Therefore either f is reducible or it is a constant multiple of F', and its degree is 2^k. Therefore:

If an irreducible algebraic equation with rational coefficients can be solved by quadratic surds, its degree is a power of 2. Note that *all* its roots are conjugate surds.

Duplication and Trisection.

These theorems have been stated with regard to algebraic equations whose coefficients are rational, that is, belong to the domain $[1]$. They can be extended to apply to equations whose coefficients belong to any given domain $[1, a, b, \ldots]$.

It follows that the ancient problems of the duplication of a cube and the trisection of an angle cannot be solved by means of ruler and compasses. The problem of finding the side x of a cube whose volume is twice that of a cube of given side a is equivalent to solving the cubic equation

$$x^3 = 2a^3,$$

whose coefficients are rational in $[1, a]$, but whose real root $x = a \,.\, \sqrt[3]{2}$ involves a surd which is not quadratic. And the trisection of an arbitrary angle θ carries with it the determination of $\cos \frac{1}{3}\theta$ ($= y$ say) when $\cos \theta$ ($= b$ say) is given; but since $\cos \theta = 4 \cos^3 \frac{1}{3}\theta - 3 \cos \frac{1}{3}\theta$, this is equivalent to solving the cubic equation

$$4y^3 - 3y - b = 0,$$

which is irreducible except for special values of the given cosine b. These two famous problems are therefore beyond the range of ruler and compass constructions.

III. Regular Polygons.

A good example of this theory is to determine what regular polygons can be constructed with ruler and compasses. If n the number of sides is not prime, say $n = a_1 a_2$, then the n-gon also furnishes a regular a_1-gon by taking a_1 of the vertices, each separated from the next of the a_1 by $a_2 - 1$ consecutive vertices of the n-gon. For example, the 1st, 4th, 7th and 10th vertices of a regular dodecagon form a square. So if we could construct the n-gon, we could also from it construct the a_1-gon; and if the latter is known to be impossible, it follows that the former must be impossible also, and it is useless to consider it. In the first place, therefore, we consider prime values of n. The construction can be made to depend on finding the irrational roots of the equation $x^n - 1 = 0$. This is reducible, but removing the factor $x - 1$, we have

$$x^{n-1} + x^{n-2} + \ldots + x + 1 = 0,$$

which can be shewn to be irreducible if n is prime (Enriques, p. 152). Then, if the polygon can be constructed with ruler and compasses, this equation can be solved by quadratic surds, and its degree $n-1$ must be a power of 2; hence

$$n=2^k+1.$$

Further, since n is prime, k must also be a power of 2; for if k had an odd factor l other than 1, let $k=lm$; then

$$n=(2^m)^l+1,$$

which is not prime, but has 2^m+1 as a factor. Hence k has no odd factor, and we may write

$$k=2^p, \quad n=2^{2^p}+1.$$

The only exception to the condition $k=2^p$ is the trivial case $k=0$, $n=2$.

It has also been proved, conversely, that if n is a prime of this form, the equation can be solved and the polygon constructed. For example, if $p=1$, $n=5$, one root is

$$x=\tfrac{1}{4}\{-1-\sqrt{5}+\sqrt{-10+2\sqrt{5}}\},$$

and the corresponding geometrical construction may be carried out as in the fourth book of Euclid.

Putting	$p=$	0,	1,	2,	3,	4,
we have	$n=$	3,	5,	17,	257,	65537.

The next values $p=5, 6, 7$ do not give prime values for n, and little is known of the higher numbers of the series. Hence **3, 5, 17, 257** are the only prime values of n less than **10,000** for which the regular n-gon can be constructed with ruler and compasses.

Next, if n is composite, let

$$n=a_1^{k_1}a_2^{k_2}\dots,$$

where a_1, a_2 ... are the different prime factors of n. Then, if the n-gon can be constructed, so can the a_1-gon; and since a_1 is prime, it is of the form $2^{2^p}+1$, unless $a_1=2$.

Now all the roots of $\dfrac{x^n-1}{x-1}=0$ are quadratic surds; but the left-hand side has the factor $\dfrac{x^{a^k}-1}{x-1}$ (with any suffix to k), and this again has the factor $\frac{x^{a^k}-1}{x^{a^{k-1}}-1}$; hence the rational equation of degree $a^{k-1}(a-1)$,

$$x^{a^{k-1}(a-1)}+x^{a^{k-1}(a-2)}\dots+x^{a^{k-1}}+1=0,$$

can be solved by quadratic surds. Now this equation can be shown to be irreducible when a is prime, so its degree $a^{k-1}(a-1)$ is a power of 2, and therefore both a^{k-1} and $a-1$ are powers of 2. Now if $a=2$, both these conditions hold for all values of k; but if $a=2^{2^p}+1$, then $a-1$ is a power of 2, but a is not, and a^{k-1} is not, unless $k=1$, when $a^{k-1}=a^0=1=2^0$. Thus we must have either $a=2$ or $k=1$, and n can have no repeated factor other than 2; we may write

$$n=2^k(2^{2^{p_1}}+1)(2^{2^{p_2}}+1)\dots,$$

where p_1, p_2 ... are different integers, and all the numbers $2^{2^p}+1$ are prime.

Conversely, if a_1, a_2 are two numbers prime to one another, and the regular a_1-gon and a_2-gon can be constructed, so can the regular a_1a_2-gon. For we can find two positive integers N_1, N_2 such that $N_1a_2-N_2a_1=\pm 1$. Then the difference between the angles subtended at the centres of the polygons by N_1 sides of the first and N_2 sides of the second is

$$N_1\frac{2\pi}{a_1}-N_2\frac{2\pi}{a_2}=(N_1a_2-N_2a_1)\frac{2\pi}{a_1a_2}=\pm\frac{2\pi}{a_1a_2},$$

which is the angle subtended at the centre of a regular a_1a_2-gon by one of its sides; so that we can construct this angle, and therefore the whole polygon. This can be extended to any number of factors prime to one another. Also any angle can be bisected any number of times, so that the factor 2^k can be provided for. Therefore if n has the form given above, the regular n-gon can certainly be constructed.

Thus with ruler and compasses we can construct regular polygons of

	2,	3,	4,	5,	6,		8,		10,		12,			15,	16,	17,			20,	... sides,
but not of						7,		9,		11,		13,	14,				18,	19,		... sides,

Regular 17-gon.

Here is a construction for the regular polygon of **17** sides (Richmond).

Take perpendicular radii OA, OC of a circle centre O; bisect OC and bisect again, making $OD=\frac{1}{4}OC$; join DA. Bisect $\angle ODA$ and bisect again, making $\angle ODE=\frac{1}{4}\angle ODA$. Draw a perpendicular at D to DE, and bisect the right

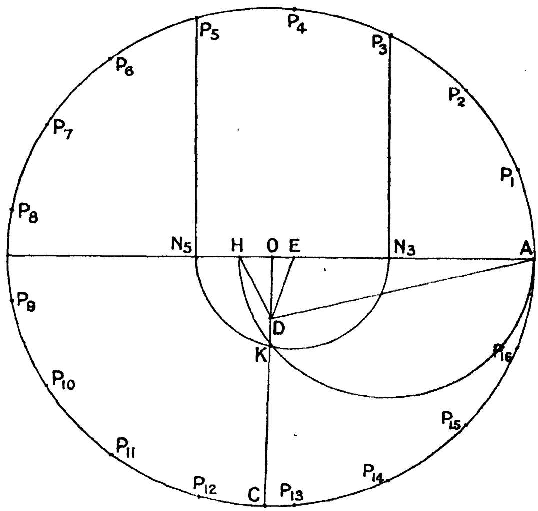

FIG. 6.

angle, making $\angle EDH=\frac{1}{4}\pi$, its arms meeting OA in E, H. Describe the circle on AH as diameter to meet OC in K, and describe the circle E(K) to meet OA in N_3, N_5. Draw ordinates through these points perpendicular to OA to meet the circle O(A) in P_3, P_5; then these are the third and fifth vertices of a regular **17**-gon, of which A is the last.

CHAPTER III.

RULER CONSTRUCTIONS.

In this chapter we consider the actual carrying out of some of the ruler constructions which we have found to be possible when a set of points are given in a plane. First of all, an important distinction must be made according as the positions of the points have or have not any special relations to one another. If we project all the points and lines that we obtain, upon any other plane, the result is a figure in this second plane, which can also be constructed with a ruler if we are given in the second plane the projections of the points which are given in the first plane. The two constructions correspond step by step, and every incidence of a point with a line in either plane corresponds to the incidence of their projections in the other plane.

Now if the data have no special relations, the only properties possessed by the figure which we first constructed are due to the various incidences secured by the steps of the construction; for these steps consist either of drawing a new straight line which passes through two points already obtained, or of obtaining a new point which lies on two straight lines already drawn. Since all these incidences occur in the second figure also, it follows that all the properties of each figure hold of the other; they are said to be *descriptive* or *projective properties*. They are unchanged by any projection or by any series of projections, and so belong to a whole set of figures, any one of which can be obtained from any other by one or more projections. All the figures of the set are exactly alike as regards projective properties.

If a figure possesses any properties which are not of this nature, they are called *non-projective*, or, more usually, *metrical*. These cannot be due to the steps of the construction, when this is carried out with ruler only, and must be due to the data, which are therefore a set of points whose positions are not general, but which enjoy certain special relations which do not hold between their projections, and so do not lead to the same properties of the second figure. But to a metrical property of the first figure there must correspond *some* property of the second, and the latter must be a projective property, since it holds of all the projections of the first figure; it must therefore hold of all the projections of these projections, among which the first figure itself is included. So there exists a projective property, arising out of the metrical property, which holds of all the projective set of figures, including the first figure; and when applied to this original figure, it must give us exactly the same information about it as the metrical property from which we started. Thus *any metrical property can be stated in projective language*, and then it is true of all the projective set to which the original figure belongs; stated in metrical language, which is usually simpler, it is true of the original figure only.

Conversely, from a figure having a projective property we can obtain a figure having a corresponding metrical property, by projecting so that certain elements assume special positions. The most familiar case is that of projecting a straight line to infinity, when two straight lines intersecting at a point on that particular straight line become two parallel straight lines. Here "intersecting at a point on that particular straight line" and "parallel" are respectively the projective and metrical ways of stating the same property.

We can at once apply this way of looking at things to the projection used on p. 18. The first figure there has one metrical property, namely that the straight line FG is at infinity, which is not true of its projection **fg**; we can avoid mentioning this if for "parallel straight lines" we substitute "straight lines meeting on FG", and in particular, if for "straight lines parallel to the axes" we substitute "straight lines through F or G"; to these there correspond, in the second figure, "straight lines meeting on **fg**, or in **f** or **g**" respectively. Then a ruler construction for carrying out any rational operation upon the projective coordinates of points in the second plane can be deduced from those given in chapter II for Cartesian coordinates by merely making these changes in the wording. This gives a theoretical method of solution for any linear problem whatever.

The simplest metrical properties are concerned with length and angular measure, each regarded as a ratio. Among the simplest metrical statements that can be made are: that a given segment of a straight line is bisected at a given point, that two straight lines are parallel, that two straight lines are at right angles. The idea of area is metrical, and so is that of similarity; the usual definition of a circle is entirely metrical. The simplest projective property is incidence, as explained above, and next come the cross-ratios of ranges of four points on a straight line or of pencils of four straight lines through a point.

I. Projective Properties.

Desargues' Theorem.

An example of a projective property concerned with incidence only is given by *Desargues' theorem*: *If two triangles are in perspective, the pairs of corresponding sides meet in collinear points.* For a full discussion, see Mathews, chap. V.

Two triangles ABC, A′B′C′ are said to be in perspective if the three straight lines AA′, BB′, CC′ joining corresponding vertices all pass through one point O, called the centre of perspective or of homology. The theorem states that if the pairs of corresponding sides BC, B′C′; CA, C′A′; AB, A′B′ meet in X, Y, Z respectively, then the three points XYZ lie on a straight line called the axis of perspective or of homology.

This is true whether the two triangles ABC, A′B′C′ are in the same plane or in different planes; in the second case, the axis XYZ is the straight line of intersection of the planes of the two triangles.

We can apply Desargues' theorem to draw a straight line through a given point to pass through the point of intersection of two given straight lines, when for any reason it is not possible to lay the edge of a ruler against this point of intersection (see chapter VI). In fig. 7, let AA′, BB′ be the given straight lines and O their inaccessible point of intersection. If C is the given point, we first take any two points A, B, one on each of the given straight lines, and then construct a triangle A′B′C′ which is in perspective with ABC, with O as centre of perspective.

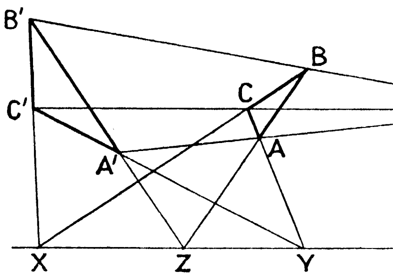

FIG. 7.

Then CC′ is the required straight line. Take any convenient axis of perspective and let it cut the sides of the triangle ABC in XYZ. Draw any straight line through Z to meet the given straight lines in A′, B′ respectively. Join YA′, XB′ to meet in C′; then CC′ passes through O, the point of intersection of AA′, BB′. In particular, if AA′, BB′ are parallel, so that O is at infinity, we can thus draw a third parallel through any given point C. Several other constructions for the general problem are given on p. 102, ex. 1.

Cross-ratio.

Most projective properties are connected with cross-ratio. The fundamental fact that a cross-ratio is unaltered by projection has already been assumed in several places. It is easily proved by reference to the angles at the vertex of projection.

Let there be given a range of four points A, B, C, D on a straight line, and let them be projected from any point O on to another transversal. Draw OP perpendicular to AB. Then, since

$$AC \,.\, OP = 2\triangle AOC = OA \,.\, OC \sin AOC,$$

the cross-ratio of the range is

$$\{AB, CD\} = \frac{AC \cdot DB}{CB \cdot AD} = \frac{AC \cdot OP \cdot DB \cdot OP}{CB \cdot OP \cdot AD \cdot OP}$$

$$= \frac{OA \cdot OC \sin AOC \cdot OD \cdot OB \sin DOB}{OC \cdot OB \sin COB \cdot OA \cdot OD \sin AOD}$$

$$= \frac{\sin AOC \cdot \sin DOB}{\sin COB \cdot \sin AOD},$$

and therefore depends only on the angles at O, and not

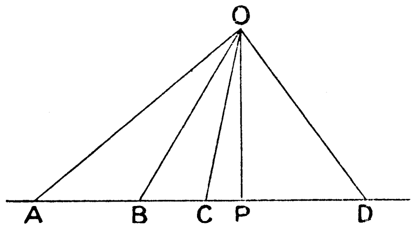

FIG. 8.

on the particular transversal AB, but is the same as the cross-ratio of the range cut out on any other transversal.

The first problem which presents itself is the construction of a range of given cross-ratio: more definitely, given three points OAF in a straight line, to find a point X in the

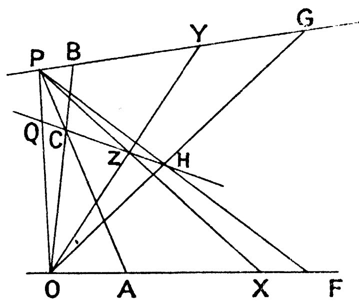

FIG. 9.

same straight line such that $\{OF, XA\} = \frac{OX \cdot AF}{XF \cdot OA} = \lambda$, where λ is a given number. The values $\lambda = \mathbf{0}, \mathbf{1}, \infty$ are obtained by making X coincide with O, A, F respectively. If the assigned value λ is given as the cross-ratio {PG, YB} on another straight line, we can carry out the construction as follows.

Let PA, OB meet in C, and PF, OG in H. Let CH meet OP in Q and OY in Z. Then PZ meets OA in X; for

$$\{OF, XA\} = \{QH, ZC\} = \{PG, YB\},$$

the vertices of projection being P and O respectively.

If the given and required ranges are on the same straight line, we can first project the former on to any other straight line, and then proceed as before.

Harmonic Range.

The case $\lambda = -1$ is specially important; then OF is divided internally and externally in the same ratio at A and P, and the range is said to be *harmonic*. Its importance depends largely on the fundamental *harmonic property of the quadrilateral: Any diagonal of a complete quadrilateral is divided harmonically by the other two diagonals.* This property can also be taken as the definition of a harmonic range (see Mathews, chap. VI).

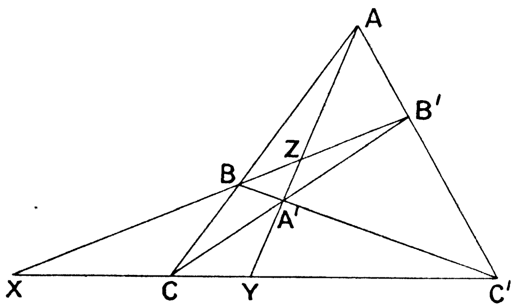

FIG. 10.

Let X, Y, Z be the points of intersection of the diagonals AA′, BB′, CC′ of the quadrilateral formed by the four straight lines ABC, AB′C′, A′BC′, A′B′C. Then the theorem states that each of the ranges

$$\{AA', YZ\}, \quad \{BB', ZX\}, \quad \{CC', XY\}$$

is harmonic. This is proved by projecting the range lying on one diagonal from each end in turn of a second diagonal on to the third.

Involution.

The harmonic property of the quadrilateral is a particular case of the more general *involution property of the quadrangle* (Mathews, p. 83) : *The pairs of opposite sides of a complete quadrangle cut any transversal in involution.* And this is in turn a particular case of a similar theorem on conics through four points (see p. 63). An *involution* is a set of pairs of points on a straight line, the two points of each pair being said to be *conjugate* to one another, such that the cross-ratio of any four of the points is equal to that of their four conjugates.

Let P, P′ ; Q, Q′ ; R, R′ be the intersections, with any transversal PQ, of the pairs of opposite sides AB, A′B′ ; AB′, A′B ; AA′, BB′ of a quadrangle ABA′B′. Then for example,

$$\{PQ, RR'\} = A\{BB', ZR'\} = A'\{Q'P', RR'\} = \{P'Q', R'R\},$$

and, similarly, the involution property can be verified for any four points of the range. If we now make the transversal coincide with CC′, we can identify P and P′ with C ;

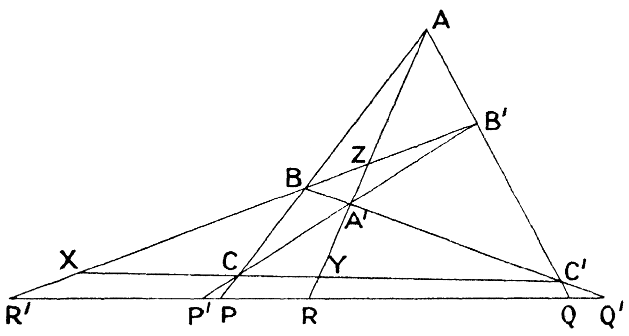

FIG. 11.

Q and Q′ with C′ ; R with Y ; R′ with X. The involution property just proved becomes {CC′, YX} = {CC′, XY}, which expresses that the range is harmonic.

If two pairs PP′, QQ′ of the involution are given, the quadrangle property enables us to construct the conjugate R′ of any given point R of PQ. Take any point A of the plane, not lying on the straight line PQR, and join AR; take any other point A′ on AR. Let AP, A′Q′ meet in B, and let AQ, A′P′ meet in B′; join BB′ and let it meet PQ in R′, which is the required conjugate to R. The point R′ so obtained is quite independent of the choice of the auxiliary points A, A′, and is completely determined by means of the equation $\{PQ, RR'\} = \{P'Q', R'R\}$ as the conjugate of R in the involution determined by the two given pairs of points PP′, QQ′.

Homography.

An involution is a particular case of a *homography*, defined as follows. Let there be a correspondence established between pairs of points P, P′; Q, Q′; R, R′; ... belonging respectively to two different ranges PQR..., P′Q′R′..., on the same or on different straight lines. Then if the cross-ratio of any four points of the first range is equal to that of the four corresponding points of the second range, the two ranges are said to be *homographic*, and the correspondence between the points is a homography. Two ranges in perspective give the simplest case of homography, but two homographic ranges need not be in perspective.

In order to define a homography completely, we need to be given three points P, Q, R of the first range and the three corresponding points P′, Q′, R′ of the second. Then if any point X is taken belonging to the first range, and X′ is the corresponding point of the second,

$$\{P'Q', R'X'\} = \{PQ, RX\};$$

that is, X′ is the point which with three given points makes a range of cross-ratio equal to a given cross-ratio, and it can be constructed with ruler only by the method of p. 39. Then as the variable point X of the first range moves along the straight line which bears that range, the corresponding point X′ moves along the straight line which bears the second range. If these two straight lines coincide, then X, X′ are in general different points of this line; when X is at P, then X′ is at P′; when X moves along and comes to the position P′, then X′ will in general take up some position P″ different from P′ and from P. But it may happen that for every position of P, this point P″ coincides with P; that is to say, to every point P′ of the straight line there corresponds the same point P, whether we regard P′ as belonging to the first range or to the second. Then the points of the straight line fall into conjugate pairs, and we need not specify which point of the pair belongs to either range; the cross-ratio of any four points is equal to that of their four conjugates, and the homography is an involution.

Common Points.

For example, let the two ranges (X), (X′) on the same straight line PQ be each in perspective, with centres of perspective C, C′ respectively, with the same range (Y)

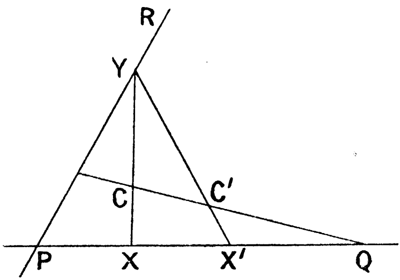

FIG. 12.

on the straight line PR. The three ranges are all homographic, and the common points of the first two are P the point of intersection of the two given straight lines PQ, PR, and Q the point of intersection of CC′, PQ.

As another example, the points of the same axis, whose Cartesian coordinates are x and λx, where λ is any constant, form homographic ranges. Then if p is the coordinate of P, that of P′ is λp and that of P″ is $\lambda^2 p$. If this homography is an involution, we must have $\lambda^2 p = p$ for all values of p, or $\lambda = \pm 1$. If $\lambda = +1$, the two ranges are identical; if $\lambda = -1$, each range is the reflexion of the other in the origin. In the homography, the origin corresponds to itself, and so does the point at infinity; in any homography on one straight line, or in any involution, there are always two points X_1, X_2, real or imaginary, which coincide with their corresponding or conjugate points X_1', X_2', but the construction requires the solution of a quadratic equation, and cannot be carried out with ruler only, except in special cases.

The two double points of an involution have the important property of dividing harmonically any pair of conjugates PP'; for if X_1, X_2 are their own conjugates, $\{X_1X_2, PP'\} = \{X_1X_2, P'P\}$, which is the characteristic property of a harmonic range. An involution can be defined as the set of pairs of points PP' which divide a given pair X_1X_2 harmonically.

We can prove analytically that any homography on one straight line has two common points. Suppose that the positions of X and X' are determined by their Cartesian coordinates x, x' measured from any fixed origin in the straight line. Then x, x' are related by some equation, and since x' can be obtained by a ruler construction when x is given, this equation must be soluble in the form $x' =$ a rational function of x, and therefore when simplified it is linear in x'; similarly, it is linear in x, and must therefore be of the form

$$axx' + bx + cx' + d = 0, \quad \text{(I)}$$

where a, b, c, d are constants, which gives

$$x' = -\frac{bx+d}{ax+c}, \quad x = -\frac{cx'+d}{ax'+b}.$$

If p is the coordinate of P, etc., where PQR, $P'Q'R'$ are the given points which determine the homography, we have $\{PQ, RX\} = \{P'Q', R'X'\}$, or

$$\frac{(r-p)(q-x)}{(q-r)(x-p)} = \frac{(r'-p')(q'-x')}{(q'-r')(x'-p')},$$

which can be rearranged in the form (I). The coefficients a, b, c, d are rational functions of the six given coordinates $p, \ldots r'$, and these coefficients, or any rational functions of them, can be constructed with ruler only from the six given points.

Now the common points X_1, X_2 are found by putting $x = x'$ in equation (I), and their coordinates x_1, x_2 are therefore the roots of the quadratic equation

$$ax^2 + (b+c)x + d = 0, \quad \text{(II)}$$

so that there are always two common points, real, coincident or imaginary, and finite or infinite.

