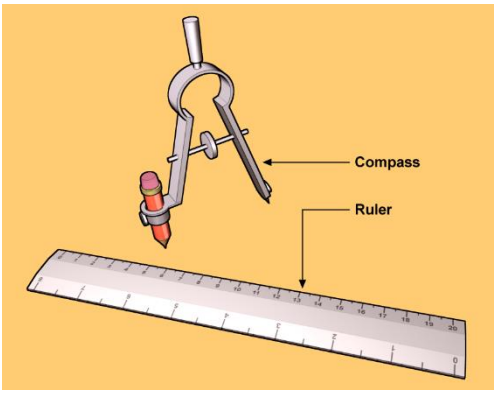
9. RULER AND COMPASS CONSTRUCTIONS

§9.1. The History of the Problem

Many geometric constructions can be carried out with just two tools – a ruler and a compass (and, of course, a sharp pencil!) The classic examples are bisection – bisection of intervals and of angles.



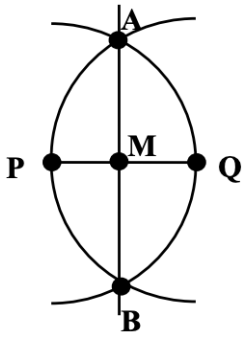
Although many ruler and compass constructions have a practical use, we're taking here a purely theoretical approach where we use idealised instruments. Our pencil is assumed to be infinitely sharp, our straight edge is a perfect straight line and our compass draws arcs of perfect circles.

A ruler and compass construction is carried out in the Euclidean Plane. We begin with a finite set of points which we call 'marked' points because they have been specifically identified. As we proceed with our constructions we mark additional points. Although our pencil produces infinitely many additional points when we draw a line, only those that occur as intersections in the construction will be considered as marked.

At each stage in the construction we'll have a finite set of marked points. We can extend the construction as follows:

- Draw the line through any two distinct marked points. (In practice this line would be a line segment, with two endpoints but in theory the line would extend indefinitely.)

- Draw the circle with one marked point as centre, passing through another marked point.



Having additional marked points in our construction we can now draw more lines and more circles and so continue the construction as far as we wish. Note that we don't have to include all possible lines or circles at each stage in our construction. We usually have a goal and we choose just what we need to construct to achieve our goal.

Example 1: Bisect the interval PQ by ruler and compass.

Solution:

(1) With centre P draw the circle passing through Q.

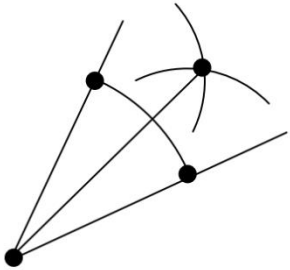
(2) With centre Q draw the circle passing through P.

Suppose that the two points of intersection of these circles are A and B.
(3) Draw AB. The point of intersection, M, of PQ and AB is the required midpoint.

The line AB is, in fact, the perpendicular bisector. The proof that PM = MQ and that AB $\perp$ PQ uses the elementary geometry of congruent triangles.

Note that ΔAPQ is equilateral, so this is also a construction for the equilateral triangle on a given base.

Example 2: Bisect an angle by ruler and compass.
Solution:



With centre at the vertex of the angle draw an arc, cutting each arm of the angle in a point. Now with these points as centres draw arcs with the same radius. Join the point of intersection of these arcs with the vertex of the angle. This will be the bisector of the angle.

The Greeks considered arithmetic and geometry as being two different ways of looking at the same number system, and geometrical constructions to perform arithmetic operations were considered very

natural. Since rational numbers were the only numbers they could conceive of arithmetically, they assumed that these are the only numbers that can be obtained geometrically.

So it came as a great shock for them to discover that $\sqrt{2}$ is irrational, that is a number that had a proper geometric existence but didn't have a proper arithmetic one. But then irrational numbers came to be accepted, though they didn't lose their geometric flavour. Number was considered to be a geometric concept. And it was accepted as an unwritten axiom that only numbers which arose in the context of geometric constructions could possibly exist.

Certain problems were posed where a construction "must exist but is hard to find". The three most famous were the "doubling of a cube", the "trisection of an angle" and the "squaring of a circle". The numbers involved in these three problems clearly exist, they thought (it's interesting that their intuition must have been using some primitive notion of continuity – a concept that took another couple of thousand years to become fully developed), so clearly there had to be a corresponding construction. And 'construction' meant a construction using ruler and compass.

Doubling the Cube:

There's a legend that, when asked for a way of stopping a plague that was attacking the city of Delos, the Oracle of Delphi advised that the altar of Apollo

should be doubled in size. The altar was in the shape of a cube and although its sides were doubled the plague continued. The Oracle then revealed that the citizens of Delos had not done as instructed since they had increased the altar eight-fold. What was required was to double the *volume*.

The problem of doubling the cube is thus: Given the side of a cube, construct the side of a cube with twice the volume.

Trisecting an Angle: Divide any given angle into three equal pieces.

For certain angles, such as 90°, it can easily be done. The above construction of an equilateral triangle gives a 60° angle and the above construction for angle bisection would give a 30° angle. But the problem is to do it for *any* given angle.

Squaring the Circle: Construct a square whose area is exactly equal to that of a given circle.

The methods allowed in all these constructions are the use of a ruler and compass. The compass is to be used to draw circles through certain points and passing through others. The ruler must be used solely as a straight-edge for joining points by straight lines, not for measurement. For this reason, ruler and compass

constructions are often called constructions by ‘straight-edge and compass’.

The reason for disallowing measurement is the question of accuracy. The accuracy with which we can measure lengths is limited by the scale of the markings. Even if we had the means to magnify the scale we’d have to end up making judgements. The whole philosophy of ruler and compass constructions is to have a procedure that is theoretically exact.

There do exist ruler and compass methods for getting quite good approximations to the solutions to all of the above three problems. But that’s not the point. With a question of existence it’s no good saying that these numbers *approximately* exist.

We need methods that are *exact*. And it’s been shown that all three problems are insoluble by ruler and compass methods. The geometric concept of number in terms of constructability thus proved to be inadequate and so the more general concept of real number gradually emerged.

The objects in a ruler and compass construction are points (denoted by A, B, C, ...) and lines (denoted by *a*, *b*, *c* ...). Lines are straight lines and circles, which can either be given, or can be drawn in the course of the construction. Other curves, such as parabolas, can only occur if they’re given at the beginning of the construction.

Here is some notation that we can use to describe a ruler and compass construction. They can be thought of as instructions in a computer program.
It is assumed that points with different labels are distinct.

LINE(A, B) = straight line through A, B.

CIRCLE(A, B) = circle with centre A, through B.

A = *p* ∩ *q* indicates that A is the point of intersection of lines *p* and *q*. If the lines are parallel we get an error message and the construction terminates.

A, B = *p* ∩ *q* indicates that A and B are the points of intersection of lines, or circles, *p* and *q* where one, or both, is a circle. Again if there fail to be two such points we get an error message. So we cannot find the point of contact of a tangent, *t*, to a circle, *c*, simply by intersecting the tangent and the circle. I suppose the reason for that restriction is that in practice it's hard to decide on the precise intersection in the case of a tangent. Or perhaps because we could insist that the tangent doesn't actually *intersect* the circle.

CHOOSE A, B, … ∈ *p* describes A, B, C, … as arbitrary distinct points on *c* (distinct from any other marked points) where *p* is a line, circle or any other curve that is given at the start of a construction. Note

that the choices can't be made mindlessly. Apart from the points being distinct from each other and from other previous points, there are other cases where a bad choice will result in the program failing.

If we were to make these constructions into a computer package we'd need to prepare error messages to prevent inappropriate choices. So we adopt the rule that if we get to a stage where we're asked to draw a line between two identical points, or a circle where the centre is the same as a point on the circle, or an intersection where there are fewer points of intersection than stated, the program gives an error message and reverts to the last CHOOSE … instruction asking the user to make another choice or, if there is no such statement, aborting the construction altogether. However implementing these programs by hand we should have enough intuition to avoid such bad choices.

So if, for example, we have circles and the program asks for two points of intersection when there is only one (touching circles) or none at all (one circle inside the other) we get such an error.

This programming language has subroutines. Having shown a program MIDPOINT for obtaining the midpoint of a line we can use it in subsequent programs.

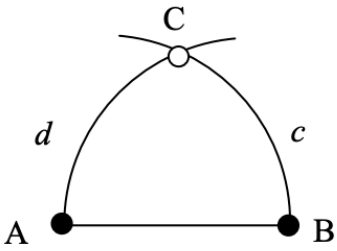
Now the validation for these constructions depends on knowing some basic facts about circles,

parabolas and ellipses. I refer you to my notes *Geometry vol 1*.

§9.2. Constructions With Lines

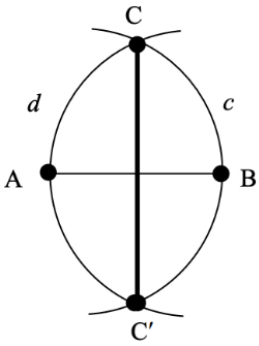
The following are some examples of ruler and compass constructions.

(1) C = EQUILATERAL(A, B) is a point that makes ABC an equilateral triangle.
c = CIRCLE(A, B)
d = CIRCLE(B, A)
C = *d* ∩ *e*



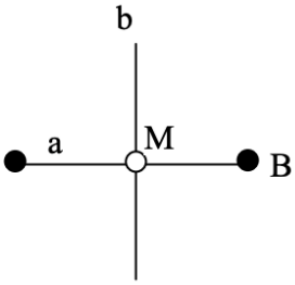
(2) *b* = BISECT(A, B) is the perpendicular bisector, *b*, of AB.

c = CIRCLE(A, B)
d = CIRCLE(B, A)
C, C′ = *c* ∩ *d*
b = LINE(C, C′)



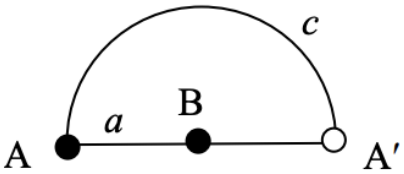
(3) M = MIDPOINT(A, B) is the midpoint of the interval AB.

a = LINE(A, B)
b = BISECT(A, B)
M = *a* ∩ *b*



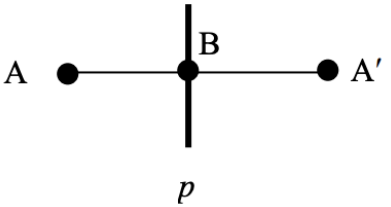
(4) A′ = REFLECT(A, B) is the reflection of A in B

a = LINE(A, B)
c = CIRCLE(B, A)
A′ = *a* ∩ *c*



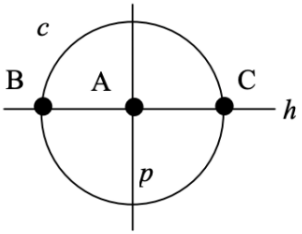
(5) *p* = PERP(A, B) is the line through B that's perpendicular to AB

A′ = REFLECT(A, B)
p = BISECT(A, A′)



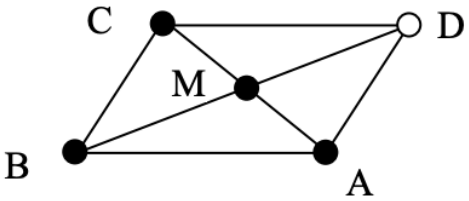
(6) *p* = PERP(A, *h*) where A ∈ *h*, is the line through A, that's perpendicular to *h*.