In general, x_1, x_2 involve the square root $\sqrt{(b+c)^2-4ad}$, and so cannot be constructed with ruler only, except in particular cases, as for example when one root is known to be rational. If $a=0$, and also $b+c=0$, both roots are infinite. In this case the relation (I) between x and x' reduces to

$$x'=x+\frac{d}{b},$$

so that the second range is equal to the first displaced through a constant distance d/b; it is clear geometrically that there are no finite common points.

If the homography is an involution, the value of x' corresponding to any assigned value p of x is the same as the value of x when the value p is assigned to x'; in other words, the bilinear relation (I) is symmetrical in x, x', and the condition for this is $b=c$. Then the relation is

$$axx'+b(x+x')+d=0,$$

and the coordinates x_1, x_2 of the double points are the roots of the quadratic

$$ax^2+2bx+d=0.$$

The point O midway between the double points is called the *centre* of the involution; its coordinate is

$$x_0=\frac{x_1+x_2}{2}=-\frac{b}{a}.$$

We can change the origin to O by writing

$$x=\xi-\frac{b}{a},\quad x'=\xi'-\frac{b}{a},$$

and the relation between the new coordinates ξ, ξ' reduces to

$$\xi\xi'=\frac{b^2}{a^2}-\frac{d}{a}=\left(\frac{x_1+x_2}{2}\right)^2-x_1x_2=\left(\frac{x_1-x_2}{2}\right)^2$$
$$=e^2,\text{ say,}$$

where e is the distance of either double point from the centre.

The centre is defined by a metrical property, and is the conjugate of the point at infinity. It can be constructed with ruler only, provided we can draw parallels (p. 47). The involution property expressed in metrical language is that the product of the distances of any pair of conjugates from the centre is constant. If the involution is projected on to any other straight line, conjugate pairs project into conjugate pairs, and double points into double points, but the first centre does not project into the second centre, except in the special case when the point at infinity on the first line projects into the point at infinity on the second line. The constant e may be real or imaginary; if it is real, so are the double points, and conjugate points are on the same side of the centre; if e is pure imaginary, there are no real double points, and conjugate points are on opposite sides of the centre.

II. Metrical Properties.

(i) Lengths.

A figure can have metrical properties only when some of its elements have special positions. The harmonic property of the quadrilateral (p. 40) is still the basis of discussion. First let one of the diagonal points, Z of fig. 10 say, lie at infinity; then the straight lines AA′, BB′ are parallel, and Z may be considered to be defined as the

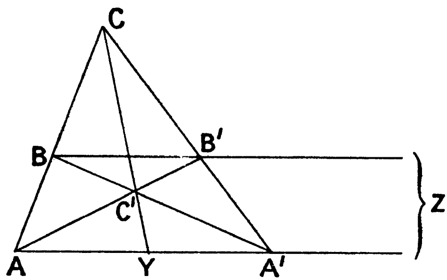

FIG. 13.

point of intersection of these parallels. But if we wish to define Z by referring only to points lying on AA′, we must fall back on the harmonic property expressed by the equation

$$\{AA', YZ\} = \frac{AY \cdot ZA'}{YA' \cdot AZ} = -1.$$

Since Z is at infinity, in this case the ratio $\frac{ZA'}{ZA}=\mathbf{1}$, and therefore

$$AY = YA'.$$

The point Z at infinity may be considered as the fourth harmonic of the three points A, A′ and the midpoint Y of the bisected segment AYA′. Thus a pair of parallels, and a bisected segment on one of them, are equivalent data. If AA′, BB′ are two given parallels, we can bisect a given segment AA′ on one of them: take any point B on the parallel BB′, and any point C on AB; then to complete fig. 13, join AC to meet BB′ in B′; join AB′, A′B to meet in C′; join CC′ to meet AA′ in Y, which is the required midpoint of AA′. Conversely, if we are given a bisected segment AYA′, we can draw a parallel to AA′ through any given point B: take any point C on AB; then to complete the figure, join A′B, YC to meet in C′; join AC′, A′C to meet in B′; join BB′, which is the required parallel to AA′ through the given point B. In the first case, we are given the point Z at infinity by means of the second straight line BB′ passing through it, and then we can construct the fourth point Y of the harmonic range AA′, YZ; in the second case, we are given Z by means of the other three points AA′Y of a harmonic range to which it belongs, and then we can draw BB′ to pass through Z. In either case, the point Z is determined, in the sense that we can draw the straight line joining it to any given point, that is, we can draw a parallel to AA′ through any given point. This is exactly what is meant by saying that a finite point, such as B′ in the last construction, is determined as the intersection of the two straight lines AC′, A′C: it is so determined that we can use it in the next step of that construction, which is to join it to B.

Infinity.

It is important to notice that if AA′ is produced to any *finite* point Z, the segment AZ is greater than the segment A′Z, because their difference AA′ is greater than zero *and* their ratio AZ/A′Z is greater than unity. For finite segments, the two properties of zero difference and unit ratio are inseparable, and each implies that the segments are equal. The converse is what is expressed by saying that the whole is greater than the part; this is not so much an axiom as the *definition of finiteness*. As Z recedes along AA′, the difference AZ − A′Z remains finite and constant, being equal to AA′; but the ratio AZ/A′Z is always greater than unity and tends to unity as its limit. When Z takes up its position at infinity, this ratio passes to its limit and is equal to unity, the difference remaining finite; but both the segments AZ and A′Z are infinite. They have become equal, which for infinite segments means that their ratio is infinite and not necessarily that their difference has any special value. The two properties of unit ratio and zero difference have become separable, and this fact is the *definition of infiniteness*: if from a whole a part is removed, and the whole is equal (that is, in a unit ratio) to the remaining part, then the whole and this part are both infinite.

Addition of Segments.

When the point at infinity on AA′ is determined, so that we can draw parallels to AA′, then besides bisecting any segment, we can transfer it to any other part of the straight line. Let it be required to lay off a segment equal to AA′ from any point B of AA′ or AA′ produced.

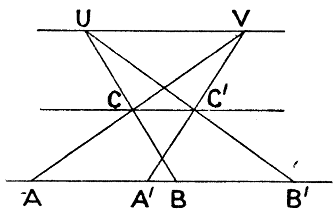

FIG. 14.

To AA′ draw any two parallels CC′, UV; take any point V on one of these, and join VA, VA′ to cut the other parallel in C, C′; join BC to meet UV in U, and join UC′ to meet AB in B′. Then B′ is the required point, for by similar triangles,

$$\frac{BB'}{CC'}=\frac{BU}{CU}=\frac{AV}{CV}=\frac{AA'}{CC'},$$

so that

$$BB'=AA'.$$

If we had wished to lay off the segment on the other side of B, we should have joined BC′ instead of BC.

Thus a bisected segment on a given straight line enables us to construct the sum or difference of any two given segments of that line, or, by repetitions, any linear function with integral coefficients of any number of segments of the same straight line.

For example, given a bisected segment ABA′ and another point C of the straight line AB, let it be required to lay off from a given point X of AB a segment XY of length $3\mathfrak{b} - \mathfrak{c}$, where $AB = \mathfrak{b}$, $AC = \mathfrak{c}$.

We might first draw two parallels to AB by means of the bisected segment ABA′; then transfer the segment CA′ to the position FA, and the segment FB to the position XY. The whole construction may be arranged more compactly as follows.

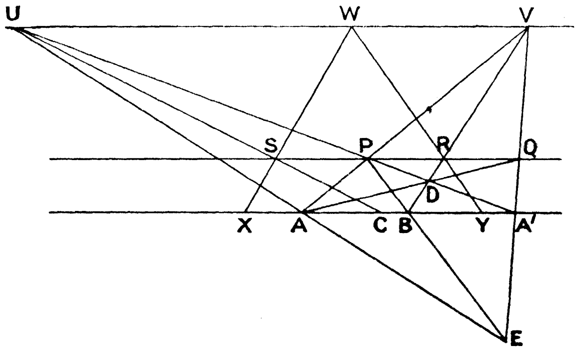

FIG. 15.

Draw any straight line APV through A, and take any two points P, V on it; join PA′, VB to meet in D, and join DA, VA′ to meet in Q. Join PB to meet VA′ in E, and join EA to meet PA′ in U. Then PQ, UV are each parallel to AB. Let VB meet PQ in R. Join UC to meet PQ in S, join XS to meet UV in W, join WR to meet AB in Y. Then XY is the required segment; the proof is left to the reader.

If we are only given two segments AA′, BB′ of the straight line, which are equal but not adjacent, we cannot draw parallels. Suppose, for example, that the four points occur in the order named above, so that the given segments do not overlap. Then taken in pairs AA′, BB′, they determine an involution whose centre is Y the midpoint of AB′ or of A′B, and whose double points are real; taken in pairs AB, A′B′, they determine an involution with Y as

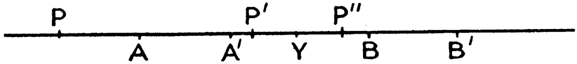

FIG. 16.

centre and imaginary double points; taken in pairs AB′, A′B, they determine an involution with its centre and one double point at infinity, and with Y as the other double point. If any point P is conjugate to P′ in the first of these involutions and to P″ in the second, then P′, P″ are conjugate in the third, and are equidistant from Y. Thus we can construct the reflexion of any point in Y, but we cannot construct Y itself, nor the point Z at infinity on AA′, for Y, Z are given by two unknown roots of a quadratic equation. But if the given segments are adjacent, A′, B, Y all coincide, one root of the quadratic has a known rational value and Z can be constructed.

But if we are given three equal segments AA′, BB′, CC′, these three pairs of points determine a homography in

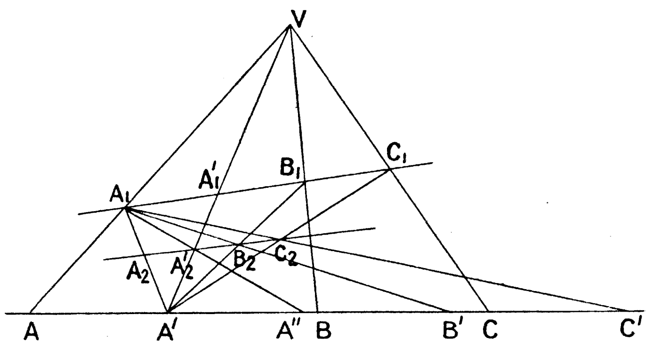

FIG. 17.

which the distance between corresponding points is con stant (p. 45). We can therefore construct the point A″, which corresponds to A′ when A′ is regarded as a point of the first range ABC; we make {A′B′C′A″} = {ABCA′}, by the construction of p. 39. Then $A'A'' = AA'$, and we have a bisected segment $AA'A''$, and can therefore draw parallels. The actual construction of A'' is shown in fig. 17. The range $ABCA'$ is first projected from any vertex V into $A_1B_1C_1A_1'$ on any straight line; then from A' as vertex into $A_2B_2C_2A_2'$, and then from A_1 into $A'B'C'A''$.

Parallelogram Given.

But none of these data help us to draw parallels to any straight line in a direction different to that in which we are given two parallels or a rationally divided segment. In other words, we are only given one point on the straight line at infinity, namely the point in which it meets the given parallels, and so we cannot consider it as fully determined, nor can we obtain the point in which it meets any other straight line, which is the point at infinity on the latter, and so we cannot draw parallels in a second direction. But if we are given two pairs of parallels in different directions, then we have two points on the straight line at infinity, which is therefore completely determined, and we can obtain its intersection with any other straight line, and so draw parallels in any direction. Since two pairs of parallels in different directions are the sides of a parallelogram, we may say that *if a parallelogram is given, parallels can be drawn to any straight line with ruler only.* This is the theorem which was assumed in chapter II, p. 14. In fact, if we have, or can obtain, a bisected segment BYB', and if we can also draw through these three points three straight lines BX, YZ, $B'X'$ all parallel to a second direction, these three equidistant parallels meet any straight line EF in a bisected segment XZX', by means of which we can draw parallels to EF. When a parallelogram is given, each diagonal is bisected by the other; we can draw parallels to one diagonal through the ends of the other, and these cut off from an arbitrary straight line a segment which is bisected by the parallel diagonal.

In the figure, AYA', BYB' are the diagonals of the given parallelogram $ABA'B'$, bisecting one another in Y. Using the bisected segment AYA', we draw through B a straight line BX parallel to AA' by the construction of p. 47; the sides AB, $A'B$ of the given parallelogram are used as sides of the complete quadrilateral required, and AA' is one diagonal. We therefore draw any straight line YC_1C through Y to meet AB, $A'B$ in C, C_1; join $A'C$, AC_1 to meet in B_1, and join BB_1, which is the required parallel. In constructing the other parallel through B', we can use the

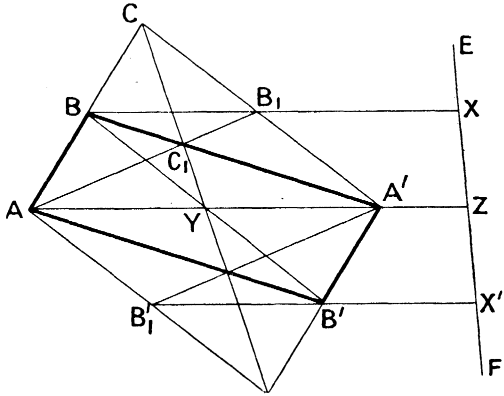

FIG. 18.

quadrilateral which is the reflexion of the first in the point Y, which includes the straight lines AB', $A'B'$, CY produced, and the third parallel $B'B_1'$. Then these three parallels meet any straight line EF in points XZX', which furnish a bisected segment, by means of which we can draw parallels to EF.

Rotations.

Then, as we saw in chapter II, we can perform any rational operations upon the Cartesian coordinates of the given points, including exchange of abscissa and ordinate; but these are not absolute lengths, but ratios of segments of the axes to the unit segments of the same axes; we cannot assume that these units are equal, or that the scale is the same for ordinates as for abscissae; we can transfer coordinates from one axis to the other, but we cannot transfer lengths. If we are able to choose equal lengths as units along the two axes, this is due to a separate, additional metrical relation between the data. When it is given, we can at once draw a rhombus, which is the special form assumed by the unit parallelogram, and its diagonals are a pair of straight lines at right angles. We can now transfer a length from a straight line in any direction to its reflexion in either diagonal of the unit rhombus; for the distance between the points whose coordinates are $(\mathfrak{p}, \mathfrak{q})$ and $(\mathfrak{r}, \mathfrak{s})$ is equal to the distance between $(\mathfrak{q}, \mathfrak{p})$ and $(\mathfrak{s}, \mathfrak{r})$. But unless the angle α between the axes has some special value, we cannot do more. For if the distance $\mathfrak{d}$ between the two points $(\mathfrak{p}, \mathfrak{q})$ and $(\mathfrak{r}, \mathfrak{s})$ could be transferred to an axis, $\mathfrak{d}$ would belong to the domain $[1, \mathfrak{a}, \mathfrak{b}, \ldots]$ determined by the coordinates of the given points, to which domain $\mathfrak{p}$, $\mathfrak{q}$, $\mathfrak{r}$, $\mathfrak{s}$ also belong. But the formula for the distance between two points in oblique coordinates is

$$\mathfrak{d}^2 = (\mathfrak{r} - \mathfrak{p})^2 + (\mathfrak{s} - \mathfrak{q})^2 + 2(\mathfrak{r} - \mathfrak{p})(\mathfrak{s} - \mathfrak{q}) \cos \alpha,$$

which gives a value for $\cos \alpha$ belonging to the same domain, except in the cases $\mathfrak{r} = \mathfrak{p}$ or $\mathfrak{s} = \mathfrak{q}$, when the segment $\mathfrak{d}$ is already parallel to one or other axis. But in general, the cosine of the angle between the axes is not rational, nor a rational function of the given coordinates; if this is so, we have a special case.

For example, if OA = OB and $\cos \alpha = \frac{1}{2}$, $\alpha = \frac{1}{3}\pi$, we have an equilateral triangle OAB, which is half the unit rhombus, and by drawing parallels, we can construct a regular hexagon ABCDEF, which gives three pairs of perpendiculars,

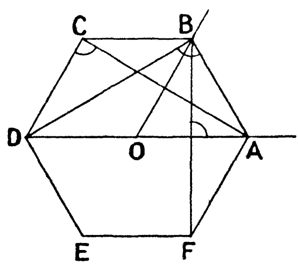

FIG. 19.

AB, BD; AC, CD; AD, BF. We can now reflect any segment in any of the diagonals of the hexagon, and two such reflexions give a rotation through an angle $\frac{2}{3}\pi$; we can erect an equilateral triangle on each side of any given segment, and by joining the vertices obtain a perpendicular, so that we can draw right angles and trisect them, but not bisect them.

If the units are not equal, but the axes are rectangular, then the unit parallelogram is a rectangle, and its diagonals give a new pair of axes, not at right angles, along which we have equal segments; so that the case of rectangular axes with unequal units is equivalent to the case of oblique axes with equal units. But if we are given both these metrical data at once, the unit parallelogram is a square, and its diagonals give a second pair of rectangular axes with equal units. We can reflect any segment in the sides and in the diagonals of the unit square, and therefore turn it through a right angle; for the straight line joining the points whose coordinates are (**p**, **q**) and (**r**, **s**) is now equal and perpendicular to that joining (**q**, **r**) to (**s**, **p**). We can erect a square on any given segment, and draw its diagonals; we can draw right angles and bisect them, but not trisect them.

(ii) **Angles.**

The chief interest of metrical properties of angles is based on the right angle, and the theory of right angles can be worked out in projective language, in terms of pencils in involution. The involution property involves cross-ratios only, and therefore belongs equally to a range of points on a straight line and to a pencil of rays through a point, by the principle of duality (Mathews, chap. I). Reciprocal to the involution property of the quadrangle (p. 41) there is the *involution property of the quadrilateral: The pairs of opposite vertices of a complete quadrilateral subtend a pencil in involution at any point.* The proof of p. 41 for the reciprocal theorem applies to this one also, if we interchange "point" and "straight line" wherever they occur; and just as in the other theory, we can construct with ruler only the conjugate of any given ray in the involution determined by two given pairs of rays at the same vertex.

Orthogonal Involution.

The application of the theory of involution to right angles is based on the theorem: *Pairs of straight lines at right angles through any point form a pencil in involution,* called the *orthogonal involution* at the point. For if OA′ is perpendicular to OA, etc., the pencil O{A′B′C′D′} is the same as O{ABCD} turned through a right angle; the angles between corresponding pairs of rays are equal, and the two pencils have the same cross-ratio. But this involution is not determined unless two different pairs of straight lines at right angles are given, or can be drawn, at the same vertex O; then we can construct the perpendicular to any other straight line through O, as its conjugate in the involution determined by the two given pairs.

The accompanying figure gives the actual ruler construction of the perpendicular to a given straight line **r** through a point O at which two right angles are given.

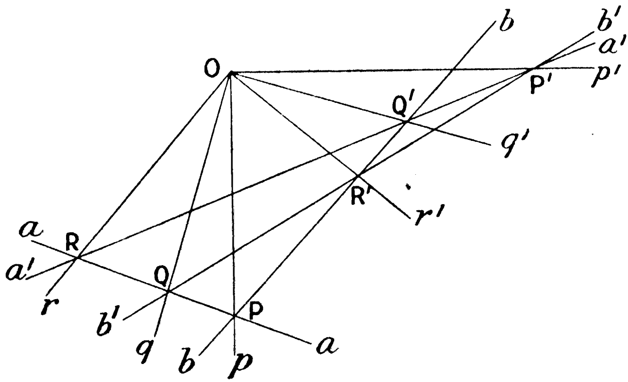

FIG. 20.

Small letters denote straight lines; **p**, **p**′; **q**, **q**′ are the two given pairs of straight lines at right angles, and **r** the straight line to which a perpendicular **r**′ is required. Take any straight line **a** of the plane, not passing through the common vertex O, and let it meet **p**, **q**, **r** in P, Q, R; take any straight line **a**′ through R and let it meet **p**′, **q**′ in P′, Q′. Draw the straight lines **b** joining PQ′ and **b**′ joining P′Q; let **bb**′ meet in R′. Then the straight line **r**′ joining OR′ is the required perpendicular to **r**. For **p**, **p**′; **q**, **q**′; **r**, **r**′ are the pairs of straight lines joining O to the pairs of opposite vertices P, P′; Q, Q′; R, R′ of the complete quadrilateral **aba′b′**, and are therefore in involution; and since two pairs of rays are at right angles, this is the orthogonal involution at O, and the two rays of every other conjugate pair, including r, r′, are also at right angles to each other. The figure is lettered so as to be reciprocal to fig. 11. The steps of the two constructions correspond exactly (see p. 42).

Focus of a Parabola.

As another example, consider the problem: to find the focus of a parabola touching four given straight lines. Since the focus is uniquely determined, this is possible by a ruler construction. The form of the question implies that the metrical data are given. For a parabola is distinguished from other conics by the fact that it touches the straight line at infinity, or by some equivalent property, which implies that the straight line at infinity is

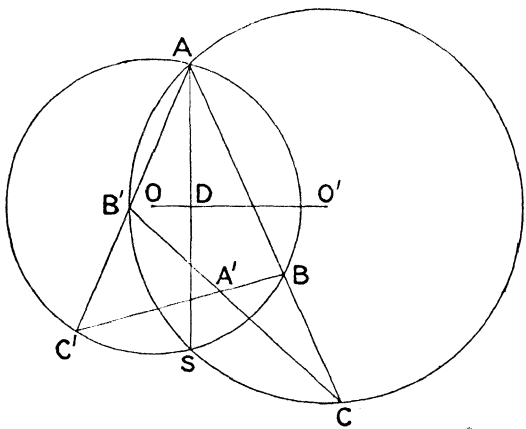

FIG. 21.

determined. And any definition of a focus involves right angles directly or indirectly. The usual focus and directrix definition of a parabola is as the locus of a point P such that its distance from a given point S is equal to its perpendicular distance PN from a given straight line NX. This assumes that we have the means of comparing a segment in an arbitrary direction PS with one in a fixed direction PN, so that we must have some instrument beyond a ruler, equivalent to an Einheitsdreher (p. 71) ; and we suppose that this has been used first of all to provide the metrical data for drawing parallels and perpendiculars.

In order to construct the focus S from the four given tangents AB, AB′, A′B, A′B′, we can use the known property that S lies on the circumcircle of each of the four triangles ABC′, etc., formed by three out of the four tangents, and it therefore lies on the common chord of each pair of these circles. We can construct these common chords without actually describing the circles, because each pair of circles has one common point already given. Find the centres O, O′ of the circles ABC′, AB′C in Euclid's way, by bisecting the sides at right angles. Draw AD perpendicular to OO′, and produce it to S, making DS = AD ; then AD is the common chord, and S is the required focus.

Summary of the Theory of Ruler Constructions.

When no metrical data are given, we can carry out all rational operations upon the coordinates of the given points, these coordinates being projective, that is, expressed as cross-ratios.

If we are given that two straight lines are parallel, or that one segment is bisected, or one equivalent metrical fact, we can replace the corresponding coordinate by a Cartesian coordinate, which is a ratio of lengths instead of a cross-ratio. We can draw parallels in one direction, and carry out rational operations upon lengths of segments of any one of these parallels, but we cannot transfer lengths from one parallel to another. A single point at infinity is determined, but not the whole straight line at infinity. If we are given a separate metrical fact with regard to each of two different directions, we can operate upon a Cartesian system of two coordinates ; we can draw parallels in any direction, and transfer a length from one parallel to another, but not from one direction to another. The straight line at infinity is completely determined.

If we are also given one right angle, or two equal lengths in different directions, we can choose a pair of oblique axes with the same unit of length, and we can reflect any length in the arms of the right angle. We cannot draw right angles with their arms in other directions.

If two pairs of parallels and two right angles are given, we can draw any parallels and perpendiculars. We can reflect a length in any axis, but we cannot transfer it into an arbitrary direction ; the straight line at infinity is determined. This gives the greatest powers that can be obtained with ruler only. Since the sides of a square and its diagonals give two pairs of parallels and two right angles, *a square gives all the metrical data for ruler constructions.*

CHAPTER IV.

RULER AND COMPASS CONSTRUCTIONS.

I. Ruler and Compasses.

WHEN compasses are used as well as ruler, we can obtain at once all the metrical data required for that part of the construction in which the ruler is used. When a circle is drawn with a given centre, any diameter gives a bisected segment, so we can draw parallels to any straight line; and any two diameters give the corners of a rectangle, so we can draw right angles; and all the radii are equal, so we can have the same unit of measurement in all directions. We can also, by the use of ruler and compasses, bisect an arbitrary angle, but not trisect it, and carry out all the other constructions of the first six books of Euclid.