CHOOSE B ∈ *h*
c = CIRCLE(A, B)
C = *c* ∩ *h*
p = PERP(C, A)

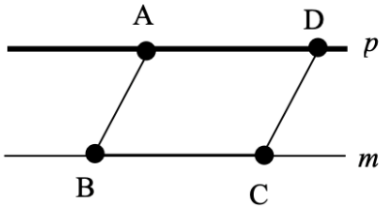


(7) D = PGRAM(A, B, C) is the point D such that ABCD is a parallelogram

M = MIDPOINT(A, C)
D = REFLECT(B, M)

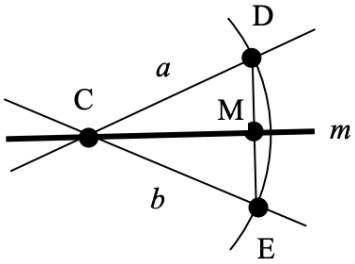


(8) *p* = PARALLEL(A, *m*) is the line through A that's parallel to *m*



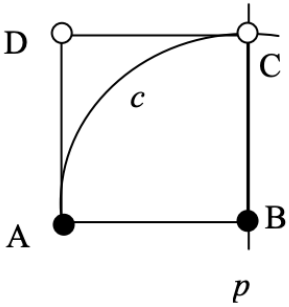
CHOOSE B, C ∈ *m*
D = PGRAM(C, B, A)
p = LINE(A, D)

(9) *m*, *n* = BISECT(*a*, *b*) are the bisectors of the angles formed between lines *a*, *b*.



C = *a* ∩ *b*
CHOOSE D ∈ *a*
c = CIRCLE(C, D)
E, F = *b* ∩ *c*
M = MIDPOINT(D, E)
m = LINE(C, M)
N = MIDPOINT(D, F)
n = LINE(C, N)

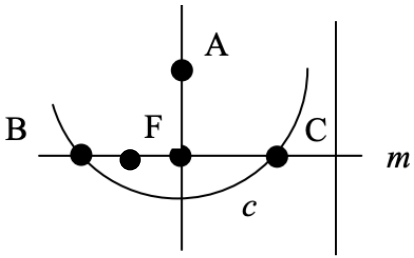
(10) C, D = SQUARE(A, B) are the pair of points C, D such that ABCD is a square.



p = PERP(A, B)
c = CIRCLE(B, A)
C = *p* ∩ *c*
D = PGRAM(A, B, C)

(11) F = FOOT(A, *m*) is the foot of the perpendicular from A to *m* where A $\notin$ *m*

CHOOSE B $\in$ *m*
c = CIRCLE(A, B)
B, C = *m* $\cap$ *c*
F = MIDPOINT(B, C)



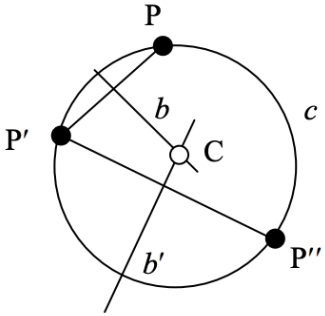
§9.3. Constructions With Circles

Any circle that we construct in the course of a Ruler and Compass Construction will already have its centre identified. But suppose we're given a circle at the outset. How do we find its centre?

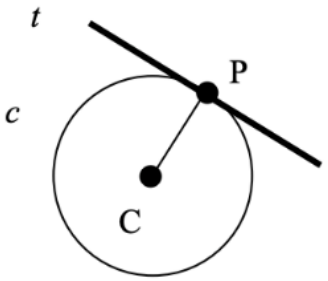
We simply take three distinct points P, P′, P″ on the circle and join P to P′ and P′ to P″. The perpendicular bisectors of PP′ and P′P″ intersect at the centre.

(12) C = CENTRE(*c*) is the centre, C, of the circle *c*

CHOOSE P, P′, P″ $\in$ *c*
b = BISECT(P, P′)
b′ = BISECT(P′, P″)
C = *b* $\cap$ *b*′



If we're given a circle and a point on it, how do we construct the tangent at that point? We construct the centre, join the point to the centre and then draw the line perpendicular to this line.

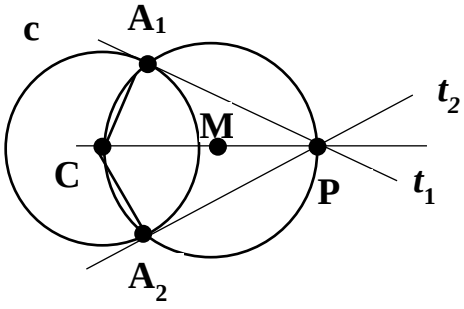


(13) *t* = TANGENT(*c*, P) is the tangent, *t*, to the circle *c* at a point P lying on the circle.

C = CENTRE(*c*)
t = PERP(C, P)

(14) t_1, t_2 = TANGENTS(*c*, P) are the two tangents to the circle *c* passing through a point P. If P lies on, or outside the circle, we get an error message.

C = CENTRE(*c*)
M = MIDPOINT(C, P)
d = CIRCLE(M, C)
A_1, A_2 = c ∩ d
t_1 = LINE(A_1, P)
t_2 = LINE(A_2, P)



§9.4. Constructions With Parabolas

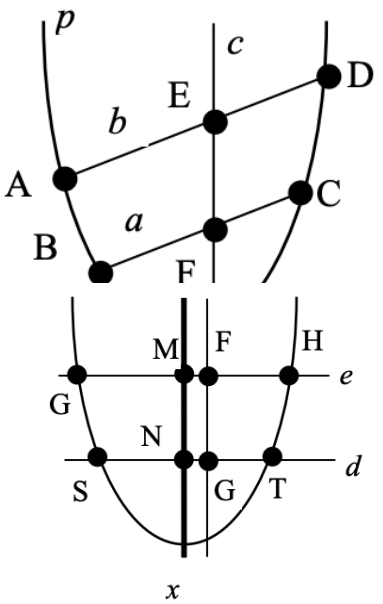
We could be given other curves at the outset, curves that can't be constructed by ruler and compass. For example we might be given a parabola and be asked

to construct its axis. Here we need to have some knowledge of the geometry of the parabola.

You may recall that the midpoints of parallel chords of a parabola lie on a line that is parallel to the axis. This leads to the following construction.

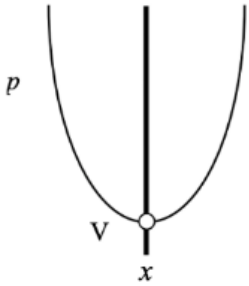
(15) *x* = AXIS(*p*) is the axis of the parabola *p*

CHOOSE A, B, C ∈ *p*
a = LINE(B, C)
b = PARALLEL(A, *a*)
D = *b* ∩ *p* so D ≠ A
E = MIDPOINT(A, D)
F = MIDPOINT(B, C)
c = LINE(F, G)
d = PERP(F, G)
e = PERP(G, F)
G, H = *e* ∩ *p*
M = MIDPOINT(H, K)
S, T = *d* ∩ *p*
N = MIDPOINT(S, T)
x = LINE(M, N)



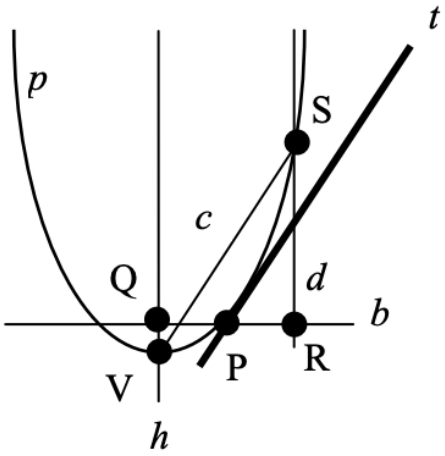
(16) V = VERTEX(*p*) is the vertex of the parabola *p*

x = AXIS(*p*)
V = *x* ∩ *p*



(17) *t* = TANGENT (P, *p*) is the tangent to the parabola *p* at the point P.

V = VERTEX(*p*)
CHOOSE P ∈ *p*
h = AXIS(*p*)
b = PERP(P, *h*)
Q = *h* ∩ *b*
R = REFLECT(Q, P)
d = PERP(R, *b*)
S = *d* ∩ *p*
c = LINE(V, S)
t = PARALLEL(P, *c*)



Why does this construct work?
We may take the parabola to have the equation $x^2 = 4ay$ in the *x*-*y* plane. Then V = (0, 0).

We may take P = $(2at, at^2)$ where $t \neq 0$;
h is $x = 0$; *b* is $y = at^2$; Q = $(0, at^2)$;
R = $(4at, at^2)$; *d* is the line $x = 4at$; S = $(4at, 4at^2)$;
c has slope *t*, which is the slope of the tangent at P.

§9.5. Constructions With Ellipses

Again, although we can't construct ellipses by ruler and compass, we can be given an ellipse and be asked to make certain constructions upon it by ruler and compass. We begin with the problem of locating the centre of the ellipse. As with circles we do this with

midpoints of chords. But with ellipses we need four chords that are parallel in pairs. And keep in mind that perpendiculars have no significance with ellipses.

Now the first step will be to choose 6 points on the ellipse.

(18) O = CENTRE(*e*)
CHOOSE A_1, A_2 B_1, B_2, C_1, $C_2 \in e$
h_1 = LINE(A_1, B_1)
k_1 = PARALLEL(C_1, h_1)
C_1, D_1 = $k_1 \cap e$
M_1 = MIDPOINT(A_1, B_1)
N_1 = MIDPOINT(C_1, D_1)
p_1 = LINE(M_1, N_1)
h_2 = LINE(A_2, B_2)
k_2 = PARALLEL(C_2, h_1)
C_2, D_2 = $k_2 \cap e$
M_2 = MIDPOINT(A_2, B_2)
N_2 = MIDPOINT(C_2, D_2)
p_2 = LINE(M_2, N_2)
O = $p_1 \cap p_2$

Next we want to be able to locate the foci. This isn't difficult if we have the major and minor axis. And, once we have one of these axes it's child's play to find the other. But how do you locate one of the axes to begin with? You just take a circle with the same centre as the ellipse and intersect it with the ellipse. Ideally this would

give four points and these would form the corners of a rectangle. The lines joining the midpoints of opposite sides will be the axes.

Of course the circle may lie completely inside the ellipse, or vice versa, but we can ensure that they do intersect by choosing a point on the ellipse and drawing a circle that passes through these points. It's possible that the circle and ellipse are tangential in which case there would only be two points of intersection. In this case we get an error message and the user is asked to choose again.