As we have seen, the steps of a ruler and compass construction which require compasses are each equivalent to the solution of a quadratic equation.

(i) The solutions of the equation

$$\mathbf{x}^2 = \mathbf{a}, \quad \text{or} \quad \mathbf{x} = \pm\sqrt{\mathbf{a}},$$

can be constructed as the mean proportionals between **a** and **1** by means of an ordinate in a circle on **1 + a** as diameter.

Take a unit length AO and produce it to Q, making $OQ = a$; draw a circle on AQ as diameter, and let it meet the perpendicular at O to AQ in X_1, X_2.

Then $$OX_1 = \sqrt{a}, \quad OX_2 = -\sqrt{a}.$$

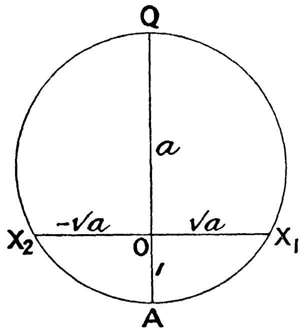

FIG. 22.

(ii) The solutions of the equation

$$x^2 - 2px + q = 0, \quad \text{or} \quad x = p \pm \sqrt{p^2 - q},$$

can be constructed by laying off a length $\sqrt{p^2 - q}$ in both senses from the end of a segment of length p. If OX_1,

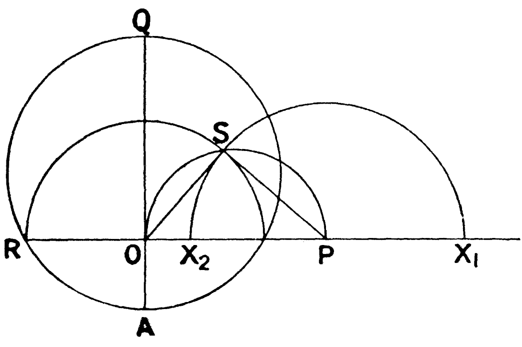

FIG. 23.

OX_2 are the roots, and $OP = p$, all measured along the same axis, P is the midpoint of X_2X_1, and $PX_1 = -PX_2 = \sqrt{p^2 - q}$.

We therefore take a unit length AO and produce it to

Q, making $OQ = q$; draw a circle on AQ as diameter, and let it meet the perpendicular at O to AQ in R; then $RO = \sqrt{q}$. Produce RO to P, making $OP = p$, and draw a circle on OP as diameter to meet the circle O(R) in S; then OSP is a right-angled triangle, with hypotenuse $OP = p$ and one side $OS = OR = \sqrt{q}$; therefore the other side $PS = \sqrt{p^2 - q}$. Describe the circle P(S) to meet OP in X_1, X_2; then the distances of X_1, X_2 from O are the roots of the given equation.

The construction fails if the circle O(R) does not meet the semicircle OSP in real points; this can only be when the latter lies entirely within the former, and $p < \sqrt{q}$; then the surd is imaginary.

V. Staudt has given a prettier and less obvious construction.

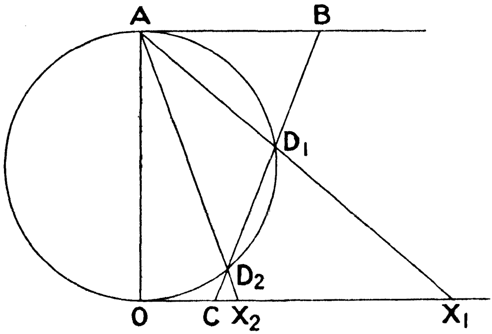

FIG. 24.

Take a circle of unit radius, and from two parallel tangents OC, AB cut off lengths $OC = \frac{q}{2p}$, $AB = \frac{2}{p}$, measured from the points of contact O, A. Join BC to cut the circle in D_1, D_2; join AD_1, AD_2 to cut OC in X_1, X_2; then OX_1, OX_2 are the roots of the given equation.

This can be verified analytically. Take OC, OA as axes of x and y; then the equation of AB is $y = 2$,
and of the circle OD_2D_1A, $x^2 + y^2 - 2y = 0$.

If X_1 has the coordinates $(x_1, 0)$, where x_1, x_2 are the roots of the given quadratic, the equation of AX_1 is

$$\frac{x}{x_1} = \frac{2 - y}{2};$$

this meets the circle again in D_1, whose coordinates we find to be $\frac{4x_1}{x_1^2+4}, \frac{2x_1^2}{x_1^2+4}$, and similarly for D_2.

The equation of D_1D_2 is therefore

$$\begin{vmatrix} x, & y, & 1 \\ 4x_1, & 2x_1^2, & x_1^2+4 \\ 4x_2, & 2x_2^2, & x_2^2+4 \end{vmatrix} = 0.$$

But since x_1, x_2 are the roots of $x^2-2px+q=0$, we may write $2p$ for x_1+x_2 and q for x_1x_2; by use of these relations, the equation of D_1D_2 may be reduced to the form

$$4(px-y)+q(y-2)=0.$$

To find the abscissa of B, put $y=2$; then $x=AB=\frac{2}{p}$;

to find the abscissa of C, put $y=0$; then $x=OC=\frac{q}{2p}$; which justifies the construction.

Conics through Four Points.

Just as there is a fundamental cross-ratio property of the quadrilateral, which underlies the discussion of linear constructions, so there is a cross-ratio property of the circle, which underlies much of the theory of quadratic constructions: *Four fixed points of a circle subtend a pencil of constant cross-ratio at a variable fifth point of the circle.* This follows from the fact that the angle subtended by any two of the fixed points at the variable point is constant as the latter moves round the circumference. Hence the angles between the rays of the pencil are constant, and since the cross-ratio can be expressed in terms of the sines of these angles, the cross-ratio of the pencil is constant also. Thus we can give a meaning to the cross-ratio of four points upon a circle, defining it as the cross-ratio of the pencil subtended by the four points at any fifth point of the same circle. If four points lie upon a straight line, they subtend a pencil of constant cross-ratio at any vertex whatever; if they do not lie upon a straight line, the cross-ratio can have any value if the vertex can be any point of the plane; but if the four points all lie on a circle, and the vertex is taken to lie on the same circle, we get a definite, constant value for the cross-ratio.

This is a projective property, and is therefore true of any conic; for a conic can always be projected into a circle. Thus we have the *projective definition of a conic*, as the locus of a point at which four fixed points subtend a pencil of fixed cross-ratio.

From this there follows a theorem which is a generalization of the involution property of the quadrangle. In the first place: If a quadrangle is inscribed in a circle, the two intersections of any straight line with the circumference are a pair of conjugate points of the involution in which the pairs of opposite sides cut the chord. For let PP′ be

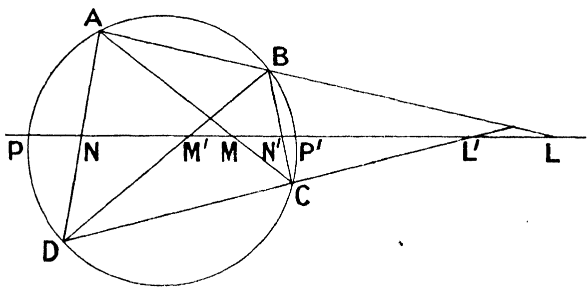

FIG. 25.

the chord and ABCD the quadrangle, with intersections as in fig. 25. Then

$$\{PLMP'\} = A\{PBCP'\} = D\{PBCP'\} = \{PM'L'P'\} = \{P'L'M'P\},$$

so that PP′ belong to the involution determined by the pairs LL′, MM′. This is a projective property, and is therefore true of the range cut out on a transversal by the sides of a quadrangle and any conic through the four vertices; hence all such conics cut the transversal in pairs of conjugate points of the same involution determined by the pairs of opposite sides. But these pairs of sides are themselves particular, degenerate cases of conics through the four points, so the theorem can be stated in the simple and general form: *Conics through four fixed points cut any transversal in involution.*

Hence if five points ABCDE on a conic are given, the other point of intersection X of any straight line EX through one of the five points is determined as the conjugate of E in the involution cut out on EX by the sides of the quadrangle ABCD. It can therefore be constructed with ruler only. Thus we have a ruler construction for any number of points on the conic through five given points, for we can find its intersection with any number of straight lines through any of the given points. Thus with ruler only, and without calculation, we can plot the curve with as much accuracy as time, patience and other limitations permit, and this is the nearest approach to actually drawing the curve that we can ever get with Euclidean instruments.

But if we wish to find the two intersections of a conic with an arbitrary straight line not passing through any of the five points, the problem cannot be solved with ruler only. Let it be required to find the intersections XY of a given straight line with the conic through five given points ABCDE. Take any point P on XY; join AP and determine

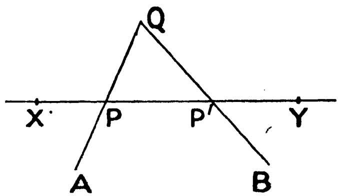

FIG. 26.

its other intersection Q with the conic; join BQ to meet XY in P′. Then if P lies on the conic, it coincides with Q and also with P′; and conversely, if P coincides with P′, it lies on the conic. If not, take any four positions P_1, P_2, P_3, P_4 of P, and let {P} stand for their cross-ratio, with a similar notation for Q, P′. Then, since $ABQ_1 \ldots Q_4$ lie on the same conic,

$$\{P\} = A\{Q\} = B\{Q\} = \{P'\},$$

so that P, P′ describe homographic ranges, and the points required are their common points X, Y, which demand the quadratic construction given below. But in the case in which XY passes through one of the five given points, E say, then one of the common points is E, and the other only demands a ruler construction, as we have just seen.

To the cross-ratio property of points on a conic there corresponds the cross-ratio property of its tangents: *Four* *fixed tangents to a conic cut out a range of constant cross-ratio on a variable fifth tangent to the conic.* And there is a construction for tangents to the conic touching five fixed straight lines, which corresponds step by step to any construction for points on the conic through five fixed points.

Common Points of a Homography.

In order to complete this discussion, we must go back to the theory of homography, and give the construction for the common points of two homographic ranges. Let there be given two homographic ranges $PQR\ldots$, $P'Q'R'\ldots$ on the same straight line PQ, whose common points X, Y are required. Take any circle, and any point V upon it; with V as vertex project the ranges into $\mathbf{pqr}\ldots$, $\mathbf{p'q'r'}\ldots$ on

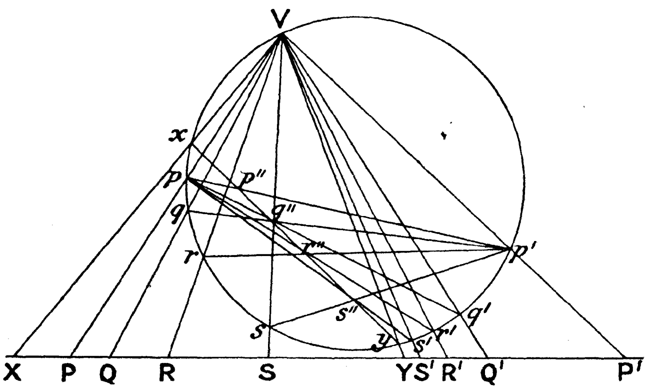

FIG. 27.

the circle. Now if three pairs of corresponding points P, P'; Q, Q'; R, R' are given on the straight line, we know that the point S' corresponding to any point S is determined; and it may be found thus.

Join $\mathbf{pq'}$, $\mathbf{p'q}$ to meet in $\mathbf{q''}$; join $\mathbf{pr'}$, $\mathbf{p'r}$ to meet in $\mathbf{r''}$; join $\mathbf{q''r''}$ to meet $\mathbf{pp'}$ in $\mathbf{p''}$, to meet $\mathbf{p's}$ in $\mathbf{s''}$, and the circle in $\mathbf{x}$, $\mathbf{y}$. Join $\mathbf{ps''}$ to meet the circle in $\mathbf{s'}$, and join $V\mathbf{s'}$ to meet PQ in S'. Then S' is the point corresponding to S; for

$$\{PQRS\} = V\{\mathbf{pqrs}\} = \mathbf{p'}\{\mathbf{pqrs}\} = \{\mathbf{p''q''r''s''}\}$$
$$= \mathbf{p}\{\mathbf{p'q'r's'}\} = V\{\mathbf{p'q'r's'}\} = \{P'Q'R'S'\}.$$

We have not only the two given homographic ranges on PQ, but also the two ranges $\mathbf{pqr}$... , $\mathbf{p'q'r'}$... on the circle, homographic with the former and with each other; and the common points on the straight line project into common points on the circle. But the common points on the circle are $\mathbf{x}$, $\mathbf{y}$, where $\mathbf{q''r''}$ meets the circumference, and therefore the common points of the given homography are X, Y, the projections from V of $\mathbf{x}$, $\mathbf{y}$.

One common point X is at infinity if V$\mathbf{x}$ is parallel to PQ; in this case any problem that requires a finite common point of the homography has only one solution, and demands only a linear construction. Both X and Y are at infinity if $\mathbf{q''r''}$ touches the circle where it is met again by the parallel to PQ through V. In general, the common points are real, coincident, or imaginary according as $\mathbf{q''r''}$ cuts the circle in real points, touches it, or does not meet it.

The construction is simpler if the homography is given on a circle instead of on a straight line, as then we do not need the projection from V. One particular case is historic. Let the range P′Q′R′... be equal to the range PQR... shifted round the circumference through a constant angular distance α. Then PQ′, P′Q are parallel, $\mathbf{q''}$ is at infinity,

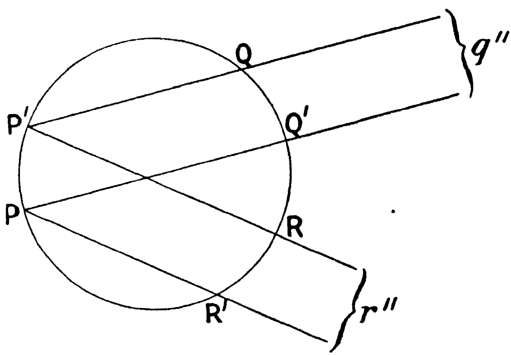

FIG. 28.

and so is $\mathbf{r''}$ and the whole straight line $\mathbf{q''r''}$. The common points X, Y are the intersections of the circle with the straight line at infinity, and are imaginary. Thus there are no real common points except in the case $\alpha = \mathbf{0}$ or $\mathbf{2n\pi}$, when the two ranges coincide, and every point is the same as its corresponding point.

Porisms.

Now suppose, for example, that we wish to draw a triangle whose corners lie on one circle, and whose sides touch a concentric circle. Take any point P on the outer circle, and draw successive tangents PQ, QR, RP′ to the inner circle to meet the outer circle again in Q, R, P′. Then

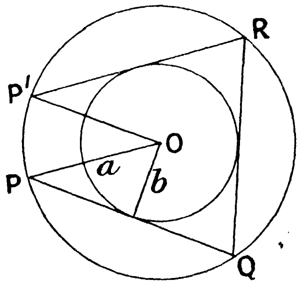

FIG. 29.

if P, P′ coincide, PQR is the required triangle, and the problem is solved; if not, P and P′ describe homographic ranges on the circle, and the solution is given by the common points. But if O is the centre of the circles, and $\mathbf{a}$, $\mathbf{b}$ the outer and inner radii, $\angle POP' = 6\cos^{-1}(\mathbf{b}/\mathbf{a}) =$ constant; the homography is of the kind just mentioned, and no real positions of P, P′ can coincide, unless we have the condition

$$6\cos^{-1}\frac{\mathbf{b}}{\mathbf{a}} = 2n\pi, \quad \text{or} \quad \mathbf{a} = 2\mathbf{b},$$

and in that case, P, P′ always coincide, and there are an infinite number of triangles of the kind required, one with a corner at any point P of the outer circle. The problem is *poristic*, that is to say, although the number of independent conditions is equal to the number of variables, yet these conditions are in general inconsistent, and there is no solution, except in the case when a special relation holds among the constants, when the conditions are no longer independent, and there are an infinite number of solutions. The same is true when the triangle is replaced by a polygon of any given number of sides; it can also be extended to any pair of circles, and, by projection, to any pair of conics: *The problem of circum-inscribing a polygon to two given conics is poristic.*

Double Points of an Involution.

In the case of an involution, the construction for the double points can be simplified, for one pair of conjugate points of the involution, taken in reverse order, gives a second pair of corresponding points of the homography. In fig. 27 we can use Q′, Q, in that order, in place of R, R′, and the figure reduces to fig. 30, in which q″r″ is a diagonal of the quadrilateral whose sides are pq, p′q′, pq′, p′q.

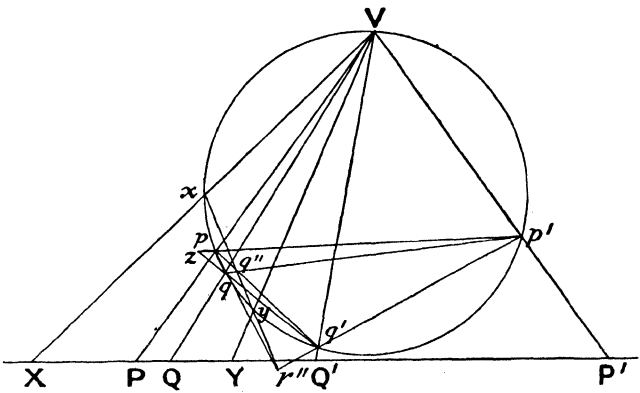

Fig. 30.

This figure enables us to see when two given pairs of points PP′, QQ′ determine an involution whose double points are real. This is so if q″r″ meets the circle in real points. Let pp′, qq′ meet in z; then since q″r″ divides harmonically the two chords zpp′, zqq′ through z, therefore q″r″ is the polar of z, and meets the circle in real points if z is outside the circle and in imaginary points if z is inside the circle.

Now z is inside the circle, provided q, q′ lie one on each of the two arcs into which p, p′ divide the circumference. In this case, on whatever arc V may lie, one and only one of the straight lines Vq, Vq' lies between Vp and Vp', so that one and only one of Q, Q' lies between P and P'. Hence *the double points are imaginary if the given pairs overlap*; and are real if one pair lies wholly between the other pair (as when the points are in the order $PQQ'P'$), or when each pair is outside the other pair (as $PP'QQ'$).

The graphical solution of any quadratic equation can be made to depend on finding the double points of an involution. The general equation

$$x^2 - 2px + q = 0$$

can be regarded as the equation for the coordinates of the double points of an involution, in which the coordinates x, x' of any pair of conjugate points are connected by the relation

$$xx' - p(x + x') + q = 0.$$

We can at once obtain two pairs of conjugates; for example,

$$x = 0,\ x' = \frac{q}{p}; \qquad x = 1,\ x' = \frac{q-p}{p-1}.$$

If p, q are given, we can find these four points by a ruler construction, and thence find the double points of the involution by the construction given above, and so solve the quadratic equation.

One Fixed Conic.

Now that construction only requires one circle to be drawn, which is an arbitrary circle; and only uses projective properties of that one circle, namely the cross-ratio property and the property of meeting any straight line in two points. Hence the circle can be replaced by any conic actually drawn; the rest of the construction requires a ruler only. From this it follows that if a single conic of any species is drawn once for all, we can find the double points of any involution of which two pairs of conjugates are given, and hence we can construct the roots of any quadratic equation with given coefficients. But this is exactly the power which the use of compasses adds to the use of a ruler; which proves that *a single fixed conic can completely replace the use of compasses.* The case in which the fixed conic is a circle is discussed in chapter VII, where some actual constructions are given.

II. Other Elementary Instruments.

Geometrical construction of to-day is a very different thing from what it was in the time of Euclid. A modern schoolboy collects far more apparatus, when he sits down to write out the theorem of Pythagoras, than its discoverer ever possessed; but both for accuracy and for general utility, the two Euclidean instruments still hold the field. Moreover, most other modern instruments depend for their manufacture on ruler and compass constructions, and in all their commoner uses they do not enable us to draw any figure that could not have been drawn, if enough time and trouble were taken, by Euclid's methods.

In this section we shall refer to a few of the instruments that are commonly used along with ruler and compasses. It would carry us too far to describe the devices that extend the range of geometric constructions to the solution of cubic, higher and transcendental problems, such as trisectors and integraphs. But there are three instruments that call for some notice, as providing short cuts in Euclidean constructions, without adding to the range; these are *dividers*, *parallel ruler*, and *set-square*. We shall determine the extent of their powers by means of the analytical ideas of chapter II.

(i) Dividers.

When we speak of a pair of compasses as one of Euclid's instruments, we have to remember that the ordinary modern article is a complex instrument, combining the powers to do two quite different things. First we may make the two legs stand upon two given points A, B, whose distance apart is equal to the radius of the required circle; then we pick up the compasses and set them down again with one leg standing upon a given point C, which is to be the centre of the circle, and then we describe the circumference with the other leg. Now when Euclid marked his circles, probably on the sand, he required to be given the centre C and a point A on the circumference of the required circle before he could describe it. It was not sufficient to have the radius AB given in any other position; one of the points, B say, had to coincide with C.

It is true that Euc. I. 2 removes this restriction, so that, as far as regards the possibility of a construction, there is no difference between the two forms of compasses; but the figure of Euc. I. 2 contains three straight lines and four circles, and though a simpler figure of five circles may be used instead (see p. 132), there is still a great difference between the two forms of instrument as regards the simplicity of a construction.

Mascheroni, who wrote at the end of the eighteenth century, found that the makers of scientific instruments of his day, in graduating the reading circles of telescopes where great accuracy was required, preferred to use compasses as far as possible, as being more accurate than rulers. He regarded it as a special feature of a pair of compasses that it retained accurately a radius once taken up with it, until he chose to alter it (" compas fidèle "). Where the same radius occurred more than once in a construction, he kept a separate pair of compasses for it, laying them aside until they were wanted again; for there was less chance of error from the compasses slipping than there was from having to take up the radius afresh. We shall use the terms *Euclidean* and *modern* compasses where it is necessary to make the distinction; there are quite a large number of constructions in which the modern use has no advantage over the ancient. Modern compasses combine the power of describing circles, which belongs to Euclidean compasses, with the power of carrying distances, which belongs to a separate instrument called *dividers*, whose elementary operation is exactly that described in Euc. I. 2: " from the greater of two straight lines to cut off a part equal to the less."

All that can be done with ruler and dividers can be done, possibly with greater labour, with ruler and Euclidean compasses. The converse is not true; ruler and dividers can do more than ruler alone, but not so much as ruler and compasses.

Einheitsdreher.

Dividers give all the metrical data for drawing parallels and perpendiculars; for we can draw two straight lines intersecting at O say, and lay off along them equal lengths AO, OA′, BO, OB′, and so obtain the corners of a rectangle

ABA′B′. This enables us to draw parallels, and we have only to repeat the construction to obtain a second right angle with its arms in other directions, and this enables us to draw perpendiculars. Since we can draw parallels, we can make similar figures, and dividers have no wider range of construction than an instrument, such as a ruler or strip of paper with two fixed marks on it, that will transfer from one part of the plane to another one definite segment, say the unit of length. The range is also just as wide if all we can do is to rotate the unit of length about one end which remains fixed, that is, to cut off a unit length from any straight line through a fixed origin; an instrument which can do this is called by Hilbert *Einheitsdreher* (unit rotator). For with an Einheitsdreher we can construct a rhombus as above, and then with a ruler we can draw parallels. Then let it be required to cut off, from a given segment PQ, a part PX equal to a given segment

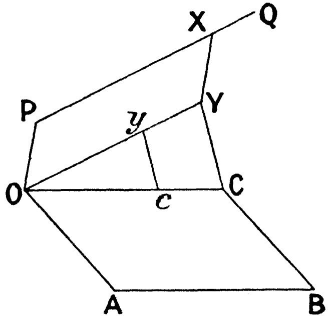

FIG. 31.