(19) a_1, a_2 = AXES(*e*)
CHOOSE P $\in$ *e*
O = CENTRE(*e*)
c = CIRCLE(O, P)
A, B, C, D = $c \cap e$
M_1 = MIDPOINT(A, D)
M_2 = MIDPOINT(B, C)
M_3 = MIDPOINT(A, B)
M_4 = MIDPOINT(C, D)
a_1 = LINE(M_1, M_2)
a_2 = LINE(M_3, M_4)

This construction doesn't tell us which is the major axis and which is the minor one, but that doesn't matter.

Now we want to find the foci. For the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, the foci are $(\pm ae, 0)$ where $b^2 = a^2(1 - e^2)$. So $a^2 = (ae)^2 + b^2$. So we draw a circle with centre (0, b) and radius a. It cuts the major axis in one of the foci.

(20) F_1, F_2 = FOCI(e)
a_1, a_2 = AXES(e)
$A_1 = a_1 \cap e$
$A_2 = a_2 \cap e$
D = PGRAM(A_1, O, A_2)
c = CIRCLE(A_1, D)
F_1, $F_2 = c \cap e$

In implementing this program we should take a_1 to be the minor axis and a_2 the major one. If the computer got them the wrong way round we'd have no points of intersection in the last step and we'd be sent back to the first step and be asked to make a swap. Again, if we are doing this by hand we'd obviously make the correct choice to begin with.

(21) *n*, *m* = NORMAL(P, *e*) are the normal and tangent to *e* at P
F_1, F_2 = FOCI(e)
a_1 = LINE(F_1, P)
a_2 = LINE(F_2, P)
n, m = BISECT(a_1, a_2)

BISECT gives two perpendicular angle bisections. One is the normal and the other is the tangent.

§9.6. Construction of an Ellipse

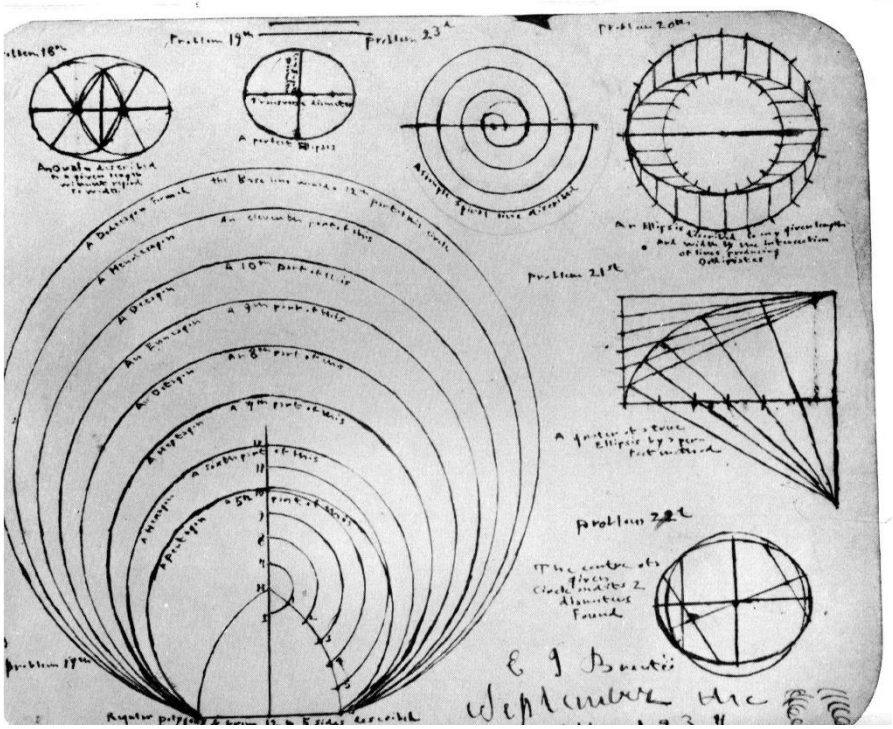
Clearly one can't draw a complete ellipse by ruler and compass. But we can construct any finite number of points on an ellipse. In the nineteenth century such methods began to have a very practical use for artists.

In the real world circles are very common, but in perspective on a canvas these usually appear as ellipses. In Victorian art, it was considered a very necessary part of the training of an artist to learn how to draw ellipses. Drawing manuals contained instructions on how to do this.

The author of *Wuthering Heights*, Emily Brontë, left behind very few manuscripts. The manuscript of her famous novel, for example, has been lost. All that remain are a few scraps of 'diary papers' and this one page with geometric drawings.

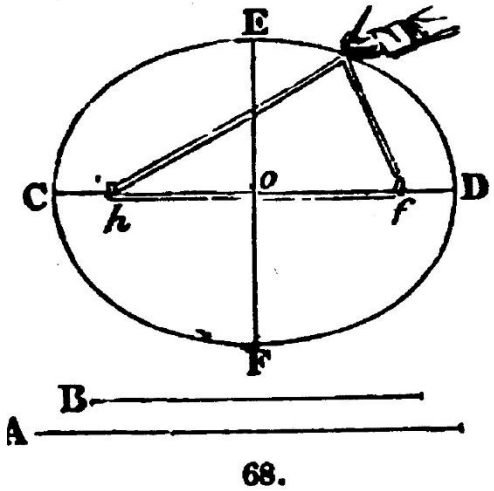
This page consisted of her working through some drawing exercises, taken from the 5th edition of *An Introduction to Perspective, Practical Geometry, Drawing and Painting* by Charles Hayter in 1832.

Problem 18 is a construction of an oval, which is only an approximation to an ellipse, so we will pass over it. Problem 19 appears to be the standard construction of an ellipse using string and pins.



You place the paper on a board and mark the points where you want the foci to go. Then you stick pins through the paper at these points and wrap a loop of string around the pins. The length of the string should be somewhat longer than twice the distance between the pins.

Place a pencil inside the loop, and move it across the paper, keeping the string taut. This will trace out a true ellipse (or ellipsis as Emily called it). This construction is based on the fact that the sum of the distances of any point on an ellipse to each of the foci is constant – in fact it is equal to the length of the major axis. This is proved in my notes on the ellipse in *Geometry vol 1*. By varying the length of the string one can obtain differently shaped ellipse with the same foci.



68.

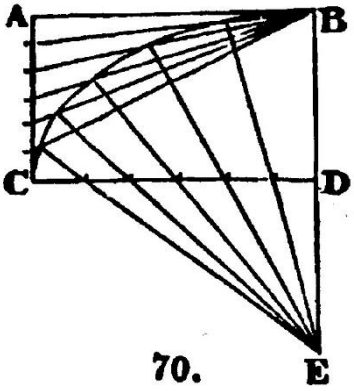
Emily's manuscript is in the Brontë Parsonage Museum in Haworth in Yorkshire. I once examined the page and held it up to the light. Sure enough there were faint pin-holes at the foci in this diagram.

Problem 20 constructs an ellipse from its auxiliary circle, using 24 equally spaced points around the circle. This would require a 15° angle to be constructed, which

can be done by bisecting a 60° angle twice. There is no evidence that Emily carried out such a construction. Probably she merely divided a 90° angle into 6 equal parts by eye. You can probably work out the method used.

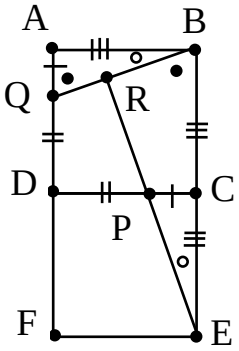
Problem 21 is interesting. It is a true ruler and compass construction, but rather than constructing the whole ellipse it constructs any finite number of points on it. The ellipse has to be completed by eye. Also, rather than constructing the whole ellipse it focusses on one quadrant.

You construct a rectangle ABCD. You then divide adjacent sides AC and CD into a certain number of equal parts – the same number for each side. Emily followed Hayter in using 6 parts. Actually, rather than constructing the rectangle first, one can decide on the length of each part and just replicate this length with the compasses. This is probably what Emily did. Here is a reproduction of the diagram from Hayeter's book, which is a little clearer than in Emily's manuscript.



Then you take a point E on CD produced so that CE = BC and join E to the internal points on CD and B to the internal points on AC. The intersections of corresponding lines will lie on an ellipse.

Theorem 1: Let ΔABCD and ΔDCEF be squares.
Let P lie on CD and Q lie on AD such that DP = DQ.
Then R = BQ ∩ EP lies on the circle with centre C, passing through B and D.



Proof: ΔABQ ≡ ΔCEP (2 sides and right angle).
Hence ∠ABQ = ∠CEP.
∠AQB = ∠QBC (alternating angles).
Hence ΔERB is right-angled at R.
∴ R lies on the circle with centre C, passing through B, E.

Theorem 2: Let $\mathcal{P}$(ABCD) and $\mathcal{P}$(DCEF) be rectangles with AB = *k*AD.
Let P lie on CD and Q lie on AD such that DP = *k*DQ.
Then R = BQ ∩ EP lies on the ellipse with centre C, passing through B and D.
Proof: We apply the linear transformation

$$(x, y) \to (kx, y).$$

Of course, if k ≠ 1, ∠ERB will not be a right-angle, but R will lie on the image of the circle, which will be an ellipse.

In Problem 22 we are given an ellipse and are required to find the centre and the diameters. Emily's notation is *The centre of a given Circle and its two diameters Found*. It is not clear why she mentions a circle when the method works for any ellipse, including a circle as a special case. Hayter describes the method as applying to any 'oval'.

The method depends on the fact that for any ellipse (including a circle) the midpoints of parallel chords pass through the centre and the midpoint of any chord through the centre will be the centre.

One then draws a circle, with its centre being the centre of the ellipse and any radius that causes the circle to intersect the ellipse in four points. These points will be the vertices of a rectangle whose sides are parallel to the axes of the ellipse, and so the axes themselves can be found.

§9.7. Impossible Constructions

There are limitations as to what can be constructed by ruler and compass. For example a line can be divided into any finite number of equal parts. However, although any angle can be bisected by ruler and compass, not all angles can be trisected.

But how can we possibly say what can't be done? Might not some very clever mathematician at some stage in the future come up with such a construction?

Impossibility proofs pervade a lot of branches of mathematics. Mathematicians seem to take great

pleasure in proving that such and such cannot be done. But they should not be thought of as negative 'spoilsports'. Instead they are great time-savers. Think of how much time can be saved by preventing other mathematicians wasting time on a project which cannot succeed.

To prove that something is impossible we use Proof By Contradiction. We suppose that it is possible, and prove that if this is so then something else is possible – something which is blatantly impossible.

Theorem 3: There is no ruler and compass construction that can trisect any given angle.

Proof: If any angle can be trisected by ruler and compass then a 20° angle can be constructed in this way, and hence a line of length $2\cos 20°$ can easily be constructed.

Let $c = \cos(2\pi/9)$ and $s = \sin(2\pi/9)$.

Then $(c + is)^3 = c^3 + 3ic^2s - 3cs^2 - is^3$.