AB. Let O be the origin; draw OC, OY parallel to AB, PQ respectively, and cut off from these unit lengths Oc, Oy. Complete the parallelogram OABC; draw CY parallel to cy, and YX parallel to OP to meet PQ in X. Then PX is the required segment, for, by parallels,

$$PX = OY = OC = AB.$$

Hilbert in his *Grundlagen* gives a proof that the range of constructions possible with this instrument corresponds to the set of quadratic surds which are essentially real, so as to remain real however the quantities on which they depend vary within their domain ; the full proof is too hard to give here. For example, we can bisect any angle ; this is a quadratic problem which always has two real solutions, the internal and external bisectors, however the given angle may vary. Cut off equal lengths OA, OA′, OB from the vertex O along the arms of the angle ; complete the rhombuses OACB, OA′C′B ; then OC, OC′ are the

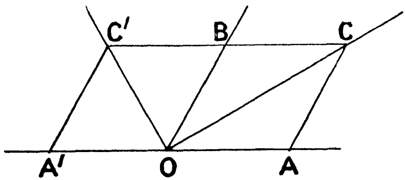

FIG. 32.

two bisectors. Again, if the centre and radius of a circle are given, we can draw straight lines through the centre, and mark off on each its two intersections with the circumference, at a distance from the given centre equal to the given radius ; we can obtain the intersections of the circle with any diameter, for these are always real ; but we cannot obtain its intersections with an arbitrary straight line which does not pass through the centre, for these intersections become imaginary for some positions of the given straight line.

We can see more exactly what can be done with an Einheitsdreher if we turn to analysis. Take the centre as origin ; we can construct the two intersections of the circle $x^2+y^2=1$ with any straight line through the centre $y=mx$ that we can draw, that is, such that m belongs to the domain. The coordinates of these points of intersection are

$$\left(\pm\frac{1}{\sqrt{1+m^2}},\quad \pm\frac{m}{\sqrt{1+m^2}}\right),$$

and since we can carry out rational operations on these coordinates, the effect is to add $\sqrt{1+m^2}$ to the domain, which is therefore formed from the coordinates of the given points by rational operations, together with this new operation of passing from m to $\sqrt{1+m^2}$, where m is any quantity belonging to the domain. More generally, if m, n are any two quantities of the domain, we can construct in turn $\frac{m}{n}$, $\sqrt{1+\left(\frac{m}{n}\right)^2}$, and $n\sqrt{1+\left(\frac{m}{n}\right)^2}$ or $\sqrt{m^2+n^2}$; so that the new operation is equivalent to taking the square root of the sum of squares. This operation is less general than that of taking the square root of an arbitrary quantity of the domain, for $\sqrt{m}$ is an imaginary quantity when m has some of its admissible values, namely real negative values; but $\sqrt{m^2+n^2}$ remains real for all admissible values of m and n.

(ii) **Parallel Ruler.**

Two straight edges rigidly connected, so as to make a constant angle with one another, form an instrument various types of which are common. If the angle is a right angle, we have a T-square; $\frac{1}{6}\pi$, $\frac{1}{4}\pi$, $\frac{1}{3}\pi$ are usual in set-squares; these will be discussed below. If the angle is zero, we have a parallel ruler.

First of all, this can be used as an ordinary ruler. Next, it gives the metrical data, for if we lay it down anywhere and rule along both edges, and then do it again with the edges in a different direction, we get a rhombus, of which each altitude is equal to the breadth of the ruler, which we shall take as the unit of length, and the diagonals of the rhombus give a right angle. The positions of the two edges are completely determined as soon as they are made to pass through two given points of the plane; we must distinguish between two uses of the parallel ruler, according as the same edge is made to pass through both of the given points, or a different edge through each.

First Use Equivalent to Dividers.

In the first use, one edge is made to coincide with a given straight line, joining the given points AB; and we rule a parallel straight line at unit distance from AB; in fact we can rule two, one on each side of it. Since there are these two positions, this elementary operation with the parallel ruler is equivalent to solving a quadratic equation, but as before, it is one whose roots are always real. The first use is exactly equivalent to the use of dividers; for we can cut off a unit distance from any straight line OA, by first drawing the perpendicular OB and then drawing a parallel to OB at unit distance from it, to meet OA in the point required; and conversely, dividers enable us to draw a parallel to a given straight line OB at unit distance from it, by first cutting off a unit length from the perpendicular OA. With dividers, we can find the intersections of the unit circle with any diameter, and then draw the tangents at these points, perpendicular to the radii; with a parallel ruler we can draw the tangents to the unit circle parallel to any given radius, and then determine their points of contact as the feet of the perpendiculars through the centre.

Second Use Equivalent to Compasses.

In the second use of the parallel ruler, one edge is made to pass through one given point A, and the other through B. This is only possible provided $AB \geqq 1$. Then by ruling along the second edge we obtain, by an elementary operation, the solution of the problem: through a given point B to draw a straight line BC at unit distance from another given point A, that is to say, to draw from B a tangent to the circle centre A of unit radius. This has two solutions in general, and is again equivalent to solving a quadratic equation; but it is one whose roots become imaginary when B lies within the unit circle, that is, when the parameter $m = AB$ takes real values less than unity, which are permissible values for the distance between two arbitrary points of the plane. Now if C is the point of contact of the tangent, which is the foot of the perpendicular from A to BC, we have $BC = \sqrt{m^2 - 1}$, so that we can construct surds of this form. But then we can construct any quadratic surd, for if a is any given quantity of the domain, and $m = \frac{a+1}{a-1}$, we can obtain in turn m, $\sqrt{m^2-1}$, and $\frac{a-1}{2}\sqrt{m^2-1}$ or $\sqrt{a}$. Thus a parallel ruler, in its second use, can wholly replace compasses.

For example, let it be required to find the intersections of a given straight line PQ with a circle whose centre O and radius are given, but which is not to be drawn. We assume that the radius is equal to the breadth of the ruler; if this were not so, we could first construct a figure similar to the required figure but with a unit circle, and then alter the scale to the desired size.

If X, Y are the required intersections, the tangents to the circle at X, Y meet in R the pole of PQ. These tangents can be drawn at once if R is first obtained. Now R lies on TT′, the polar of any point P of PQ, and also on OS, the

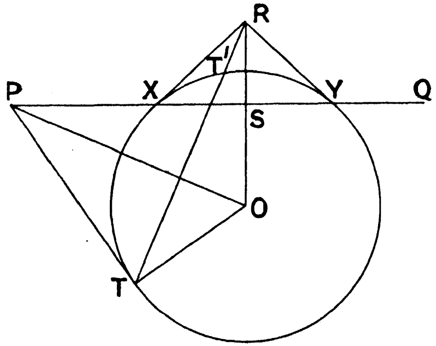

FIG. 33.

perpendicular from O to PQ. We therefore take any point P on PQ, and draw PT through P at unit distance from O; draw OT perpendicular to PT; then TR perpendicular to OP and OSR perpendicular to PQ to meet in R; draw RX, RY at unit distance from O to meet PQ in X, Y, which are the required points.

In this second use of the parallel ruler, we are only using a single point of the first edge, and the essential parts of the instrument are one straight edge and a point at a fixed distance from it, which can be made to coincide with any given point of the plane, while at the same time the straight edge is made to pass through a second fixed point which is far enough away from the first.

(iii) **Set-square.**

If the ruler has two straight edges which are not parallel, but inclined at a constant angle α, its position is not determined unless we make its edges pass through three fixed points, two on one edge and one on the other; the elementary operation is to draw through a given point A a straight line making the angle α with the straight line joining two other given points BC. We can construct a parallelogram and draw parallels; in particular, we can draw a rhombus with BC as diagonal and an angle 2α at B; the other diagonal is then at right angles to BC, and we have all the metrical data. But so far we have added nothing to the range of ruler constructions, except that $\tan\alpha$, if irrational, must be added to the set of independent coordinates, of which the rational functions can be constructed.

But if we assume the power to make a marked point of one edge coincide with a fixed point of the plane, while the second edge passes through another fixed point, then since the marked point is at a fixed distance from the second edge, this use of the set-square is exactly equivalent to the second use of a parallel ruler, and so can replace the use of compasses. A different use of the set-square, which can also replace the use of compasses, is when we assume the power to make each edge pass through one of two given points A, B, and also the vertex, or point of intersection of the edges, lie on a given straight line **c**. The two positions of the vertex are the two intersections of **c** with the arc of a circle standing on AB capable of containing an angle α; and since we can determine the intersections of any straight line with this fixed circle, we can (p. 69) carry out all constructions that are possible with ruler and compasses.

CHAPTER V.

STANDARD METHODS.

THE only quite general methods which apply to all ruler and compass problems are those which follow the steps of the analytical discussion in each case; these are seldom the best. But there are certain standard methods, each of which applies to a more or less well-defined group of cases, though the groups overlap considerably. The idea underlying many of these is the *separation of properties*. If the problem is definite, we have to construct certain elements which have certain relations to the data, and each of these required elements has to satisfy just as many independent conditions as are necessary to determine it: two for a point or a straight line, three for a circle. If these conditions have more than one solution, there may be further requirements, which exclude some, but not all, of the elements which satisfy the other conditions. Now it is often possible to separate the conditions into two distinct statements, each expressing a single requirement or group of requirements, and to consider each statement by itself. Then in general, each statement determines an infinite set of elements, and the solution of the problem is given by the common elements of the two sets.

(i) **Method of Loci.**

The simplest example is the method of loci, when a point is to have two properties, each of which is satisfied by all points of a certain locus; then the point or points of intersection of these two loci are the solutions of the problem. In applying this, the art is so to separate the properties that the two loci are both easy to discover and possible to draw: they must for our purpose consist of straight lines and circles. The method fails to give a ruler and compass construction if one of the loci is a curve other than a circle. Easy examples of this method are given by the usual constructions for the circumcentre and incentre of a triangle, and many similar problems.

Ex 1. The circumcentre S of a triangle ABC is defined as a point equidistant from the three corners A, B, C. This can be separated into the two statements, SA = SB and SA = SC. The locus of points having the first of these two properties is the perpendicular bisector of AB; the locus of points having the second is the perpendicular bisector of AC; the single point S common to these two loci is the solution of the problem.

Ex 2. To construct the radical axis of two given circles. This is defined as the locus of points from which equal tangents can be drawn to the two circles; it is known to be a straight line perpendicular to the straight line joining the centres A, B. It can therefore be drawn if a single point X on it is obtained, by drawing the perpendicular from X to AB; or we may obtain two such points X, Y and join them.

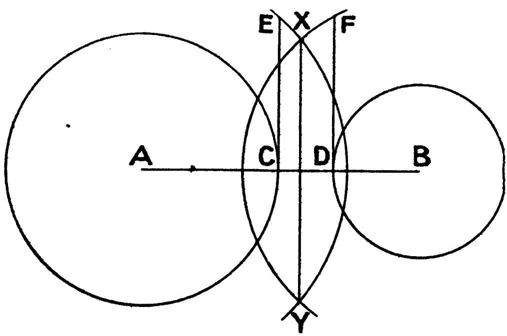

Fig. 34.

If X is the point from which the tangent to each circle is of a given length t, then X may be constructed by the method of loci. The point X has two properties: (i) the tangent to the first circle is of length t, (ii) the tangent to the second circle is of length t. The locus corresponding to the statement (i) is a circle concentric with the first given circle ; in order to describe it, we first find one point E on its circumference, by cutting off from any tangent a length CE = t measured from the point of contact C, and then describe the circle A(E). Similarly the locus corresponding to (ii) is another circle B(F). We can take either of the points of intersection of these two loci to be X and the other to be Y ; the radical axis of the two given circles is the straight line XY. The whole construction can be arranged as follows.

Let the line of centres AB meet the circumferences in C, D. Draw CE, DF at right angles to AB and cut off CE, DF, each equal to any convenient length t. Describe the circles A(E), B(F) to meet in X, Y ; join XY, which is the radical axis required. We must take t large enough for the circles A(E), B(F) to meet in real points. This is always possible except when the given circles are concentric, and then the radical axis is wholly at infinity. For a shorter construction of the radical axis see pp. 108, 117.

(ii) **Method of Trial and Error.**

This is somewhat similar to the method of loci ; we replace the point to be constructed by a pair of related points, which between them satisfy all the requirements of the problem, so that it would be solved if the pair coincided. That is to say, we regard the required point as a particular case of the pair of points, which has the separate additional property that the two points of the pair coincide. In most of the cases in which this method leads to a ruler and compass construction, the locus of the first point of each pair is a straight line or circle, and the second point has the same locus ; and the corresponding positions form two homographic ranges. Then the solution of the problem is given by the common points of the homography, which can be constructed by the method of the last chapter (p. 65). One of the differences between this method and the method of loci is that there we have to find a point of the first locus which coincides with some point, no matter which, of the second locus ; while here we have to find a point of the first range which coincides with a definite corresponding point of the second range, namely the point corresponding to the first by means of the construction adopted.

Ex. 1. We have already had an example of this method on p. 64, in determining the intersections of an arbitrary straight line with the conic through five given points.

Ex. 2. To draw a triangle XYZ with its sides parallel to fixed directions and its corners lying upon the sides of a given triangle ABC.

Start from any point X of BC; draw in succession XY, YZ, ZX′, parallel to the proper fixed directions, to meet CA in Y, AB in Z, and BC in X′ respectively. Then the problem is solved if X′ coincides with X. This will not

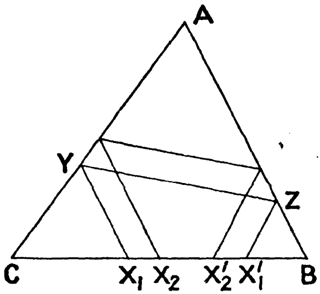

FIG. 35.

happen in general, but if X moves on BC, and describes a certain range, X′ describes a homographic range also on BC; for the cross-ratio of any four positions of X is equal to those of the corresponding positions of Y, of Z and of X′. The common points of the homography X, X′ give the solution of the problem.

If the lines are arranged as in fig. 35, when X moves from X_1 to X_2 towards X′, both Y and Z move towards A, and X′ moves toward X; thus X, X′ come together at some point between B and C, so that there is at least one real finite common point of the homography. It follows that the other common point is also real, but it lies at infinity; in general, the problem has one and only one solution. With a different figure, there may not be a common point between B and C, but in the general case, as X goes to infinity along BC produced, Y, Z and X′ all go to infinity also, and X, X′ both ultimately coincide with the point at infinity on BC, which is the second common point of the homography. Certain special cases are an exception to this. If the first fixed direction is parallel to BC, then Y always coincides with C, and Z, X′ are fixed points; similarly X′ is fixed if YZ is parallel to CA, or ZX′ to AB. These cases we exclude, and also the similar cases in which XY, YZ or ZX′ is parallel to CA, AB or BC respectively, when X′ is at infinity always.

These facts about the common points may also be proved analytically. Let $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ be the sides of the triangle ABC, and let $CX = \mathbf{x}$, $CX' = \mathbf{x}'$. Since the triangles XCY, YAZ, ZBX′ have their sides in fixed directions, they are of constant shapes, and the sides of any one of them are in constant ratios. We may therefore introduce three constants α, β, γ, such that

$$CY = \alpha CX = \alpha \mathbf{x},$$
$$AZ = \beta AY = \beta(\mathbf{b} - \alpha \mathbf{x}),$$
$$BX' = \gamma BZ = \gamma(\mathbf{c} - AZ) = \gamma\{\mathbf{c} - \beta(\mathbf{b} - \alpha\mathbf{x})\},$$
$$\mathbf{x}' = CX' = \mathbf{a} - BX' = \mathbf{a} - \gamma\{\mathbf{c} - \beta(\mathbf{b} - \alpha\mathbf{x})\};$$
$$\therefore\ \mathbf{x}' + \alpha\beta\gamma\mathbf{x} = \mathbf{a} - \gamma\mathbf{c} + \beta\gamma\mathbf{b} = \text{constant}.$$

The excluded cases correspond to the values $\mathbf{0}$ and ∞ of α, β or γ. To find the common points, put $\mathbf{x} = \mathbf{x}'$; this gives one finite solution

$$\mathbf{x} = \frac{\mathbf{a} - \gamma\mathbf{c} + \beta\gamma\mathbf{b}}{1 + \alpha\beta\gamma};$$

if we regard this as a degenerate quadratic equation, the other root is infinite.

In the particular case $\alpha\beta\gamma = 1$, the equation of the homography is symmetrical in $\mathbf{x}$, $\mathbf{x}'$, and we have an involution; starting from X′ and repeating the construction, we get back to X. This condition is satisfied, for example, if XY, YZ, ZX′ are parallel to the sides of the pedal triangle PQR formed by joining the feet of the perpendiculars from A, B, C on to the opposite sides of the triangle. Then the pedal triangle itself is the solution of the problem.

In this case each of the triangles XYC, AYZ, X′BZ is similar to ABC, and we have

$$\alpha=\frac{\mathbf{a}}{\mathbf{b}},\quad \beta=\frac{\mathbf{b}}{\mathbf{c}},\quad \gamma=\frac{\mathbf{c}}{\mathbf{a}},\quad \alpha\beta\gamma=\mathbf{1}.$$

Also XY, YZ are equally inclined to CA, and similarly at Z, X′, so that XYZX′Y′Z′X is the path of a ray of light

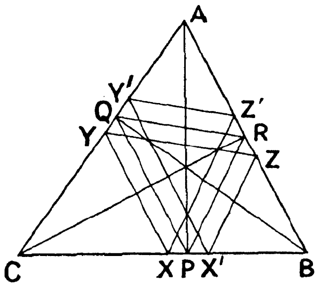

FIG. 36.

incident along Z′X and reflected at the sides in turn ; if such a ray after three reflexions is parallel to its original direction, then after six reflexions it repeats its path.

Another exceptional case is when $\alpha\beta\gamma=-\mathbf{1}$; then both common points are at infinity, and

$$XX' = \mathbf{x}' - \mathbf{x} = \mathbf{a} - \gamma\mathbf{c} + \beta\gamma\mathbf{b} = \text{constant},$$

so that X, X′ can never coincide unless this constant vanishes, and then they always coincide. When this special relation $\alpha\beta\gamma=-\mathbf{1}$ holds, the problem is poristic (p. 67), and there is no solution unless the further relation $\mathbf{a}-\gamma\mathbf{c}+\beta\gamma\mathbf{b}=\mathbf{0}$ is satisfied, and then there are an infinite number of solutions. But if both these conditions hold, the three given directions must be all the same.

(iii) Method of Similar Figures.

This same problem can also be treated by the method of similar figures, in which the properties separated from the rest are the size and position of the required figure. We construct a figure similar to the required one in some other part of the plane, and then copy it on the right scale and in the right position. The similar figure is some-times constructed by reversing the data and desiderata of the problem. This *method of the inverse problem* can often be used together with other methods.

Ex. 1. In the example we have just been discussing, let any three straight lines in the given directions form a triangle LMN. Through the corners of this draw parallels to the sides of the given triangle ABC, to form a third

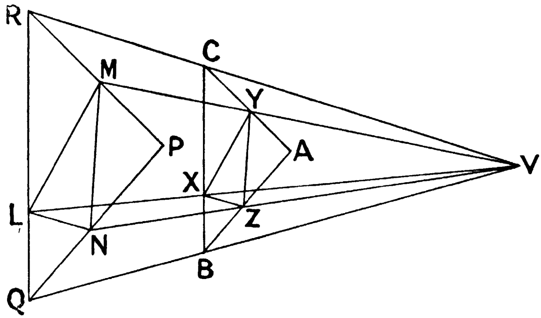

FIG. 37.

triangle PQR. Then PQRLMN is a figure similar to the required figure ABCXYZ, and X divides BC in the same ratio as L divides QR ; if QB, RC meet in V, then VL meets BC in the required point X. The two figures are not only similar, but similarly situated, with V as centre of similarity; and Y, Z can be determined either by drawing parallels to LM, LN through X, or by joining VM, VN to meet CA, AB respectively.

Ex. 2. To draw a triangle OXY similar to a given triangle LMN, having one vertex at a given point O and the others on two given straight lines AB, BC.

We can complete the construction by drawing a similar figure, if we can find the point P which has the same relation to the triangle LMN that B has to OXY ; for then PM, PN correspond to BA, BC. Now P lies upon the arc of a circle standing on MN capable of containing an angle equal to ∠ABC ; and we can also determine the straight line LP by means of its second intersection Q with the same circle. For Q corresponds to D the point of intersection of OB with the circle XBY and we have

$$\angle MNQ = \angle XYD = \angle XBD \text{ or } ABO,$$

which is a given angle. Hence we can draw NQ and determine Q, draw LQ and determine P, and then construct OXBY similar to LMPN.

Now let AB, BC be adjacent sides of a parallelogram ABCE, and let O be its centre; and let MN be a side and

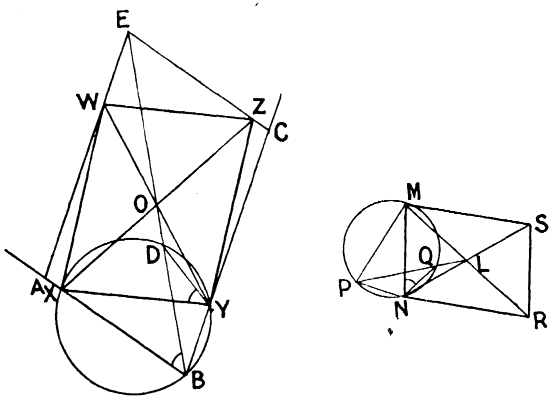

FIG. 38.

L the centre of another given parallelogram MNRS. Then, if XO, YO are produced to meet CE, EA in Z, W respectively, XYZW is a parallelogram similar to MNRS, and we have a solution of the problem: In a given parallelogram to inscribe a parallelogram similar to another given parallelogram.

Ex. 3. **Eccentric Circle.**—To find the intersections of a given straight line with a conic whose focus, directrix and eccentricity are given. In this case the method of loci would apply at once if we had an instrument for drawing the conic, for the two properties of lying on the straight line and lying on the conic have for their obvious loci the straight line and the conic respectively. With ruler and compasses we cannot draw the second locus, but we can fall back on the method of similar figures.

Let the given straight line meet the given directrix QN in Q and the conic in X, X′; let S be the focus and e the eccentricity. Join SQ. Take any point **p** on QX, draw **pn** perpendicular to QN, and describe the circle centre **p**, radius **e** . **pn** ; this is the *eccentric circle* of **p** ; let it meet QS in **s**, **s**′. Draw SX, SX′ parallel to **sp**, **s′p** to meet

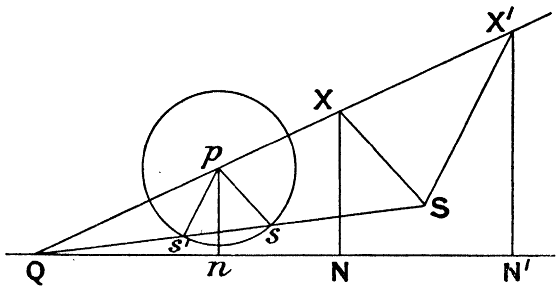

FIG. 39.

QX in XX′ ; then these are the required points. For if XN is drawn perpendicular to QN, since the figures SXQN, **spQn** are similar, we have

$$\frac{\mathrm{SX}}{\mathrm{XN}} = \frac{\mathbf{sp}}{\mathbf{pn}} = \mathbf{e},$$

so that X lies on the conic, and similarly X′. The solutions are real, provided SQ meets the eccentric circle of **p** in real points. If SQ touches this circle, the given straight line touches the conic, and the construction gives the point of contact.

(iv) Methods of Perspective, Translation, Rotation and Reflexion.

These are all particular cases of the method of similar figures, in which abstraction is made of some but not all of the properties which determine the size and position of the required figure. Examples 1 and 3 of pp. 84, 85 have already illustrated the method of perspective. It is often useful to apply these methods to part only of the figure, and so transform it into something simpler.

Ex. 1. To find the locus of a point such that the difference between its perpendicular distances **x**, **y** from two given straight lines **a**, **b** is a constant length **d**.

Translate **b** parallel to itself through a distance equal to **d**; it will take up one of two positions **b**′, **b**″ according to the sense of the translation. Then the required points are equidistant either from **a**, **b**′ or from **a**, **b**″, and there-

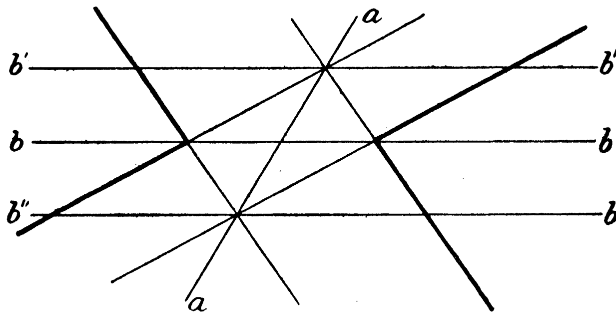

FIG. 40.

fore lie on one or other of the four bisectors of the angles between these two pairs of straight lines. If **x**, **y** are always taken to be positive, the figure shows the parts of these bisectors on which $x - y = d$; on the other parts, either $x + y = d$ or $y - x = d$.

Ex. 2. To describe a circle of given radius **a** to cut off intercepts PQ, RS of given lengths from two given straight lines AB, BC.