But by De Moivre's Theorem

$$(c + is)^3 = \cos(2\pi/3) + i \sin(2\pi/3) = -\frac{1}{2} + \frac{\sqrt{3}}{2} i.$$

Equating real parts we get $c^3 - 3c(1 - c^2) = -\frac{1}{2}$ and so

$$8c^3 - 6c + 1 = 0,$$

That is, $(2c)^3 - 3c + 1 = 0$.

Hence $2\cos 20°$ is a zero of the cubic $x^3 - 3x + 1$.

It is easy to show that this is a polynomial of smallest degree that has $2c$ as a zero, because if it was a

zero of a quadratic, the cubic would have to factorise over the rational numbers.

Never mind the details here. I'm just giving you an overview. The techniques required to do this properly involve field extensions. If you're really interested in the full story you can find it in my notes on *Galois Theory*.

Now associated with any complex number that is a zero of some polynomial with rational coefficients is a number called the **degree**. It is in fact the smallest degree for a polynomial that has that number as one of its zeros. So the number $2\cos 20°$ has degree 3. So what's impossible about that?

It is shown in my Galois theory notes that the degree of any number that can be constructed by ruler and compass, starting with points with rational coordinates is a power of 2. Whenever we intersect a line with a circle, or two circles, the degree of the points of intersection at most double their degree (the degree might stay the same). So with repeated steps in a ruler and compass construction the degree must be a power of 2.

And the simple fact that would have to be true if trisection was possible by ruler and compass, is that 3 is a power of 2. This is blatantly false and so the assumption that any angle can be trisected in this way must therefore be false.

There are three unsolved problems that date from antiquity.

(1) To trisect any angle by ruler and compass.

(2) To double a cube.
There's a legend that when asked for a way to stop a plague that was attacking the city of Delos, the Oracle of Delphi advised that the altar of Apollo should be doubled in size. The altar was in the shape of a cube and although its sides were doubled the plague continued. The Oracle then revealed that the citizens of Delos hadn't done as instructed since they had increased the altar eight-fold. What was required was to double the *volume*. The problem of doubling the cube is thus to construct a line interval of length $\sqrt[3]{2}$.

(3) To square a circle.
This meant to construct a square whose area is that of a unit circle – in other words, to construct a line of length $\sqrt{\pi}$.

All three of these remained unsolved for a simple reason. They can all be shown to be impossible! I have outlined the proof that problem (1) is impossible because 2cos 20° has degree 3 and 3 is not a power of 2.

Problem (2) is impossible because $\sqrt[3]{2}$ has degree 3. This number is a zero of the rational polynomial x^3 –

2, but nothing of lower degree. Again this comes down to the fact that 3 is not a power of 2.

Problem (3) is a little trickier. It is possible to show that $\sqrt{\pi}$ doesn't have a degree at all, because there is no polynomial of any degree that has rational coefficients and has $\sqrt{\pi}$ as a zero. So if it doesn't have a degree at all, it can't have a degree that's a power of 2.

§9.8. Regular Polygons

If you look closely at Emily Brontë's page of constructions you will notice a large diagram that appears to suggest that there are ruler and compass constructions for constructing regular polygons with anything from 5 to 12 sides. She has left out 3 and 4 sides because it's obvious that equilateral triangles and squares can be constructed.

In fact this implication is false. It is true that a regular pentagon can be constructed. So can a regular decagon. And a regular octagon can be constructed by repeatedly bisecting a 90° angle. But all the rest on that page are non-constructible.

So which regular polygons can be ruler and compass constructed? The answer is contained in the following theorem. But first you need to know what is a Fermat prime.

A **Fermat prime** is a prime number, *n*, that is a power of 2 plus 1. It can be shown that $2^n + 1$ is prime if

and only if n itself is a power of 2. So in fact a Fermat prime has the form $2^{2^n} + 1$. So they are very rare.

Now $5 = 2^2 + 1$ is prime, and $17 = 2^4 + 1$ is also prime. Also $257 = 2^8 + 1$ is prime and so is $65537 = 2^{16} + 1$. Fermat knew all these and he conjectured that $2^{2^n} + 1$ is prime for all *n*. He would have done better if he'd conjectured that these are the only Fermat primes. They are certainly the only ones that are known.

Theorem 4: A regular *n*-gon can be constructed by ruler and compass if and only if *n* is a power of 2, or a power of 2 times a product of distinct Fermat primes.
Proof: I omit the proof. But you can see why there can be any number of factors of 2. If you can construct a regular polygon the you can construct one with twice as many sides, because all you have to do is to bisect a certain angle.

So the only regular *n*-gons, with up to a hundred sides, that are constructible, are those where
n = 3, 4, 5, 8, 10, 16, 17, 20, 32, 34, 40, 64, 68, 80.

§9.9. String and Pins Constructions

In the early 1980s I wondered what could be constructed using string and pins as well as a ruler and compass. Of course we could dispense with the compass because a string and one pin can be used to draw circles, and even the ruler if the string can be made sufficiently

taut! But when I talk of string and pins constructions I will allow the use of a ruler and compass for convenience.

As well as being able to draw an ellipse I wondered whether it was possible to construct things that were impossible using ruler and compass using this additional equipment. I convinced myself that:
(1) Any angle can be trisected by string and pins.
(2) Not every angle can be penta-sected (divided exactly into five equal pieces) by string and pins.
(3) Every field extension, of the rationals, that arises in a string and pins construction must have the form 2^a3^b.