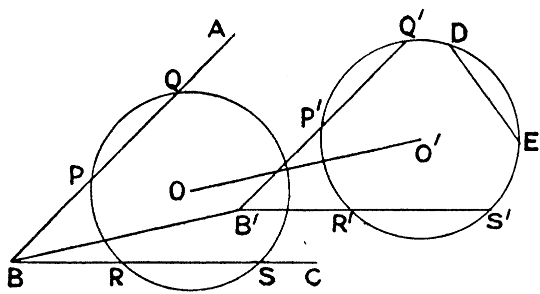

FIG. 41.

Take any circle O′(**a**) of the same radius; place it in a chord DE of length PQ; rotate DE about O through an angle equal to that between DE and AB. We thus have a chord P′Q′ equal and parallel to the required intercept PQ. Similarly place in the same circle a chord R′S′ equal and parallel to the required intercept RS on BC, and let P′Q′, R′S′ meet in B′. Join B′B, and draw O′O equal and parallel to B′B; then O is the centre of the required circle. For if the whole figure O′P′Q′R′S′B′ is translated without rotation through the step B′B, then B′ comes to B, and the straight line B′P′Q′ is brought to coincide with its parallel BA and B′R′S′ with BC; and the chords of required length are brought to lie along the given straight lines, while the radius of the circle is unaltered.

Ex. 3. To construct a triangle, given the length of the base, the sum of the sides and the vertical angle.

If ABC is the required triangle, reflect B in the external bisector of the vertical angle C. Then if D is this reflexion, ACD is a straight line, BCD is an isosceles triangle, and

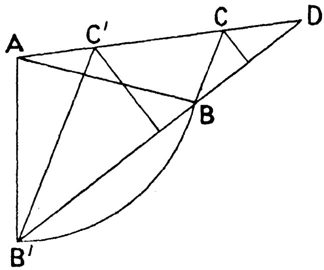

FIG. 42.

AD = AC + CB = the given sum, while ∠ ADB is half the given vertical angle ACB. We have therefore the following construction.

Take a straight line AD of length equal to the given sum, and make an angle ADB equal to half the given angle. With centre A and radius equal to the given base describe a circle to cut DB in B, B′. If these points are real, bisect DB, DB′ at right angles to meet AD in C, C′. Then either ABC or AB′C′ is the required triangle.

The method of reflexion naturally finds its classic examples in the theory of geometrical optics.

Ex. 4. **Optical Image.**—To find the path of a ray of light from one given point to another by one reflexion at a given plane surface, and to show that it is the shortest of all such broken lines joining the two points.

We have to find the point X at which the ray AXB between the given points A, B is reflected at the given plane XC. The laws of optical reflexion are that AX, XB lie in a plane through the normal at X to the reflecting

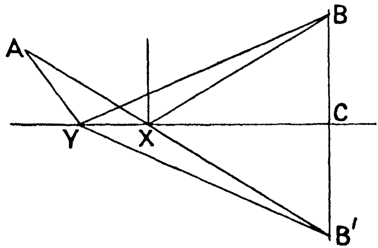

FIG. 43.

surface, and make equal angles with the normal on opposite sides. Hence if B′ is the geometrical reflexion of B in the plane XC, found by drawing BC perpendicular to the plane and producing it an equal length, then AXBB′ lie in a plane and AXB′ is a straight line. X can therefore be found as the intersection of AB′ with the given plane.

Also, since B, B′ are equidistant from any point Y of the plane, the total length of any other broken line AYB from A to B is the same as AY + YB′, which is greater than AB′ the third side of the triangle AYB′, and therefore greater than the length of the actual path of the ray.

The following problem involves much the same idea.

Ex. 5. A spider in one corner of a room wants to reach a fly on the opposite corner of the ceiling by crawling across the floor and up one wall; what is his shortest path?

Let ABCD be the floor, A the spider, E the fly, E being vertically above C. There are two types of path, according as the spider leaves the floor at a point X on the longer side, BC say, or at a point Y on the shorter side CD. While on the floor his path must be a straight line, and also while on the wall. We have to determine, first for what point

X of BC the path AXE is a minimum, then for what point Y of CD the path AYE is a minimum, and lastly, which of these two minima is the lesser. Now let the wall BCE be rotated about BC till it comes into the same plane as the floor. Then E comes to E_1 in DC produced, where

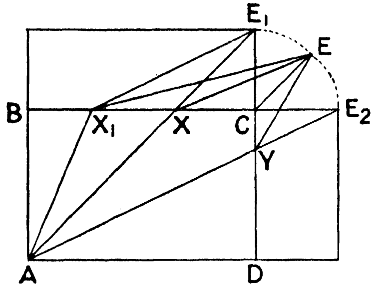

FIG. 44.

$CE_1 = CE$, and XE comes to XE_1. Then $AX + XE_1 \geqq AE_1$, and the minimum length for paths of this type is AE_1, when AXE_1 is a straight line; the point X is constructed as the intersection of AE_1 and BC, and X divides BC in the ratio AB : CE. Similarly, for paths that leave the floor at a point Y on BC, the minimum is of length AE_2, where E_2 is the point in BC produced, such that $CE_2 = CE$. The length of the required shortest path is therefore the lesser of AE_1, AE_2. Now

$$AE_1^2 = AD^2 + (AB + CE)^2 = AD^2 + AB^2 + CE^2 + 2AB \cdot CE,$$

$$AE_2^2 = (AD + CE)^2 + AB^2 = AD^2 + AB^2 + CE^2 + 2AD \cdot CE,$$

and the first is the lesser, provided $AB < AD$. The shortest path therefore leaves the floor at a point in the longer side, which divides it in the ratio of the breadth to the height of the room.

(v) Method of Projection.

The method of similar figures and its varieties are all particular cases of the method of projection, in which we separate the projective from the metrical properties of the required figure. By a suitable projection we transform the problem into one in which the same set of projective properties are combined with a different metrical set, which makes the case easier to solve. Then we reverse the projection, and pass either from the final solution of the transformed problem to the solution of the original problem, or else from each step in the construction of the second figure to the corresponding step in the construction of the required figure.

From one point of view the method of projection is an example of the separation of properties; but from another it illustrates the more important *principle of transformation*. That is to say, we do not directly solve the proposed problem, but another in which each element, whether given or required, stands in some definite relation to the corresponding element in the given problem and its solution. If the modified problem cannot be solved, or is no easier than the original problem, then the transformation adopted is not suitable. When we have solved the transformed problem, by a reversal of the transformation we pass back to the original figure, and so obtain a construction for the problem in the form in which it was given. Many examples of this have already been met with in chapter III.

Ex. 1. **Auxiliary Circle.**—To draw the tangent at a given point of an ellipse whose principal axes are given.

Project the ellipse into a circle. We may take the vertex of projection at infinity in the direction of the normal to the plane of the ellipse, and project upon a plane through the major axis inclined at an angle $\cos^{-1}\frac{CB}{CA}$ to the plane of the ellipse, where CA, CB are the given semi-major and -minor axes. The given point of the ellipse

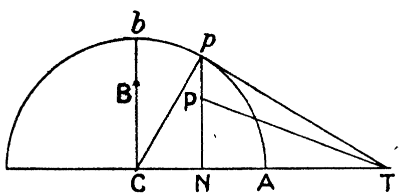

FIG. 45.

projects into a point of the circle, and the required tangent into the tangent to the circle at this point. If the plane of projection is rotated about the major axis till it coincides with the plane of the ellipse, then the circle comes to coincide with the major auxiliary circle ; corresponding points of the two figures lie on straight lines perpendicular to CA, at distances from it which are in the ratio of the axes. Points on CA correspond to themselves, and the corresponding tangents meet on CA, so that we have the following ruler construction for the problem :

Let CA, CB be the given semi-axes and P the given point. Draw PN perpendicular to CA, and produce NP to p, making $\frac{pN}{PN}=\frac{CA}{CB}$. Join Cp, and draw pT perpendicular to Cp to meet CA in T. Join PT, which is the required tangent to the ellipse at P.

(vi) Method of Inversion.

This is another important example of the method of transformation, which under certain circumstances changes circles into straight lines, and so may transform a problem into a much simpler one. Inversion replaces every point by its inverse with regard to a fixed circle, where two points P, P′ are defined to be the *inverses* of one another with regard to the circle O(k) if OPP′ is a straight line and $OP \cdot OP' = k^2$. Here O and k are called the *centre* and *radius of inversion* ; when the radius is arbitrary, we may speak of inversion with regard to the point O. If the radius k is pure imaginary, then $OP \cdot OP'$ is negative, and P, P′ are on opposite sides of the centre O.

We assume without proof the following properties of inversion, which are used in this and later chapters.

(1.) Inversion is a symmetrical transformation : if P′ is the inverse of P, then P is the inverse of P′, so that to reverse the transformation is the same as to repeat it. To every point P there corresponds one definite inverse point P′, except that to the centre O there corresponds the whole straight line at infinity.

(2.) If P is outside the circle, P′ is inside ; inversion interchanges the two regions into which the circle divides the plane. Every point on the circumference of the circle, which is the common boundary of the two regions, inverts into itself.

(3.) Any circle inverts into a circle, except that a circle through O inverts into a straight line; the centre of the original circle inverts into the point inverse to O with regard to the inverse circle.

(4.) Any straight line inverts into a circle through O, except that a straight line through O inverts into itself.

(5.) A circle orthogonal to the circle of inversion inverts into itself.

(6.) Two curves cut at the same angle as their inverses; and, in particular, contact is unaltered by inversion.

Ex. 1. To find the locus of the centre of a circle orthogonal to two given circles.

Let the given circles OAP, OBP cut in O, P. Invert with regard to O as centre and OP as radius; then P inverts into itself, and the given circles become straight lines PA′, PB′ intersecting at P. A circle centre X, orthogonal

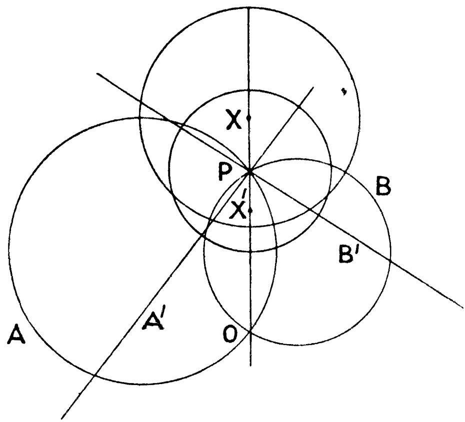

Fig. 46.

to both the given circles, inverts into a circle to which PA′, PB′ are both normals, that is, radii, so that the centre is P. Now X′, the inverse of X, is the point inverse to O with regard to this last circle, whose centre is P; hence X′ lies on OP; and since X, X′ are inverse with regard to O as centre, therefore X lies upon the same straight line OPX′. The required locus is therefore OP, the radical axis of the two given circles. It is easy to find other proofs of this. The proof by inversion is valid whether the point O and the inverse figure are real or imaginary.

Since inversion with regard to a common orthogonal circle changes each of the given circles into itself, the locus found is that of the centre of an inversion which leaves two given circles unaltered.

We can invert three circles into themselves by choosing their radical centre as centre of inversion ; but we cannot invert four or more circles into themselves unless their radical axes all meet in a point, so that they have a common orthogonal circle.

Ex. 2. To describe a circle to touch two given circles and to pass through a given point on their radical axis.

Invert with regard to the given point C, and choose the radius of inversion equal to the tangent from C to either of the given circles α, β ; it follows from (5) that

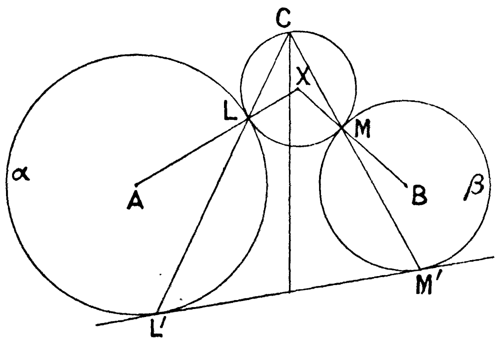

FIG. 47.

they invert into themselves. The required circle becomes a common tangent to α, β. We must therefore construct the common tangents, and then invert them with regard to C. There are in general four real common tangents and four solutions of the problem, if α, β are external to each other. But there is an exceptional case when the given point C lies upon one of the common tangents ; for then, by (4), this straight line inverts into itself, that is, into a straight line and not into a circle, so that there are only three solutions to the problem as stated. If α, β touch externally, there are three real common tangents; but one, the tangent at the point of contact, is the radical axis and passes through C, so that it does not invert into a circle, and there are only two solutions. There are in general two solutions also if α, β cut in real points, but only one if C lies on one of the two real common tangents; and if one of α, β touches the other internally, or lies wholly within it, there is no solution of the problem. The actual construction is as follows.

Draw a common tangent, by Euclid's or any other method, to touch the given circles α, β in L′, M′. If A, B are the centres and C the given point on the radical axis, join CL′, CM′ to meet α in L and β in M respectively. Join AL, BM to meet in X; then X is the centre of the required circle, XC is its radius, and it touches α at L and β at M. The proof is left to the reader.

If we regard the point C as a circle of zero radius, we have a particular case of the famous *Problem of Apollonius*:

Ex. 3. To describe a circle to touch three given circles.

If the three circles are all external to one another, there may be as many as eight solutions, for each of the given circles may be either inside or outside the touching circle. If the second circle is inside the first, and the third inside the second, as, for example, when they are concentric, there is no real solution; in other cases there may be any intermediate number; one real solution is lost when one of the touching circles becomes a straight line or reduces to a point, and two are lost when a pair of touching circles first coincide and then become imaginary.

If two of the given circles cut in real points, we may invert the whole figure with regard to one of these points; then these two circles become straight lines, and the problem becomes the simpler one: to describe a circle to touch two given straight lines and also a given circle. We shall first solve this, and then show how to pass to the less simple case when none of the given circles cut in real points.

Let LM, LN be the two given straight lines and P(p) the given circle. If Y(y) is the required circle, it touches P(p) either externally or internally. Assume first that contact is external; then suppose y increased by p: we have a circle centre Y, which passes through P and touches two straight lines L′M′, L′N′ parallel to LM, LN, each at a distance p from its parallel on the side remote from Y. Hence Y is a solution of the problem: to find a point equidistant from two given straight lines L′M′, L′N′ and from a given point P. The first requirement tells us that Y lies on one or other of the bisectors of ∠M′L′N′, and

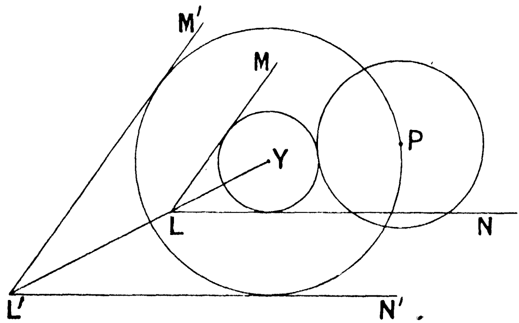

FIG. 48.

the second requirement tells us that it lies on the two parabolas with focus P and directrices L′M′, L′N′ respectively. Since Y is not only equidistant from L′M′, L′N′, but also from LM, LN, it lies on the bisector LL′ of ∠M′L′N′, and not on the other bisector. Corresponding to this assumption there may be two real positions of Y, the intersections of LL′ with a parabola whose focus and directrix are given. The determination of these is a particular case of p. 85, ex. 3. If we make a different assumption as to the position of Y, or assume that Y(y) touches P(p) internally, the parallels may have to be drawn on the other sides of LM or LN or both; so we have four cases to consider, in each of which we have to look for the intersection of one of the bisectors of ∠MLN with a parabola focus P. This leads to eight real positions of Y in the case in which P(p) cuts each of LM, LN in two real points on opposite sides of L. We can thus complete the construction of all the positions of Y that are real. The inverses of these, in the inversion used at the beginning to find LM, LN, are the positions X of the centres of the circles required by the problem of Apollonius.

Now return to the general case, in which none of the given circles are assumed to meet in real points. Let $A(\mathbf{a})$, $B(\mathbf{b})$, $C(\mathbf{c})$ be the given circles, and $X(\mathbf{x})$ the required circle. Now two circles touch if the distance between the centres is equal to either the sum of the radii, or to the difference taken positively. The conditions of tangency are therefore

$$XA = \pm\mathbf{a} \pm \mathbf{x}, \quad XB = \pm\mathbf{b} \pm \mathbf{x}, \quad XC = \pm\mathbf{c} \pm \mathbf{x},$$

with any arrangement of signs that makes the right-hand side positive in each equation. These equations still hold, if we add any quantity $\mathbf{d}$ to $\mathbf{x}$, and subtract the same quantity from $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ if the sign is the same as that of $\mathbf{x}$ in the corresponding equation, or add it if the sign is the opposite. We are now solving a different problem, the three given circles having the same centres as before, but modified radii, and the solution is a circle whose centre X is the same as in the solution of the original problem. This process may require us to consider a circle of negative radius. This must be supposed to coincide with the circle of corresponding positive radius, but where one would touch any circle internally, the other must touch the corresponding circle externally, and vice versa. In any case in which there is a real position of X, it is possible to choose $\mathbf{d}$ so that two of the modified circles meet in real points, so that X can be found by the above method. Then the corresponding radius of the required circle is $\pm XA \pm \mathbf{a}$. For if a circle exists touching the three given circles, of the three contacts there must be two which are either both external or both internal, and in the two corresponding equations the two signs are either both like or both unlike. By adding a positive or negative quantity to $\mathbf{x}$, we can therefore increase or decrease both of these two radii by any quantity, the same for both circles. Now if these two circles are external to one another, then by increasing the radii by a suitable quantity, they can be modified so as to meet in real points; while if one is internal to the other, the same can be done by decreasing the radii by a suitable quantity, the smaller radius becoming negative, and the inner circle passing outside the touching circle. If the two circles are neither external nor internal one to the other, then they meet in real points without modification.

The figure shows the case

$$XA = x + a, \quad XB = x + b, \quad XC = x - c.$$

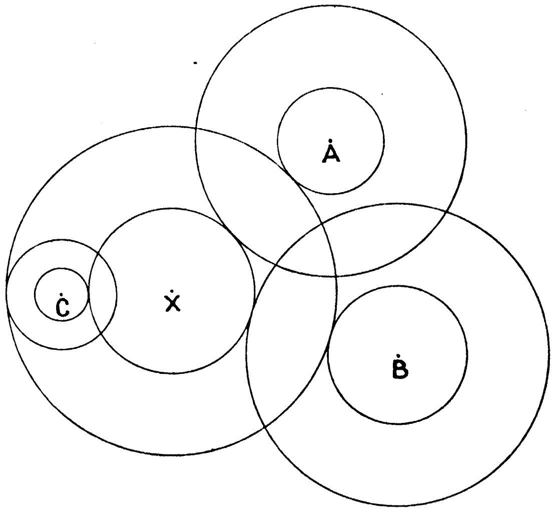

FIG. 49.

The radii of the modified circles are

$$a + d, \quad b + d, \quad c - d, \quad x - d,$$

of which the third is negative, and the corresponding circle has passed from inside to outside the touching circle.

(vii) Method of Reciprocation.

The method of reciprocation is a transformation which makes points correspond to straight lines and straight lines to points. In its simplest form, a point is made to correspond to its polar, and a straight line to its pole, with regard to a fixed circle of reciprocation. When the radius is arbitrary we may speak of reciprocation with regard to the centre of this circle. To points on a fixed straight line there correspond straight lines through a fixed point, and it may be shown that cross-ratio is unaltered by reciprocation. To any curve, regarded as the locus of a point, there corresponds the envelope of a straight line, that is, another curve. To the points on the circle of reciprocation there correspond the tangents at these points, which envelope the same circle ; this curve therefore reciprocates into itself.

The connection between reciprocation and inversion is given by the theorem : *the reciprocal of any curve is the inverse of the pedal of the same curve with regard to the centre of reciprocation.*

If SY is the perpendicular from S the centre of reciprocation upon any straight line p, then the pole P of p with

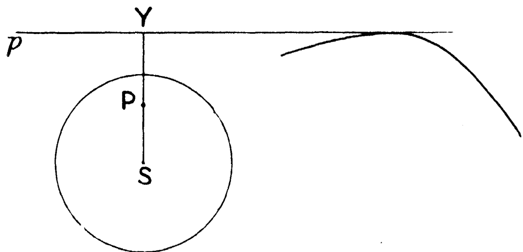

FIG. 50.

regard to the circle of reciprocation is the same as the inverse of Y. If p moves so as to envelope a curve, the locus of Y is called the *pedal* of the curve with regard to S, and the locus of P is the reciprocal of the same curve ; hence the reciprocal is the inverse of the pedal.

Ex. 1. The reciprocal of a conic with regard to the focus is a circle.

The pedal of a conic with regard to the focus S is the auxiliary circle ; hence, if we reciprocate with regard to S, the conic becomes the inverse of the auxiliary circle, which is another circle. The centre of the conic is also the centre of the auxiliary circle, which inverts into the point T inverse to S with regard to the circle reciprocal to the conic, p. 93 (3). Hence conics with a common focus can be simultaneously reciprocated into circles ; and confocal conics, which are also concentric, reciprocate into a family of circles for each of which S, T are inverse points, that is, a coaxal family with S, T as limiting points.

If the conic is a parabola, the pedal is the tangent at the vertex A, and the reciprocal is a circle through S.

This gives another construction for the intersections X, Y of a given straight line p with a parabola whose focus S and vertex A are given (p. 85). Find the pole P of p

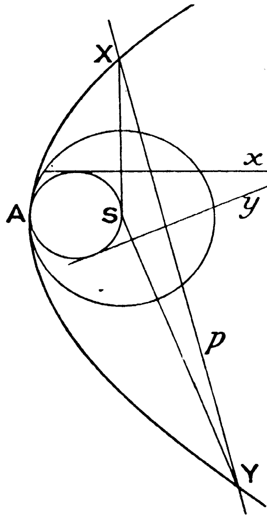

FIG. 51.

with regard to the circle S(A), and draw the tangents x, y from P to the circle on SA as diameter. Then the poles X, Y of x, y with regard to S(A) are the required points of intersection of p with the parabola.

CHAPTER VI.

COMPARISON OF METHODS.

HITHERTO we have spoken of the possibility of various constructions, and in the examples we have merely given one or more methods which lead to what is wanted, without making any comparison between them. Now whether one method is to be preferred to another depends upon many things; such as the context in which the problem arises, the extent and accuracy of our apparatus, and our skill in using it; or again, it may depend upon the ideas involved, rather than upon the operations; upon whether the construction can be easily put into words; upon the beauty and symmetry of the solution; or we may prefer a construction which combines several of these advantages in a moderate degree.

Thus it is very simple, from the point of view of language, to say in the course of a construction, "draw a figure similar to a certain figure," or "invert the whole figure with regard to a certain circle," but these may involve drawing a few dozen straight lines and circles. Euclid's proof of II. 9 and 10 are simple and attractive; but if we are dealing with the whole series II. 1-11, we shall probably treat them all by the same method of dissecting squares and rectangles. Simplicity is much the same as economy, and we have to choose what we shall economize: space in the diagram or words in the description, thought or labour or chance of error, or the indefinable quality of tiresomeness. The rest of this book is devoted to constructions which, as far as possible, avoid first, too large figures; next, certain ill-defined intersections; thirdly, more operations than need be; and lastly, in chapters VII and VIII, the use respectively of compasses and ruler.

I. Constructions in Limited Space.

It may happen that points which we might wish to use fall outside the paper, or too far off for the ruler to reach; or in surveying operations on a large scale, important points may fall in lakes or precipices or other inaccessible places. One "simple" way out of the difficulty is given by the method of inversion. Invert with regard to any circle which satisfies the two conditions: (i) every point inside it is accessible; (ii) every point or line given or required is outside it; then, by p. 92 (2), every point of the inverse figure is inside, and therefore accessible. If we have a construction for the problem which only fails because it requires inaccessible points, we can thus invert and construct the whole of the inverse figure, and finally, by inverting again, construct as much of what is required as comes on to the paper. In this way we have a proof that it is always possible to do what we want by means of lines and points which are not too far off, and so we are encouraged to look for more practical ways of doing it. Very often we can use the method of similar figures; that is, we copy the figure on a small scale, so that much more of it becomes accessible; and after completing the construction on this model, we enlarge as much of it as possible back to the former scale. The art lies in choosing the centre of similitude so as to make the change of scale as easy as possible.

Ex. 1. To join a given point A to an inaccessible point L, the intersection of two given straight lines **b**, **c**.