I never published my proofs and so it is quite possible they were flawed. But my colleague, Gerry Myerson, recently drew my attention to a paper by Aliska Gibbons and Lawrence Smolinsky *Geometric Constructions With Ellipses*, published in 2008. Because I never published this was a genuinely independent piece of work, but it goes to show how similar ideas can float around the ether and be picked up independently by different mathematicians around the same time. At the risk of sounding like I'm a Newton, something similar happened with the discovery of Calculus!

I'm indebted to the above authors because I now have proofs that I can include in my notes. I am also indebted to them for telling me that some of these sort of constructions date back to classical times. Pappus wrote of angle trisection constructions using hyperbolas.

Descartes in his 1637 treatise *La Geometrie* showed how an angle can be constructed with a parabola and a circle. I don't know of any string and pins construction for hyperbolae or parabolas, so constructions using ellipses are particular interesting.

In 1997 Carlos Videla explored the concept of **conic constructability** where, in addition to a ruler and compass, one has a method for drawing any conic. He showed that such constructions could be accomplished using only parabolas and hyperbolas. Gibbins and Smolinsky showed that all conic constructions could be achieved using only ellipses, and therefore were string and pins constructable.

The techniques for the proofs are based on field extensions, and it is more appropriate to include them in my *Galois Theory* notes. I expect to include them in the next edition of those notes.

EXERCISES FOR CHAPTER 9

Exercise 1: Describe a ruler and compass construction from scratch to draw the chord of contact of tangents from the point P, lying outside the circle *c*.

Exercise 2: Find ruler and compass constructions for the inscribed circle for a triangle ABC.

Exercise 3: Find a ruler and compass construction to construct the common tangents (where they exist) to two given circles of different radii.

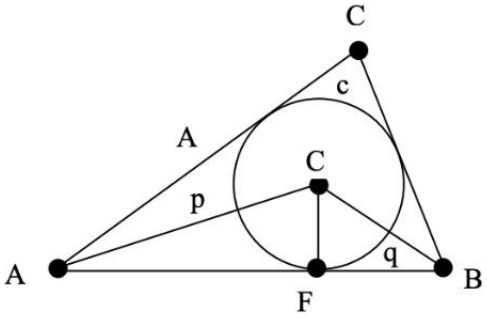
Exercise 4: If the point A does not lie on the line *m*, describe a construction for FOOT(A, *m*) which will work irrespective of what point on *m* we choose for P. [See construction (11) above.]

SOLUTIONS FOR CHAPTER 9

Exercise 1: *h* = CHORD(*c*, P)
Combining constructions (12) and (14):
CHOOSE P, P′, P″ $\in c$
b = BISECT(P, P′)
b' = BISECT(P′, P″)
C = $b \cap b'$
M = MIDPT(C, P)
d = CIRCLE(M, C)
A_1, A_2 = c $\cap$ d
h = LINE(A_1, A_2)

Exercise 2:
The centre of the inscribed circle is the common intersection of the perpendicular bisectors of the sides.
p = BISECT(C, A, B)
q = BISECT(A, B, C)
C = $p \cap q$

F = FOOT(C, A, B)
c = CIRCLE(C, F)



Exercise 3:

Let the circles be CIRCLE(C_1, P_1) of radius r_1 and CIRCLE(C_2, P_2) of radius r_2 where $r_1 > r_2$.

Construct CIRCLE(C_1, R) so that it has radius $r_1 - r_2$.

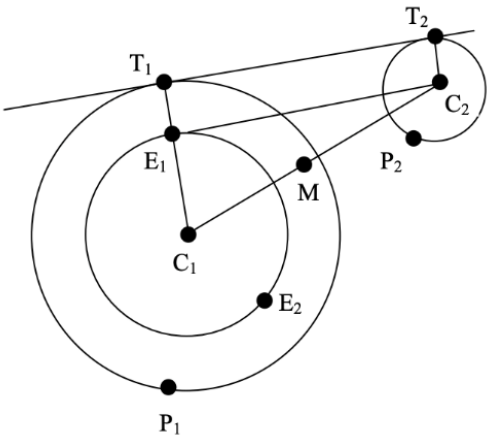
Construct M, the midpoint of C_1, C_2.

E_1, E_2 = CIRCLE(M, C_1) $\cap$ CIRCLE(C_1, R).

T_1, T_2 = LINE(E_1, C_1) $\cap$ CIRCLE(C_1, P_1).

Construct the line LINE(T_1, X) perpendicular to LINE(C_1, E_1).

This is one of the four common tangents.



Exercise 4:

All we need is a construction that will be guaranteed to result in a point on B on *m* which is *not* to be the foot of the perpendicular from A to *m*.

CHOOSE P $\in$ *m*

c = CIRCLE(P, A)

B $\in$ *c* $\cap$ *m*

No matter what point we choose for P, the foot of the perpendicular from A to *m* will have to lie *inside* the circle and so neither choice for B can be the foot of the perpendicular. We then proceed as in (11).

coopersnotes.net
LIST OF TITLES

GENERAL
- The Mathematics At The Edge Of The Rational Universe

ELEMENTARY
- Basic Mathematics
- Concepts of Algebra
- Concepts of Calculus

1st YEAR UNI
- Techniques of Algebra
- Techniques of Calculus
- Matrices

2nd YEAR UNI
- Linear Algebra
- Languages & Machines
- Discrete Mathematics

3rd YEAR UNI
- Group Theory vol 1
- Group Theory vol 2
- Galois Theory
- Geometry vol 1
- Geometry vol 2
- Topology
- Set Theory
- Number Theory
- Graph Theory
- Complex Variables

POSTGRADUATE
- Ring Theory
- Representation Theory
- Quadratic Forms
- Group Tables vol 1
- Group Tables vol 2