This problem is solved as soon as we have an accessible point X lying on AL.

(i) One method is to take any straight line AD through A and determine a point D upon it such that A, D are harmonically separated by **b**, **c**. Take any other straight line DX through D, and determine a point X on it such that D, X are again harmonically separated by **b**, **c**. These two harmonic ranges are in perspective with L as vertex, and therefore AX passes through L.

Start by taking any two straight lines B_1AC_2, B_2AC_1 through A to meet **b**, **c** with intersections as in fig. 52. Join B_1C_1, B_2C_2 to meet in D. Then AD, which need not be drawn, is a diagonal of the quadrangle $B_1B_2C_1C_2$, and it is divided harmonically by **b**, **c**, which are a pair of opposite sides of the same quadrangle. Draw any other

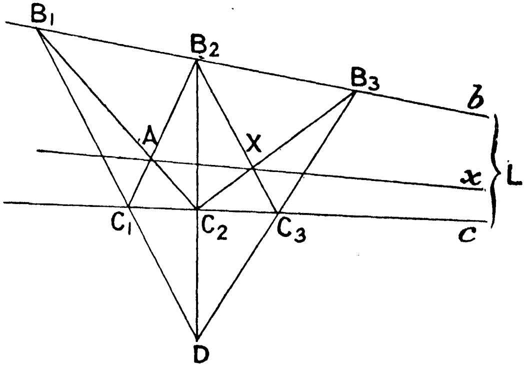

FIG. 52.

straight line DB_3C_3 through D to meet **b**, **c**, and join B_2C_3, B_3C_2 to meet in X; then DX is divided harmonically by **b**, **c**, for the same reason as before. Join AX, which is the straight line required.

This construction involves projective properties only. It may fail if A is nearly midway between **b** and **c**, for then the point D may be inaccessible for all accessible positions of B_1, B_2.

(ii) In that case we may adopt a method which depends

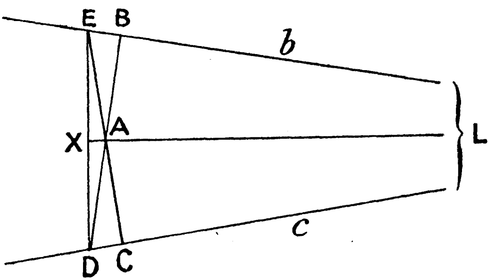

FIG. 53.

on the property of the orthocentre. Draw DAB perpendicular to **b** to meet **b** in B and **c** in D; draw EAC perpendicular to **c** to meet **c** in C and **b** in E; join DE, and draw AX perpendicular to DE. Then A is the orthocentre of the triangle LDE, and AX passes through L.

(iii) If **b**, **c** are very nearly parallel, the straight lines BD, CE, DE may be inconveniently close together; then we may use a construction depending on Desargues' theorem (p. 37).

Ex. 2. To draw the straight line joining two inaccessible points L, M, the intersections of two given pairs of straight lines **a**, **b**; **c**, **d** respectively.

(i) The required straight line is a diagonal of the quadrilateral formed by the four given straight lines. It passes through the points X, Y on the other two diagonals, which form harmonic ranges with the remaining diagonal point E and the pairs of opposite corners AB, CD. It is not

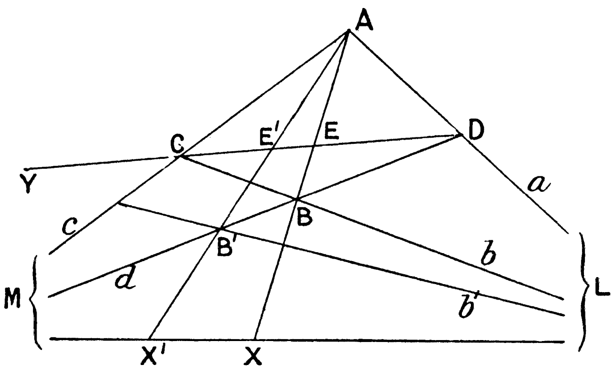

FIG. 54.

likely that both X and Y are accessible, as only one of them, X say, falls between L and M. If X only is accessible, we can obtain another point X′ on LM, moderately near to X, and therefore accessible also, by replacing **b** by another straight line **b**′, drawn through the inaccessible point L by one of the methods of ex. 1, from some point B′ of **d** moderately near to B. Then using the quadrilateral formed by **a**, **b**′, **c**, **d**, we are led as before to a second point X′ lying on the required straight line, which is therefore obtained by joining XX′.

(ii) If C or D is inaccessible, so that the point E cannot be readily obtained, we may fall back on the method of similar figures, using A as centre of similitude, for then the three straight lines **a**, **c** and AX are not altered. Take AB_1 any convenient fraction $1/\lambda$ of AB; draw B_1L_1, B_1M_1 parallel to **b**, **d** to meet **a**, **c** in the accessible points L_1, M_1

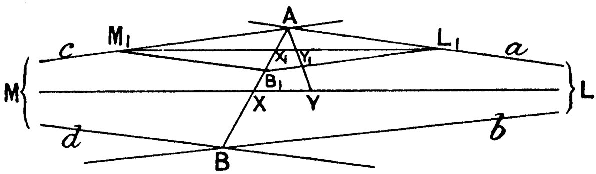

FIG. 55.

respectively. Then in the smaller figure determine the point X_1 corresponding to X, by drawing the diagonal L_1M_1 corresponding to LM to meet AB_1 in X_1. Take X on AB, making $AX = \lambda AX_1$; then X is a point on LM, and the straight line XY through X parallel to L_1M_1 is the straight line required. Or we may determine a second point Y upon it by taking any point Y_1 on L_1M_1 and producing AY_1 to Y, making $AY = \lambda AY_1$. It is convenient to take λ to be a power of 2, for then B, X and Y are found by the processes of halving and doubling repeated as often as necessary.

Ex. 3. To bisect an angle whose vertex L is inaccessible.

(i) We may use the property of the inscribed circle of a triangle. Take any points A, B on **a**, **b**; join AB,

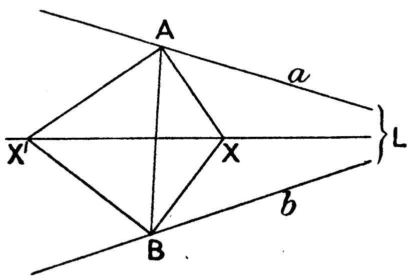

FIG. 56.

and bisect the angles LAB, LBA by straight lines that meet in X. Then X is the incentre of the triangle ALB, and lies upon the bisector of the angle L. To obtain a second point X′ on the bisector, we can most easily bisect the exterior angles at A, B by straight lines that meet in X′ the centre of the escribed circle of ALB opposite L. Then XX′ is the required bisector.

(ii) Since the bisector is the locus of a point equidistant from the given arms **a**, **b**, we can obtain a position of

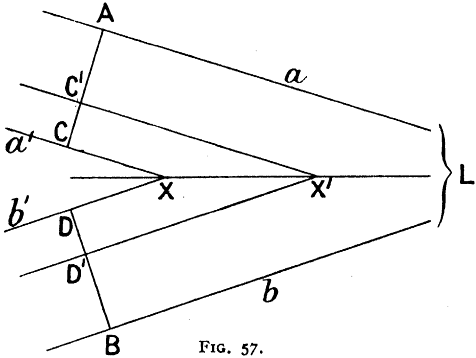

FIG. 57.

X as the intersection of parallels **a**′, **b**′ drawn at the same perpendicular distance from **a**, **b** respectively. Take any point A on **a**, draw AC perpendicular to **a** and take any point C on it. Through C draw CX parallel to **a**. Take any point B on **b**, draw BD perpendicular to **b**, making

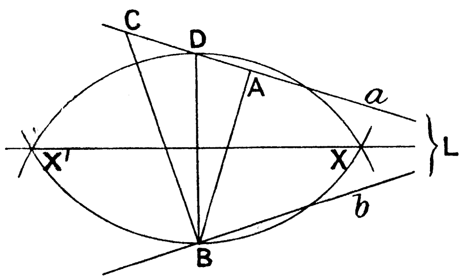

FIG. 58.

BD = AC, and draw DX parallel to **b**. From the same two perpendiculars cut off other equal lengths AC′, BD′, and draw C′X′, D′X′ parallel to **a**, **b** respectively to meet in X′. Join XX′, which is the required bisector of the angle between **a**, **b**.

Instead of constructing X′ in this way, we may use the property that XX′ is also the bisector of ∠CXD. This is the method of translation.

(iii) Since XX′ is the axis of symmetry of the pair of arms **a**, **b**, it bisects at right angles any intercept BD which is equally inclined to **a**, **b**. To construct this axis, take any point A on **a**, draw AB perpendicular to **a** to meet **b** in B; draw BC perpendicular to **b**; draw BD to bisect ∠ABC and to meet **a** in D. Then XX′, the perpendicular bisector of BD, is also the bisector of the angle between **a**, **b**; for BD, being equally inclined to AB, CB, is also equally inclined to **a**, **b**, which are perpendicular to them respectively; LBD is an isosceles triangle, and XX′, which bisects the base at right angles, also bisects the vertical angle DLB.

Ex. 4. To draw tangents to a circle from an inaccessible point.

(i) If the point L is given as the intersection of two straight lines which meet the circle in real points, let these intersections be A, B; C, D. If not, take points A, C on the circumference, and by ex. 1 draw AB, CD to pass

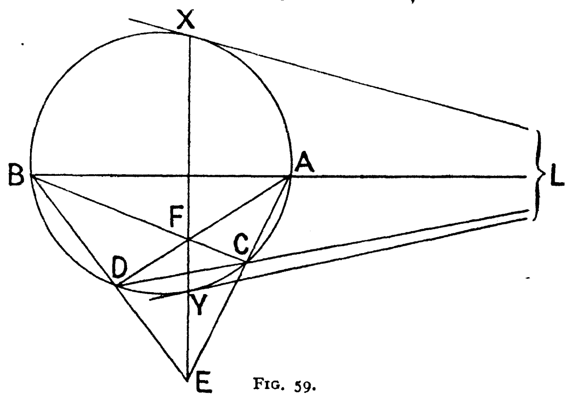

FIG. 59.

through L and meet the circumference again in B, D respectively. Let AC, BD meet in E, and AD, BC in F. If we arrange that AB is nearly a diameter and CD nearly a tangent, then both E and F are accessible. Join EF to meet the circle in X, Y. Then these are the points of contact of the tangents from L to the circle, and the tangents themselves are constructed as the perpendiculars to the radii at X, Y. For the diagonal EF of the complete quadrilateral AB, AD, BC, BD divides harmonically the two chords LAB, LCD; it is therefore the polar of L, and passes through the points of contact of the tangents from L to the circle.

(ii) Or we may use similar figures with O the centre of the circle as centre of similitude. Reduce the scale

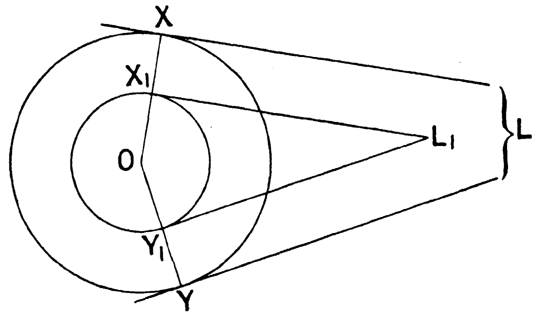

FIG. 60.

until the point L_1 corresponding to L becomes accessible, and draw tangents L_1X_1, L_1Y_1 from L_1 to the smaller circle. Then OX_1, OY_1 meet the larger circle in X, Y, the points of contact of the required tangents from L.

Ex. 5. To draw the common chord of the circles c_1, c_2, which meet in real but inaccessible points.

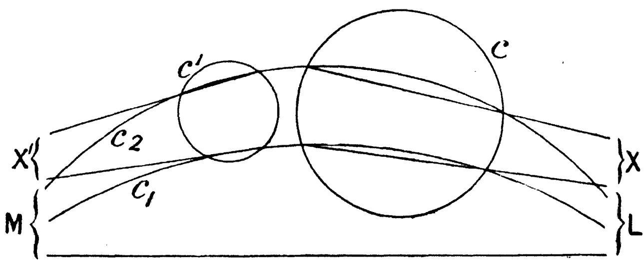

FIG. 61.

Take an auxiliary circle c which meets c_1, c_2 in real and accessible points. Then the two common chords of c, c_1;

c, c_2 meet in X, the radical centre of the three circles, which lies upon the required common chord of c_1, c_2. By using another auxiliary circle c', we can construct another point X′ of the common chord. Both X and X′ are likely to be inaccessible, and we may have to construct the straight line joining XX′ by the methods of ex. 2.

Ex. 6. One circle c_1 is partly drawn, and another c_2 is known to pass through three given points A, B, C ; both centres are inaccessible. It is required to draw the radical axis.

This can be done by the method of the last example, if c is taken to pass through A, B, and c' through B, C. The

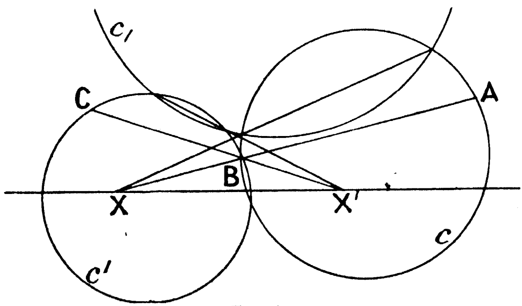

FIG. 62.

common chords of c, c_2 and c', c_2 are then the straight lines AB and BC, so that we need not have the circle c_2 drawn. Care must be taken that c, c' meet c_1 in points lying on the arc that is drawn.

II. Ill-defined Intersections.

It is always well to economize the chance of error: among many ways in which errors may arise, one of the most frequent is in determining the point of intersection of two lines, straight or curved, which meet at a very small angle. In an actual figure, in which lines are represented by strips of small breadth b say, an intersection is only determined as some point within the rhombus which is common to the two strips. If the two directions are at right angles, the rhombus is a square, and the distance of any one of its points from the centre is not greater than $\frac{1}{2}\mathfrak{b}\sqrt{2}$; but if the two directions make a very small angle ϵ, the greatest distance of a point of the rhombus from the centre is $\frac{1}{2}\mathfrak{b}\operatorname{cosec}\frac{1}{2}\epsilon$, which has no upper limit as ϵ

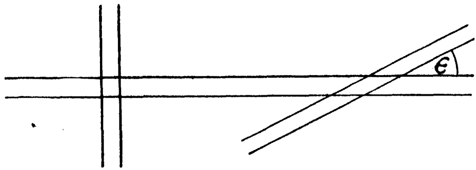

FIG. 63.

decreases, and may be a source of quite sensible error. It may be well therefore in some cases to replace such ill-defined intersections by others for which the conditions are better.

Ex. 1. The first problem is: given two straight lines AB, CD, which meet at X at a small angle, to find a straight line XY through the point of intersection meeting either at a moderate angle. We may take AC parallel to BD; for if A, C are any points on the two straight lines, and we suppose we have a bisected segment on AC, the ruler construction of p. 47, for drawing a parallel to AC through any other point B of the first straight line, does not depend on any ill-defined intersections when B is very close to AC. We can take B on AX produced, and then the problem

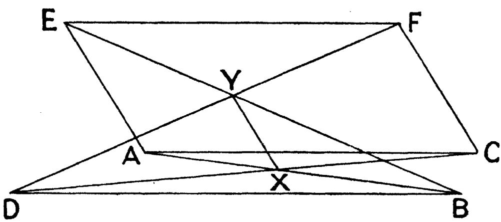

FIG. 64.

takes the form: to determine the internal diagonal point X of a given narrow trapezium ACBD.

Draw any parallelogram AEFC on AC as one side. Join BE, DF to meet in Y; through Y draw a parallel to AE or CF; then this parallel passes through X, the point of intersection of AB, CD. For if XY is joined, by similar triangles,

$$\frac{AX}{XB} = \frac{AC}{BD} = \frac{EF}{BD} = \frac{EY}{YB};$$

therefore XY is parallel to AE.

Ex. 2. To determine the two intersections X, Y of a straight line AB with a circle which it nearly touches.

Let C be the centre, and Z the pole of AB with regard to the circle. Then X, Y are the points of contact of tangents from Z to the circle, the angles CXZ, CYZ are right angles and CXZY lie on a circle whose centre is O the midpoint of CZ. Now the angle between AB and the tangent XZ to the given circle at X is equal to ∠ZCX between their perpendiculars CZ and CX; and similarly, the angle between AB and the circle O(X) is equal to ∠ZOX. But ∠ZOX at the centre of the circle O(X) is twice ∠ZCX at its circumference; hence the point X is determined by the intersection of AB with the circle O(X) with twice as much accuracy as it is determined by the intersection of AB with the given circle.

In order to construct first Z and then X, Y, take any point A on the given straight line, and on AC as diameter

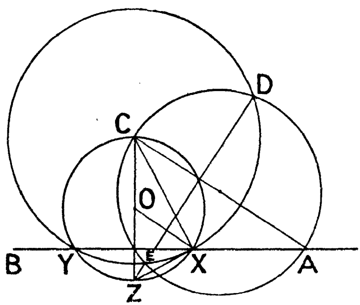

Fig. 65.

describe a circle to meet the given circle in D, E; join DE, which is the polar of A, and therefore passes through Z the pole of AB. Draw CZ perpendicular to AB to meet DE in Z; on CZ as diameter describe a circle to meet AB in the required points X, Y.

Ex. 3. To determine the second intersection X of a straight line with a circle, when one intersection A is given and the straight line is nearly a tangent.

Draw a parallel chord BC, and cut off an arc CX from the circumference equal to AB. We must take BC sufficiently far from AX for B and C to be well defined. If we

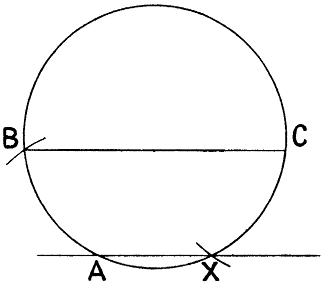

FIG. 66.

make AB and CX equal to the radius of the circle, then each of the points B, C, X is determined as the intersection of two lines that meet at an angle of nearly $\frac{1}{3}\pi$.

III. Geometrography.

If we are not specially restricted either by lack of space or by blunt tools, it is well to economize the number of operations by which a construction is carried out. The question arises at once as to what is a single operation, and how we can compare different operations with ruler and compasses. The answers to these questions are to a large extent arbitrary; we have to adopt some fixed scale of values, and decide upon the *geometrographical* or most economical construction on that understanding. The simplest way of reckoning is to count the number of straight lines and circles required by a construction, and to take this total number as a rough gauge of the amount of labour expended. Or we may consider that to describe a circle is more laborious than to rule a straight line. If we reckon that each of these two operations involves a fixed expenditure of energy, and that these are in the ratio R : C, then if a construction requires L straight lines and m circles, its measure is the quantity LR + mC, and among different constructions for the same problem, those are best for which this measure is least. If we take $\frac{C}{R} = \infty$, we try to avoid the use of compasses, so as to keep m as small as possible, however much this may entail an increase in L. We have already discussed what constructions are possible with ruler alone, which is the case $m = 0$; and m need never be greater than 1 (see chapter VII). If we take $\frac{R}{C} = \infty$, and avoid as far as possible the use of a ruler, we can always have $L = 0$ (see chapter VIII).

But we may distinguish between various operations with the same instrument. As long as we restrict ourselves to definite constructions, whenever we use a ruler we have to make its edge coincide with two given points and then rule along the edge; but often we only need to draw an indefinite straight line with one or no coincidence; and these are simpler operations. Again, to describe the circle $C(AB)$ requires us to make a compass leg coincide with a fixed point three times: first with the two ends A, B of the given radius and then with the centre C. But to describe the circle $C(A)$ only requires two coincidences, one with the centre C and one with the given point A of the circumference; and to describe a circle with given centre and any radius, or any circle through a given point, only requires one coincidence. Also, with modern compasses (p. 71), when we describe two circles in succession with the same radius, the second only requires one coincidence, at its centre.

Lemoine has worked out an elaborate system of geometrography. He distinguishes the following operations:

R_1 make the straight edge pass through one given point,
R_2 rule a straight line,
C_1 make one compass leg coincide with a given point,
C_2 make one compass leg coincide with any point of a given line,
C_3 describe a circle.

Then if these operations occur respectively L_1, L_2, m_1, m_2, m_3 times in a construction, its *symbol* is

$$L_1R_1 + L_2R_2 + m_1C_1 + m_2C_2 + m_3C_3;$$

the sum of all the coefficients $L_1 + L_2 + m_1 + m_2 + m_3$, which is the total number of operations, is the *coefficient of simplicity*, and the total number of coincidences $L_1 + m_1 + m_2$ is the *coefficient of exactitude.* As Lemoine remarks, these coefficients really measure, not the simplicity and exactitude, but the complication and the chance of error. The number of straight lines is L_2 and of circles m_3, so that the total number of lines drawn is $L_2 + m_3$, which is the difference between the coefficients of simplicity and exactitude. To join two points A, B is a complex operation, which has the symbol $2R_1 + R_2$; to describe the circle C(AB) has the symbol $3C_1 + C_3$.

With these conventions, Lemoine proceeds to give, for a large number of problems, the geometrographical constructions, those for which the coefficient of simplicity is least. It is only in easy cases that we can assert that a given construction is as simple, in the sense defined, as any other possible method; in general we can only say that it is as simple as any other known method. Further, the construction which is simplest for a certain problem standing alone may not be the simplest when the same problem occurs as part of another; for a construction less simple in itself may fit in better with the context, by making use of lines which are already drawn or which can be used later on in the course of the construction of the more complicated problem.

Other systems, more or less elaborate than this, could no doubt be given, and justified from different points of view. We shall keep to these definitions in the examples that follow in this chapter.

Ex. 1. Through a given point A to draw a straight line parallel to a given straight line BC.

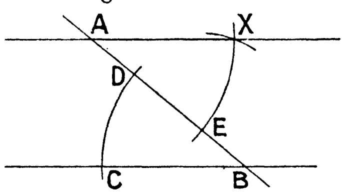

FIG. 67.

(i) A usual construction is by means of equal alternate angles. Draw any straight line through A to meet BC in B (symbol $R_1 + R_2$). With centre B and any radius ρ describe the circle B(ρ) to meet BC in C and BA in D ($C_1 + C_3$). Describe the circle A(ρ) to meet AB in E ($C_1 + C_3$), and describe E(CD) to meet A(ρ) in X ($3C_1 + C_3$). Join AX ($2R_1 + R_2$), which is the required parallel. The symbol of the whole construction is $3R_1 + 2R_2 + 5C_1 + 3C_3$; simplicity **13**, exactitude **8**.

(ii) Compare the following construction. With any centre D describe D(A) to meet the given straight line in

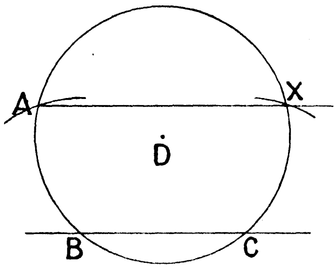

FIG. 68.

B, C ($C_1 + C_3$). Describe C(AB) to meet D(A) in X ($3C_1 + C_3$); join AX ($2R_1 + R_2$). Symbol $2R_1 + R_2 + 4C_1 + 2C_3$; simplicity **9**, exactitude **6**.

(iii) Or compare this. With any radius ρ describe the circles A(ρ) to meet BC in B; B(ρ) to meet BC in C; C(ρ) to meet A(ρ) in X. Join AX. Then ABCX is a rhom-

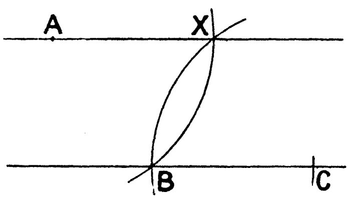

FIG. 69.

bus and AX is parallel to BC. Symbol $2R_1 + R_2 + 3C_1 + 3C_3$; simplicity **9**, exactitude **5**. Though this construction has one more line than the last, it gains by using the same radius (and that arbitrary) for all the circles, instead of two different radii, one of them assigned.

Ex. 2. Let us analyse the constructions already given for the radical axis of the circles whose common points are either imaginary or inaccessible. We suppose the centres A, B given.

(i) In the construction of p. 79, ex. 2, we begin by joining AB ($2R_1 + R_2$) to meet the circles in C, D. Then we have to draw CE perpendicular to AC. The shortest method for this perpendicular is to describe any circle G(C) through C to meet AC in H, and join HG to meet G(C) in E. Then CE (which need not be drawn) is perpendicular to AC. That would be the best construction for

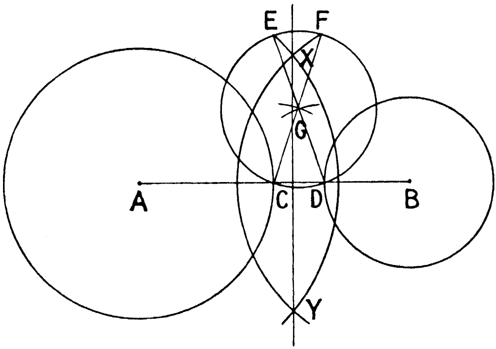

Fig. 70.

a point E on the perpendicular if there were no further construction to be made; but as we also require the second perpendicular DF, it is more economical to make the auxiliary circle pass through D as well as C, so as to do for both. Instead of taking the centre G arbitrary, we take it at the intersection of two other circles C(ρ), D(ρ), where ρ is arbitrary, ($2C_1 + 2C_3$), and then describe the circle G(ρ) ($C_1 + C_3$), which passes through C and D. Then the diameters CGF, DGE must be drawn ($4R_1 + 2R_2$), which not only gives both the perpendiculars CE, DF (which need not be drawn), but also gives them of equal length, for the figure GCDEF is symmetrical. It remains to describe the circles A(E), B(F) ($4C_1 + 2C_3$) to meet in X, Y, and to join XY ($2R_1 + R_2$), which is the required radical axis. The whole symbol is $8R_1 + 4R_2 + 7C_1 + 5C_3$; simplicity **24**, exactitude **15**.

(ii) In the construction of p. 108, ex. 5, we obtain the point X on the radical axis by using one arbitrary circle and two common chords; then X′ is found in the same way, and XX′ is joined. Symbol $10R_1+5R_2+2C_3$; simplicity **17**, exactitude **10**.

(iii) In this last construction, if the centres A, B are given, instead of X′ we can find the reflexion Y of X in AB, as the other intersection of the circles A(X), B(X), and then join XY. Symbol $6R_1+3R_2+4C_1+3C_3$; simplicity **16**, exactitude **10**.

Ex. 3. To bisect an angle whose vertex is inaccessible.

Compare the methods of p. 105, ex. 3, with this: With any point C of one arm as centre and any convenient radius ρ describe C(ρ) to meet the arms AB, CD in A, D.

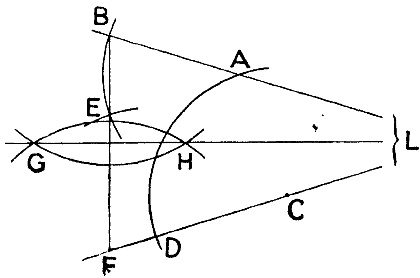

FIG. 71.

Describe A(ρ) to meet AB in B, and describe D(ρ) to meet A(ρ) in E. Join BE to meet CD in F. Describe B(ρ), F(ρ) to meet in G, H; join GH, which is the required bisector.

For the figure CAED is a rhombus of side ρ, and BAE is an isosceles triangle with sides parallel to the given arms; BEF is perpendicular to the required bisector, which must also bisect BF. Symbol $4R_1+2R_2+4C_1+C_2+5C_3$; simplicity **16**, exactitude **9**.

CHAPTER VII.

ONE FIXED CIRCLE.

ANY point that can be constructed with ruler and compasses can be constructed with ruler only if one fixed circle and its centre are given. This theorem is due to Poncelet and Steiner, and it is completely proved by the methods given on pp. 60, 61 for constructing any quadratic surd by means of one circle; and any problem may be solved by reducing its solution to that of a series of quadratic equations, and carrying out for each equation one of the constructions referred to. We shall consider some of the simpler problems that present themselves and their best solutions, meaning by "best," in this chapter, those that require fewest straight lines.

We call the fixed circle Γ and its centre O. If the centre is not given, we must have sufficient data to construct it; for example, a parallelogram, for then we can draw parallel chords and bisect them, and so obtain diameters. When O is found, all the metrical data are at hand.

The methods of this chapter are appropriate in connection with the method of trial and error (p. 80), or others which require the common points of a homography, which can be found by projecting on to the fixed circle. They are also of use in problems connected with the centres of similitude of two circles, in a way which is fully illustrated below.

We shall first consider separately the best methods for drawing parallels, which are required for all the fundamental constructions.

Lemma. Through a given point R *to draw a parallel to a given straight line* PQ.

We have first to obtain a bisected segment PZQ on the given straight line, and then construct a complete quadrilateral with one diagonal point at infinity, as on p. 47.

Case (i). PQ passes through O, the centre of Γ. The problem is to draw through R a parallel to a given diameter of Γ.

Then if P, Q are the ends of the diameter, POQ is a bisected segment already given, and we complete the construction thus. Join PR, QR ; on PR take any point U. Join UO

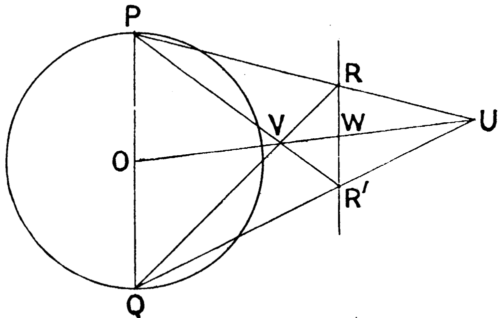

FIG. 72.

to meet QR in V ; join QU, PV to meet in R′. Then RR′ is the required parallel ; for UV, RR′ divide harmonically the third diagonal PQ of the quadrilateral PR, PR′, QR, QR′ ; and since UV bisects PQ, therefore RR′ meets it at infinity, so that PQ, RR′ are parallel. The figure consists of a quadrilateral and its three diagonals, that is, of seven straight lines, one of which is given.

An important aspect of this figure is when RR′ is regarded as the diameter of a circle γ, which is not drawn. Since the diagonals RR′, UV of the quadrilateral meet in W the midpoint of RR′, this point W is the centre of the circle γ, and U, V are the external and internal centres of similitude of Γ, γ. In the first of these similitudes, P corresponds to R and Q to R′ ; in the second, P corresponds to R′ and Q to R.

Case (ii). PQ and R have general positions.

In order to obtain a bisected segment PZQ, we first draw any straight line OZ through O to meet PQ in Z ; then, by case (i), a parallel chord ST, taking care that its intersections S, T with Γ are real. Join SO, TO and produce them to meet the circumference again in S′, T′; then S′T′ is the reflexion of ST in OZ, so that these three straight lines are

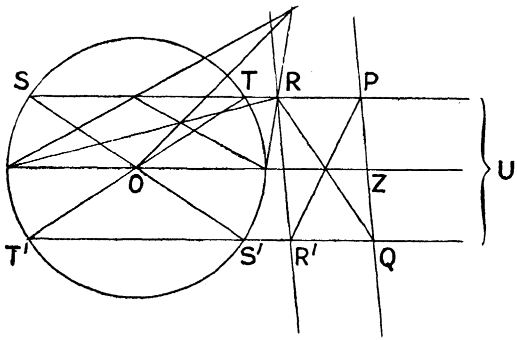

FIG. 73.

parallel and equidistant and cut off a bisected segment PZQ on the given straight line PQ. Having this we can draw the parallel RR′ as before.

For this construction we require two complete quadrilaterals, each with its three diagonals, in order to draw the two pairs of parallels OZ, ST and PQ, RR′; in addition we have the two diameters SS′, TT′ and the third parallel S′T′: if the figure is drawn without precaution, it may contain seventeen straight lines. But this may be reduced by making some lines serve two purposes. If we take OZ to pass near enough to R, the chord ST may be drawn through R and still cut Γ in real points. Then the second quadrilateral may be reduced to a parallelogram, by taking U at infinity; for then UP, UZ, UQ coincide with the three parallels already drawn; the second quadrilateral has two sides and one diagonal already in the figure, and only requires four more lines. The whole figure contains fourteen straight lines.

Case (iii). R coincides with O. The problem is to draw the diameter of Γ parallel to a given straight line.

The simplification just given fails if R coincides with O, for we cannot draw ST through the centre of Γ. The best we can do is to make one side of the second quadrilateral, QR′U say, coincide with one of the parallel chords, S′T′ say. The whole figure contains sixteen straight lines.

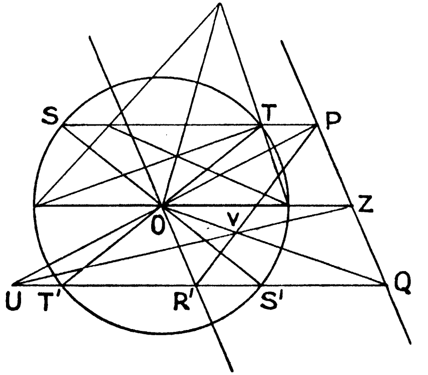

FIG. 74.

Thus when a circle γ is given by means of its centre **o** and a point **a** on its circumference, we can first draw the diameter AOB of Γ parallel to **ao**. Then to obtain other points on the circumference of γ, we make use of E, I, the external and internal centres of similitude of Γ, γ. These are the intersections of the line of centres Oo with Aa and

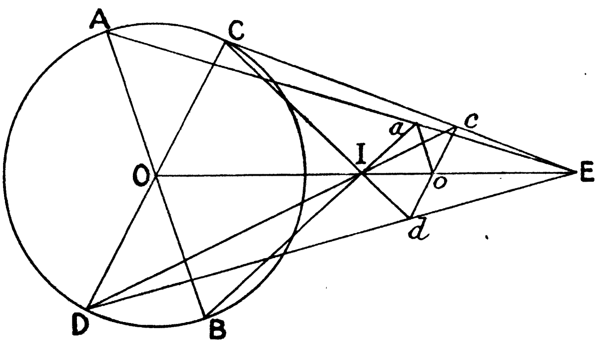

FIG. 75.

Ba respectively. Then if COD is any other diameter of Γ, the end **c** corresponding to C of the parallel diameter **cod** of γ, in the similitude whose centre is E, is the point of intersection of CE, DI; and **d** is the intersection of DE, CI.

Fundamental Problems.

In a general ruler and compass construction, new points are determined in one of three ways :

(i) as the intersection of two straight lines : this method we may use freely in this chapter ;

(ii) as the intersection of a straight line and a circle : this method we may use under the restriction that the circle must be Γ and no other ;

(iii) as the intersection of two circles : this we may not use at all.

We must therefore devise other methods for the two following problems, which are fundamental, and whose solution gives another proof of the theorem at the beginning of this chapter :

I. To determine the points of intersection of a given straight line and a given circle other than Γ.

II. To determine the points of intersection of two given circles, both in case (i) in which one circle is Γ, and is therefore already drawn, and in case (ii) in which neither circle is Γ, so that neither is to be actually drawn.

Throughout, a circle must be considered to be given or obtained when its centre and a point on its circumference are known ; but no circle except Γ must be drawn.

Problem I. To find the intersections **x**, **y** *of a given straight line* **fg** *with a circle* γ *other than* Γ.

The straight line **fg** is supposed drawn, and the circle γ given by means of its centre **o** and a point **a** on its circumference, that is, one radius **oa** is given in position and magnitude. The method employed is that of similar figures. We construct a straight line FG, which with Γ makes a figure similar to that composed of **fg** and γ ; then if FG meets Γ in X, Y, the points **x**, **y** corresponding to these in the similitude are the required intersections of **fg** and γ. The first step is to construct as above E, I the external and internal centres of similitude of Γ, γ. We use the direct similitude, which has E as centre ; then any straight line through E meets any two parallel radii of Γ, γ in corresponding points. Now suppose we have two pairs of parallel radii, AO, **ao** and CO, **co**, and let **f**, **g** be the intersections of the given straight line with **ao**, **co**. Then the corresponding points are F the intersection of **Ef**, AO, and G the intersection of **Eg**, CO. Then if FG meets Γ in X, Y, the required points **x**, **y** are the intersections of **fg** with EX, EY.

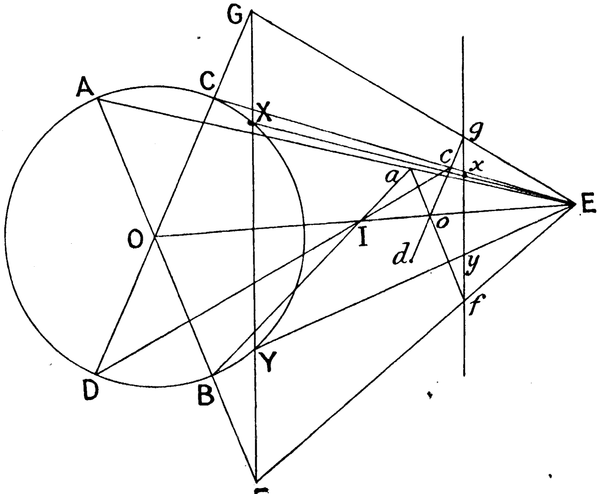

FIG. 76.

The construction succeeds or fails according as FG meets Γ in real or imaginary points; that is, according as **fg** meets γ in real or imaginary points. Hence if **x**, **y** are real, they can always be constructed by this method.

Construction in Detail.

The whole figure contains twenty-seven straight lines. We start with the one straight line **fg** already drawn. Then we carry out the construction of lemma (iii) in order to obtain the diameter AOB of Γ parallel to **ao**; this introduces sixteen straight lines. Now we can arrange that two of these sixteen, besides **ao**, AO, have positions that are used again; for the diameter of Γ first to be drawn (OZ of fig. 74) can coincide with Oo, and one of the other diameters (SS′ say of the same figure) can be used for the arbitrary second diameter CD of Γ which is required. The point **f** is now determined as the intersections of **oa, fg**. The rest of the construction requires ten more lines, given by this table :

Aa	Ba	CE, DI	**oc**	Ef	E**g**	FG	EX	EY
Oo	Oo		**fg**	OA	OC	Γ	**fg**	**fg**
E	**I**	**c**	**g**	F	G	X, Y	**x**	**y**

The first row gives the new straight lines in the order in which they are drawn, and the third row gives underneath each the new point or points determined by it. Where the new point is determined by intersection with a line already drawn, this is shown in the second row in the same column.

The straight line **oc** parallel to the diameter CD of Γ is really determined by the method of lemma (i) ; but in the complete quadrilateral Cc, Cd, Dc, Dd, of which COD, **cod** are parallel diameters, we use the harmonic range Oo, IE found independently on the third diagonal IE, and obtain **c** without having to draw the two sides Cd, Dd of this quadrilateral.

Problem II case (i). *To find the intersections of* Γ *and a given circle* γ.

This problem is the same as to find the radical axis of Γ, γ ; for this straight line is their common chord if they meet in real points, and it meets Γ in X, Y, the required intersections of the two circles. Now we can obtain two points on the radical axis most quickly as follows.

Let **o** be the centre and **co** the given radius of γ. Draw the parallel diameter COD of Γ and construct E, **I**, the external and internal centres of similitude of the two circles. Let EC, ED meet Γ again in H, K. Determine the points **d**, **h**, **k** of γ corresponding to D, H, K ; let CK, **hd** meet in L, and DH, **kc** in M. Then LM is the radical axis and meets Γ in the required points X, Y.

This follows from the fact that the four points C, K, **d**, **h** lie on an auxiliary circle, and L is the intersection of CK its common chord with Γ and **hd** its common chord with γ. For in fig. 77

$$\angle hCK = \angle hck, \text{ by parallels,}$$
$$= \angle hdk, \text{ in the circle } \gamma,$$
$$= \pi - \angle hdK;$$

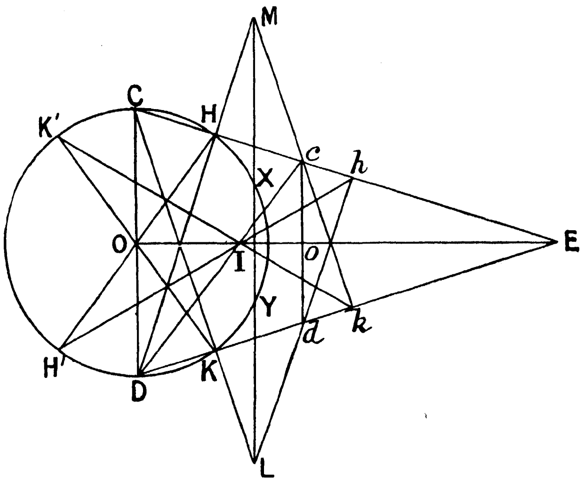

FIG. 77.

hence the points C, K, **d**, **h** are concyclic. Similarly, M lies on the radical axis of Γ, γ.

Construction in Detail.

We are given Γ and the two points **o**, **c**. Draw the diameter COD of Γ parallel to **co** by lemma (iii). The figure now contains sixteen straight lines, including **co**, COD, and one (OZ of fig. 74) which, as before, can be taken to coincide with Oo. The rest of the construction is given by the table:

Cc	DE	Dc	HO	KO	H′I	K′I	CK, **hd**	DH, **kc**	LM
Oo, Γ	oc, Γ	Oo	Γ	Γ	Cc	DE			Γ
E, H	**d**, K	I	H′	K′	**h**	**k**	L	M	X,Y

The whole figure contains twenty-eight straight lines.

Case (ii). *To find the intersections of two given circles* γ, γ_1.

We reduce this to depend on problem I by first finding **lm**, the radical axis of the two circles. Then the points of intersection **x**, **y** of **lm** with one of the circles, γ say, are the required points of intersection of γ, γ_1. We have therefore to determine two points **l**, **m** on this radical axis. Each is obtained as the intersection of two chords which meet the circles respectively in four concyclic points. We first obtain two parallel diameters **cod**, $\mathbf{c_1o_1d_1}$ of γ, γ_1. Then $\mathbf{cc_1}$, $\mathbf{dd_1}$ meet in the external centre of similitude of γ, γ_1; let the other points of intersection of γ with $\mathbf{cc_1}$, $\mathbf{dd_1}$ be **h**, **k**. Then as above (II (i) and figs. 77, 81), the four points **h**, **k**, $\mathbf{c_1}$, $\mathbf{d_1}$ are concyclic, and **hk**, $\mathbf{c_1d_1}$ meet in **l** on the radical axis of γ, γ_1. Similarly, if $\mathbf{cc_1}$, $\mathbf{dd_1}$ meet γ_1 in $\mathbf{h_1}$, $\mathbf{k_1}$, and **cd**, $\mathbf{h_1k_1}$ meet in **m**, this is another point on the radical axis, which is therefore the straight line **lm**. It is to be noticed that the two quadrilaterals **ckdh** and $\mathbf{c_1k_1d_1h_1}$ are similar and similarly situated.

But the points **h**, **k**, $\mathbf{h_1}$, $\mathbf{k_1}$ cannot be directly obtained as the intersections of $\mathbf{cc_1}$, $\mathbf{dd_1}$ with the circles γ, γ_1, since the latter are not to be described. We have to obtain them as in problem I, by means of two similitudes between Γ, γ and between Γ, γ_1 respectively; we use those with the external centres E, E_1. The quadrilaterals **ckdh**, $\mathbf{c_1k_1d_1h_1}$, inscribed in γ, γ_1 respectively, that we wish to construct, are similar (in these two similitudes) to the same quadrilateral CKDH inscribed in Γ (fig. 80). We require two pairs of parallel diameters of Γ, γ, and we take them parallel to the given radii **oa**, $\mathbf{o_1c_1}$ of γ, γ_1; these same pairs of parallels can be used again later to determine the intersections of **lm** with γ, which are the required points **x**, **y**.

The construction can be divided into stages as follows:

Stage (i), fig. 78. Construct, by lemma (iii), the diameters AOB, COD of Γ parallel to the given radii **ao**, $\mathbf{c_1o_1}$ of the given circles γ, γ_1.

Stage (ii), fig. 79. Determine, by means of the centres of similitude E, I, E_1, the other ends **b**, $\mathbf{d_1}$ of the diameters **aob**, $\mathbf{c_1o_1d_1}$ of γ, γ_1, and also the diameter **cod** of γ parallel to COD or $\mathbf{c_1o_1d_1}$. We have now the two pairs of parallel diameters AOB, **aob** ; COD, **cod** of Γ, γ, and by means of intersections with these we can find the chord of either circle corresponding to any chord of the other.

Stage (iii), fig. 80. Draw $\mathbf{cc_1}$, $\mathbf{dd_1}$, and determine their other intersections **h**, **k** with γ and $\mathbf{h_1}$, $\mathbf{k_1}$ with γ_1, by means of the corresponding chords CH, DK of Γ.

Stage (iv), fig. 81. Determine the points **l**, **m**, by means of the cyclic quadrilaterals $\mathbf{hkc_1d_1}$, $\mathbf{cdh_1k_1}$, and so construct the radical axis **lm** of γ, γ_1.

Stage (v), fig. 82. Determine, by problem I, the intersections **x**, **y** of **lm** with γ.

Construction in Detail.

We are given : the circle Γ fully drawn and its centre O ; the centres **o**, $\mathbf{o_1}$ of the given circles γ, γ_1 and the points **a**, $\mathbf{c_1}$ on their respective circumferences.

Stage (i). Draw the diameter AOB of Γ parallel to the straight line joining **ao**, by the method of lemma (iii).

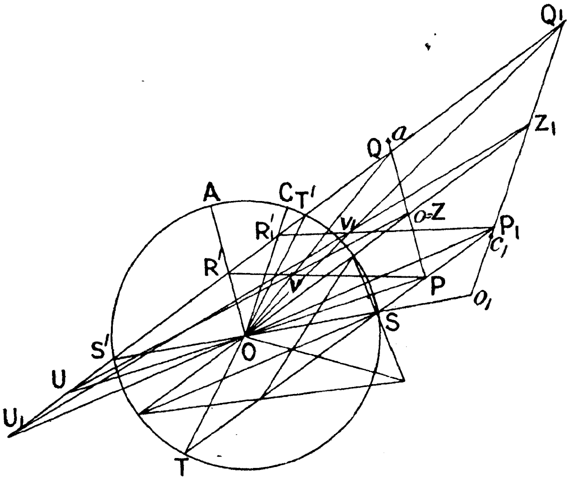

FIG. 78.

This requires the sixteen straight lines of fig. 74. Next, draw COD parallel to $\mathbf{c_1o_1}$; this also requires sixteen straight lines, but we can use the same set of three parallels as before, and therefore the same set of ten lines: namely, those called OZ, ST, S′T′, SS′, TT′ in figs. 74 and 78, and the five lines, not lettered, used in the construction of ST. The number of straight lines drawn at this stage is therefore $2 \times 16 - 10 = 22$.

These include ao, c_1o_1, AO, CO, which are needed later; we can also arrange for it to include Oo and Oo_1. For to start with, the point Z of fig. 74 can be taken at o, and then S can be taken at a point of intersection of Oo_1 with Γ, so that Oo, Oo_1 of fig. 78 coincide with OZ and SOS′ of fig. 74 respectively.

The constructions of the remaining stages are given by the following diagrams and tables. In order to keep the figure as clear and compact as possible, we use in stage (v) the inverse similitude between Γ, γ, whose centre is I, instead of the direct similitude, whose centre is E.

Stage (ii).

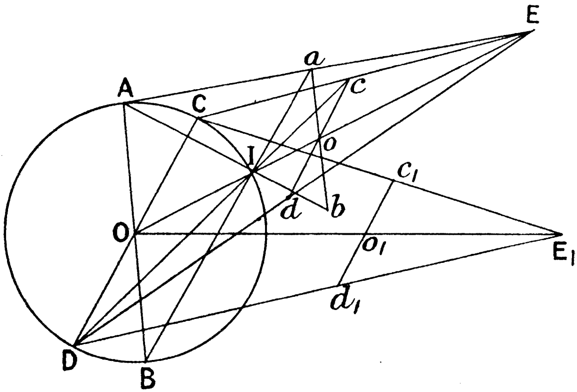

FIG. 79.

Aa	Ba	Cc_1	AI	DE_1	CE, DI	DE, co
Oo	Oo	Oo_1	ao	c_1o_1		
E	I	E_1	b	d_1	c	d

Stage (iii).

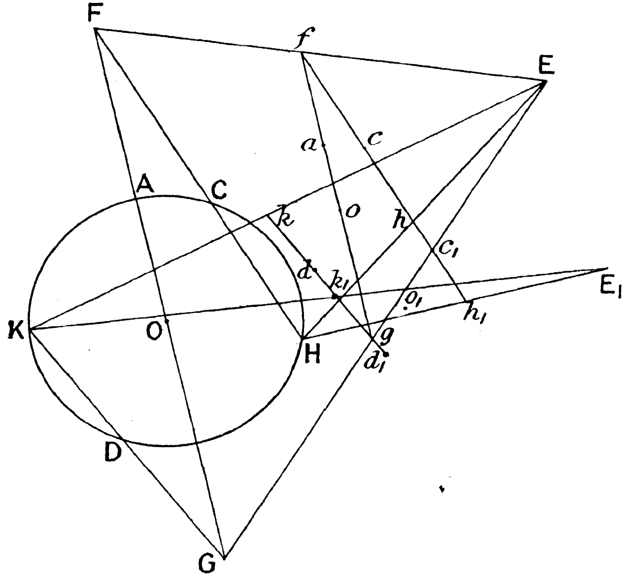

FIG. 80.

cc_1	dd_1	Ef	Eg	CF	DG	EH	EK	E_1H	E_1K
ao	ao	AO	AO	Γ	Γ	cc_1	dd_1	cc_1	dd_1
f	g	F	G	H	K	h	k	h_1	k_1

Stage (iv).

hk	h_1k_1	lm
c_1o_1	co	ao
l	m	n

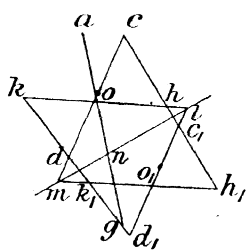

FIG. 81.

Stage (v).

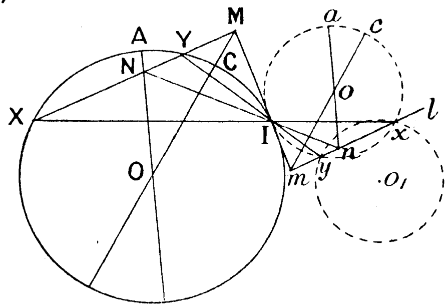

FIG. 82.

In	Im	MN	IX	IY
AO	CO	Γ	Im	Im
N	M	X, Y	x	y

The total number of straight lines is therefore

$$22 + 9 + 10 + 3 + 5 = 49,$$

but it is very possible that this might be reduced by greater ingenuity.

CHAPTER VIII.

COMPASSES ONLY.

If our only ruler were badly chipped or warped, ruling a straight line would be an altogether less desirable operation than describing any number of circles. We should then wish to avoid as far as possible having to determine a point as the intersection of two straight lines ; Mascheroni and Adler have shown that we can avoid it altogether.

Of course, no straight line can be obtained in its entirety, if we make use of compasses only ; it must be considered as given or obtained when two points lying on it are known. It is the main object of this chapter to show that we can then obtain, still with compasses only, any other points of the straight line that we may want, that is, its points of intersection with any given circle, or with any straight line given in the same way by means of two points lying on it.

Mascheroni's methods are the older ; generally speaking, they are based on the idea of reflexion in a straight line. Some of his constructions are shorter than Adler's, mainly because Mascheroni uses his compasses freely as dividers. Adler's methods are based on the idea of inversion with regard to a circle. They are much more systematic, and more powerful as regards the theory, and they use Euclidean compasses throughout. The single operation, according to Mascheroni, of describing the circle C(AB) would have to be replaced, if we use Euclidean compasses, by a construction requiring five circles. The following is as short as any :

Describe the circles A(C), C(A) to meet in D, E. Then DE is the axis which bisects AC at right angles. Describe

D(B), E(B) to meet again in X. Then X is the reflexion of B in the axis DE, and CX is equal to its reflexion AB. Describe C(X); then this is the circle required, for its centre is the required point C and its radius CX is equal to the required radius AB.

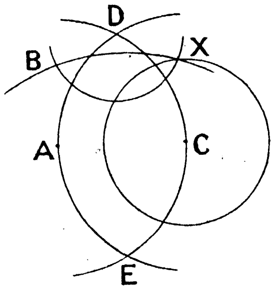

FIG. 83.

This double process, of first determining an axis, and then reflecting in it part of the figure already obtained, will occur later several times. When it forms part of another construction, we can usually arrange that some of the five circles are used either earlier or later in other parts of the construction.

Fundamental Problems.

Mascheroni's proof of the possibility of his method consists in actually giving compass constructions to replace those two, out of the three standard methods of determining points (p. 122), which involve the use of a ruler. All points, other than the data, whether required by the problem or used in the course of construction, must be determined by the third standard method only, as the intersections of pairs of circles. Just as in the last chapter, we have the two fundamental problems:

I. To determine the points of intersection of a given circle and the straight line joining two given points.

II. To determine the point of intersection of the two straight lines joining two given pairs of points.

When these are solved, we have, theoretically at least, a compass construction for any problem for which a ruler and compass construction is known. We have only to replace each step of the latter, which requires the use of a ruler, by the solution of the corresponding problem I or II.

But before giving the solutions of these fundamental problems, let us consider Adler's discussion of the matter as a whole. His proof is based on the theory of inversion. If any figure consisting of straight lines and circles is inverted with regard to a circle whose centre does not lie on any line of the figure, then the inverse figure consists entirely of circles, p. 93 (3), (4). If the original figure is that belonging to a known ruler and compass construction of any given problem, then the inverse figure consists of (i) the inverses of the given points, (ii) the inverses of the points, straight lines and circles used in the construction, and (iii) the inverses of the required points.

Now using compasses only, we can pass from the given points to their inverses (i) with regard to the auxiliary circle, whose centre is O say, by means of the first lemma given below: to find A′ the inverse of A with regard to a given circle centre O. The same proof shows that we can pass to the required points themselves as soon as we have obtained the points (iii), which are their inverses. And we shall show that we can carry out each of the steps by which we pass from (i) to (ii) and from (ii) to (iii), which correspond to the steps of the original ruler and compass construction. The latter steps are of two types: to draw a straight line AB, and to describe a circle C(D). To the first of these there corresponds in the inverse figure the step: to describe the circle A′B′O, which can be done if we first find its centre by means of the second lemma given below: to find the centre of the circle through three given points.

To a step of the second type, namely to describe the circle C(D) in the original figure, there corresponds in the inverse figure the step: to describe the circle through D′ with regard to which O and C′ are inverse points, p. 93 (3). This may be reduced to depend upon the same two lemmas. Starting with C′, D′, which we suppose already obtained, by lemma (i) we construct their inverses C, D with regard to the auxiliary circle, if they are not already in the figure, and then describe the circle C(D). If C(D) meets the auxiliary circle in real and distinct points E, F, these are unaltered by the inversion, p. 92 (2), and therefore lie on the inverse of C(D), and this inverse can be described by lemma (ii) as the circle through D′EF. Or in any case, we can find as in lemma (i) the inverse O_1 of O with regard to the circle C(D), and then the inverse O_1' of O_1 with regard to the auxiliary circle. Then O_1' is the required centre of the inverse of C(D), p. 93 (3).

If the known construction which we are inverting is a Steiner construction, the original figure contains only one circle Γ, which is arbitrary, and straight lines. Then in the inverse figure, we start with the auxiliary circle centre O and an arbitrary circle Γ′, and take the inverse of Γ′ to be Γ; then every other circle of the inverse figure is determined by three points on its circumference, one of them being O. It would not in general be convenient to take Γ itself as the auxiliary circle, for some of the straight lines of the original figure probably pass through its centre, and so would not invert into circles but into straight lines.

Lemma (i). *To find* A′ *the inverse of a given point* A *with regard to a given circle* O(B).

Case I. $OA > \frac{1}{2}OB$.

Describe the circle A(O). Since its diameter 2OA is greater than the radius OB of the given circle, A(O) cuts

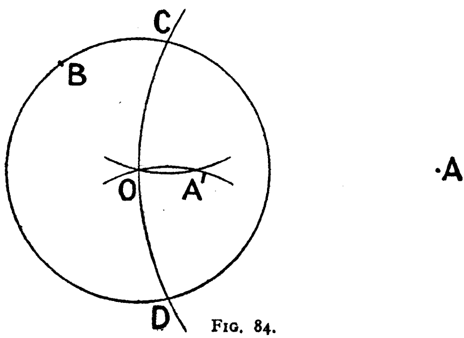

FIG. 84.

O(B) in points C, D, which are real and distinct. Describe C(O), D(O) to meet again in A′. Then A′ is the inverse of A with regard to O(B). (3 circles)

For by symmetry, OAA′ is a straight line. Also the triangles OAC, OCA′ are both isosceles by construction, and they have a common base angle at O; they are therefore similar, and

$$\frac{OA}{OC} = \frac{OC}{OA'} \quad \text{or} \quad OA \,.\, OA' = OC^2,$$

which shows that A, A′ are inverse points.

Case 2. $OA \leqq \frac{1}{2}OB$.

The construction just given fails if C, D are coincident or imaginary. We then begin by finding the point A_1 in OA produced, so that $OA_1 = 2OA$. This can be done very simply by describing first the circle A(O) and then in succession O(A), P(A), Q(A) to meet A(O) in P, Q, A_1 respectively. Then O, P, Q, A_1 are successive angular points of a regular hexagon inscribed in A(O); OA_1 is a diameter, and is twice the radius OA.

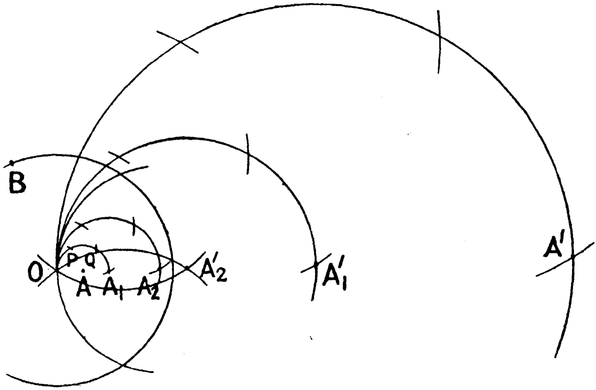

FIG. 85.

If $OA_1 > \frac{1}{2}OB$, find the inverse A_1' of A_1, by case I; then, as above, find A′ in OA_1' produced, such that $OA' = 2OA_1'$. Then A′ is the inverse of A. (11 circles)

For $\quad OA \,.\, OA' = \frac{1}{2}OA \,.\, 2OA_1' = OA_1 \,.\, OA_1' = OB^2.$

If $OA_1 \leqq \frac{1}{2}OB$, we must go on doubling, finding in succession A_2, A_3 ... A_p in OA produced, where

$$OA_2 = 2OA_1, \quad OA_3 = 2OA_2 \;\ldots\; OA_p = 2OA_{p-1} = 2^p OA,$$

until $OA_p > \frac{1}{2}OB$. We then find the inverse A_p' of A_p, and again double p times, constructing in succession the points $A'_{p-1} \dots A_2', A_1', A'$, where

$$OA' = 2OA_1' = 2^p OA_p'.$$

Then A' is the inverse of A. ($8p+3$ circles)

For $OA \, . \, OA' = 2^{-p} OA_p \, . \, 2^p OA_p' = OA_p \, . \, OA_p' = OB^2$.

The figure is drawn for $p=2$, with nineteen circles besides O(B), which is given.

Lemma (ii). *To find the centre* D *of the circle through three given points* A, B, C.

Describe A(B). Find C′ the inverse of C with regard to A(B); find D′ the reflexion of A in BC′ as axis; find D the

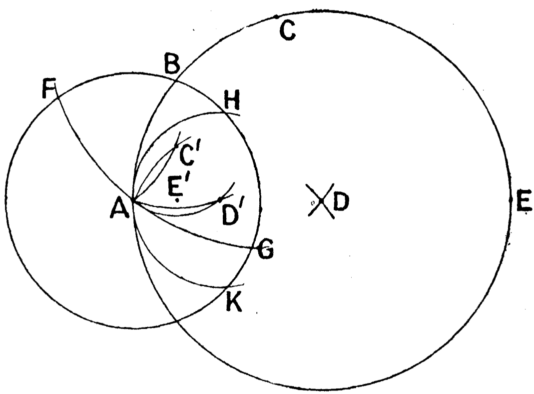

FIG. 86.

inverse of D′ with regard to A(B). Then D is the required centre of the circle ABC. (9 circles)

To prove this, let AE be the diameter through A of the circle ABC. When we invert with regard to A(B), this circle becomes a straight line, through B (which is unaltered) and C′, perpendicular to AE, so that the inverse E′ of E is the foot of the perpendicular from A on BC′, and therefore the midpoint of AD′. Now the required centre D is the midpoint of AE; hence

$$AD \, . \, AD' = \tfrac{1}{2}AE \, . \, 2AE' = AB^2,$$

so that D, D′ are inverse points with regard to A(B).

Or we may regard the straight line BC′ as a circle of infinite radius, with regard to which A, D′ are a pair of inverse points. When we invert with regard to A(B), the centre A becomes a point A′ at infinity, and D′ becomes the inverse of A′ with regard to the circle ABC, that is, its centre, p. 92 (1).

This construction is given by the table :

A(B), C(A)	F(A), G(A)	B(A), C′(A)	D′(A)	H(A), K(A)
			A(B)	
F, G	C′	D′	H, K	D

The first row gives all the circles of the figure in the order in which they are described, and the third row gives underneath each the new point or points determined by it. Where the new point is determined by intersection with a circle already described, this is shown in the second row in the same column.

Problem I. To determine the points of intersection X, Y *of a given circle* C(D) *and the straight line joining two given points* A, B.

(i) The construction suggested by the theory of inversion rests upon the fact that the required points X, Y, lying on C(D), are not altered by inversion with regard to that circle, p. 92 (2).

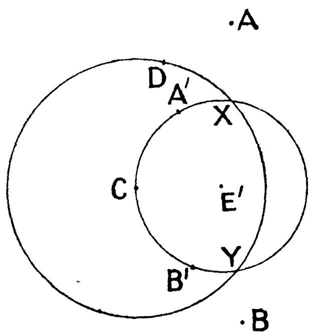

FIG. 87.

Find A′, B′ the inverses of A, B with regard to C(D). Find E′ the centre of the circle A′B′C, and describe the latter circle, which is the inverse of the straight line AB, and meets C(D) in the required points X, Y. (16 circles)

(ii) Mascheroni's construction is much shorter. The intersections of C(D) with the straight line AB are the same as its intersections with its own reflexion in AB. We therefore find C_1 the reflexion of C in AB by means of the circles A(C), B(C), and describe C_1(CD) ; this meets C(D) in the required points X, Y. (3 circles)

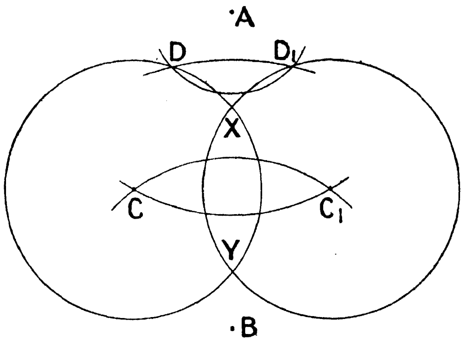

FIG. 88.

If the compasses are Euclidean, instead of describing C_1(CD), we must first find D_1 the reflexion of D in AB, and then describe $C_1(D_1)$. (5 circles)

Both these constructions fail if AB passes through the centre C ; for then the inverse of the straight line ABC is not a circle but a straight line, and cannot be drawn ; and the reflexion of C(D) in AB coincides with itself, and does not determine X, Y.

For example, if we wish to bisect a given arc DE whose centre C is given, we can find a point A on the bisecting diameter, and then we have to determine X the point of intersection of the arc DE with the diameter CX. To overcome this difficulty, complete the parallelograms DECF, EDCG, and describe F(E), G(D) to meet in A. By symmetry, A is on the bisecting diameter, and it can be proved that FX = CA. We have the construction given by the following table; the arc DE and the point C are given.

D(C), C(DE)	E(C)	F(E), G(D)	F(CA)
	C(DE)		C(D)
F	G	A	X

Problem II. To determine the point of intersection X *of the two straight lines joining two given pairs of points* A, B ; C, D.

(i) Describe any convenient auxiliary circle centre O; find the inverses A′, B′, C′, D′ of the given points; find the centres E′, F′ of the circles OA′B′, OC′D′ respectively, and describe these circles to meet again in X′; find X the

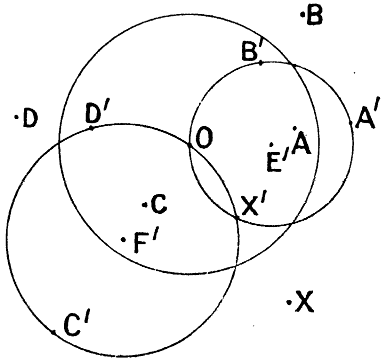

FIG. 89.

inverse of X′ with regard to the auxiliary circle. Then X is the required point of intersection of AB, CD. If carried out without precautions, this construction requires

$$1 + 4 \times 3 + 2 \times 10 + 3 = 36 \text{ circles.}$$

But this number can be very much reduced with a little care. First of all, the auxiliary circle may be taken to pass through A; then A′ coincides with A and needs no construction. Also now we need not construct B′; for in finding E′ the centre of OAB′, we require not B′, but the inverse of B′ with regard to O(A), which is B itself. In finding E′, we need only determine E the reflexion of O in AB, by means of A(O), B(O), and then E′ is the inverse of E with regard to the auxiliary circle O(A).

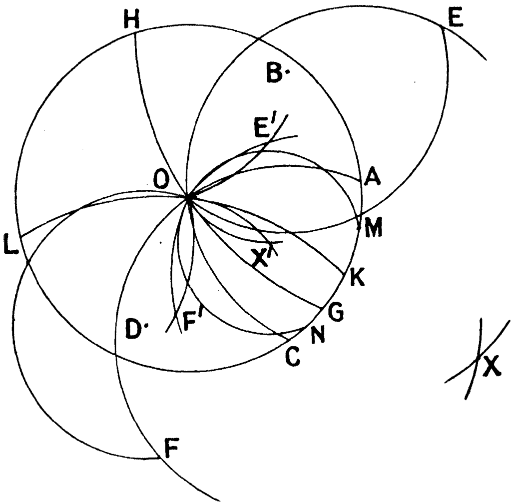

FIG. 90.

We save just as much if we also make the auxiliary circle pass through C. This is most easily done by taking O to be one of the intersections of A(C) and C(A). Then we have to begin by describing these circles, but A(C) is the same as A(O), which is used later in finding E, and similarly C(A) is used later in finding the corresponding point F.

The construction requires sixteen circles, eight of which have a radius equal to AC ; it is given by the table :

A(C) C(A)	B(O)	E(O) O(A)	G(O) H(O)	D(O)	F(O)	K(O) L(O)	E′(O) F′(O)	X′(O)	M(O) N(O)
	A(C)			C(A)	O(A)			O(A)	
O	E	G, H	E′	F	K, L	F′	X′	M, N	X

We have assumed the most favourable case for each inversion, namely that each of OE, OF, OX′ is greater than

$\frac{1}{2}$OA. The first of these conditions can always be fulfilled, by choosing O to be that one of the two intersections of A(C) and C(A) which lies furthest from the straight line AB. Then the least value of the length of the perpendicular from O to AB is $\frac{1}{2}$AC, when $\angle$BAC $= \frac{1}{2}\pi$, and the least value

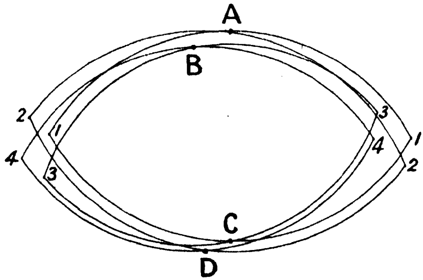

FIG. 91.

of OE is AC, so that we have OE $>\frac{1}{2}$AC. But it may happen that the position of O that is far enough from AB is too near to CD. Then instead of starting with the pair of points A, C, we may start with any other of the four pairs A, C; A, D; B, C; B, D, each of which gives two possible positions of O. But there is an unfavourable case in which all the eight positions of O fail to satisfy both the conditions OE $>\frac{1}{2}$OA and OF $>\frac{1}{2}$OA; this occurs when AB, CD are short and nearly parallel, each being inclined at an angle of nearly $\frac{1}{3}\pi$ to AC. In this case we must use case 2 of lemma (i), and the construction requires more than sixteen circles.

(ii) Mascheroni's construction is shorter if we allow the modern use of compasses, but longer if we are restricted to the Euclidean use. Find the reflexions C_1, D_1 of C, D in AB. Then the required point of intersection X of AB, CD is also the intersection of CD and C_1D_1; the triangles CXC_1, DXD_1 are similar, and

$$\frac{CX}{CD} = \frac{C_1C}{C_1C + DD_1}.$$

The length $C_1C + DD_1$ is constructed by completing the parallelogram CDD_1d, where **d** is the intersection of the circles $C(DD_1)$ and $D_1(CD)$. Then $C_1d = C_1C + Cd = C_1C + DD_1$, and we have to construct a fourth proportional to C_1d, C_1C, CD. This is done by means of similar triangles each bounded by two radii of concentric circles. Describe

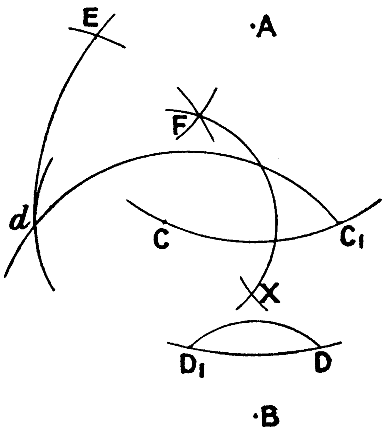

FIG. 92.

$C_1(d)$, $C_1(C)$ and place in the former a chord dE equal to CD by means of the circle $d(CD)$. The point F where the radius C_1E meets $C_1(C)$ is found as the point of contact of the latter circle and the circle $E(dC)$, or better as the point of intersection of either with $d(CE)$. Then C_1dE, C_1CF are similar isosceles triangles, and

$$\frac{CF}{dE}=\frac{C_1C}{C_1d}.$$

Hence $CF=CX=C_1X$, and X is found as the point of intersection of $C(F)$ and $C_1(CF)$.

The construction requires eleven circles as follows; three radii, namely CD, DD_1 and CF, are each used twice over.

$A(D)$ $B(D)$	$A(C)$ $D_1(CD)$	$C(DD_1)$	$C_1(d)$ $d(CD)$	$d(CE)$ $E(DD_1)$	$C(F)$ $C_1(CF)$
		$D_1(CD)$			
D_1	C_1	d	E	F	X

(iii) In adapting this construction to Euclidean compasses, I have not been able to reduce the number of circles below seventeen.

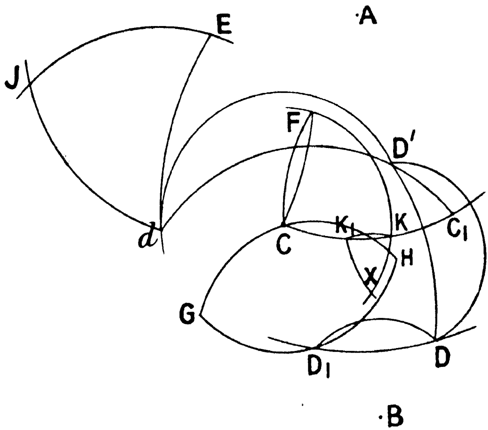

FIG. 93.

In order to construct d, we find D′, where CD′ is the reflexion of D_1D in a certain axis GH bisecting CD_1 at right angles, which has first to be determined; we can then describe C(D′), which is the same as $C(DD_1)$. The point D_1 is constructed as before, but C_1 is the intersection of the circle $D_1(D')$, just required in finding d, and A(C), which can be used again later. In order to construct F, we reflect C in an axis C_1J bisecting dE at right angles, which also bisects $\angle dC_1E$. The point J lies on d(E), which is the same as $d(D_1)$, and is already in the figure. Finally, in order to describe $C_1(CF)$, we might first reflect CF in AB, but it is shorter to reflect another radius CK of the circle C(F), so as again to make use of A(C).

The whole construction is given by the table:

A(D) B(D)	$C(D_1)$ $D_1(C)$	G(D) H(D)	C(D′) $D_1(D')$	A(C)	$d(D_1)$ $C_1(d)$	E(d)	J(C) $C_1(C)$	C(F)	B(K)	$C_1(K_1)$
				$D_1(D')$		$d(D_1)$		A(C)	A(C)	C(F)
D_1	G, H	D′	d	C_1	E	J	F	K	K_1	X

PRINTED BY ROBERT MACLEHOSE AND CO. LTD. AT THE UNIVERSITY PRESS
GLASGOW, GREAT BRITAIN